University of Hertfordshire

School of Engineering and Computer Science

Adaptive Systems Research Group

Investigating human perceptions of trust and social cues in robots for safe Human-Robot Interaction in human-oriented environments

submitted to the University of Hertfordshire in partial fulfilment of the requirements of the degree of
Doctor of Philosophy

Principal Supervisor
Prof. Kerstin DAUTENHAHN

Co-Supervisors
Dr. Kheng Lee KOAY
Dr. Michael L. WALTERS

PhD Student
Alessandra Rossi
15058449

July 2020

A voi,
che mi avete sempre incoraggiata a volare alto.

For the children who never tire of moon gazing upon the dock,
by the light of the fireflies, till the angels are dispatched by Nana,
to sprinkle sleepy dust in their eyelashes so long and fine
Were The World Mine - Tanner Cohen

I am a star ...

I am a star in the firmament
that observe the world, despises the world
and consumed in its heat.
I am the sea by night in a storm
the sea shouting that accumulates new sins
and to the ancient makes recompense.
I am exiled from your world
of pride polite, by pride defrauded,
I am the king without crown.
I am the passion without words
without stones of the hearth, without weapons in the war,
is my same force that make me sick.

Sono una stella del firmamento
che osserva il mondo, disprezza il mondo
e si consuma del proprio ardore.
Sono il mare che di notte si infuria,
il mare urlante che accumula nuovi peccati
e agli antichi rende mercede.
Sono dal vostro mondo esiliato
di superbia educato, dalla superbia frodato,
io sono il re senza corona.
Sono la passione senza parole
senza pietre nel focolare, senz'arma nella guerra,
è la mia stessa forza che mi ammala.

Hermann Hesse

Acknowledgments

I would like to thank my supervisor Prof. Kerstin Dautenhahn for providing enthusiastic support and great advice. Since the beginning she has been a very inspirational model offering her guidance and experience for letting me grow as researcher.

I am also very grateful for the amazing and relentless support received by Dr. Kheng Lee Koay. He has been always ready to help, discuss and advice me regarding all the aspects of my research. He became a very good friend in which, I know, I can always count in ... but he will probably deny it.

I would also like to thank Dr. Mick Walters for his feedback and comments to the work presented in this thesis. He also has been very patience with me and my itaglish, as Kheng Lee likes to call it.
My supervisors taught me a lot, and this would not have been possible without them!

I am very grateful to Prof. Silvia Rossi. She has been my professor of Robotics during my MSc and supervisor for my MSc final project. Then, she became my mentor and, finally, friend. She always pushed to be the best of me, supporting me and providing constant advice. She guided me, and she is keeping guiding me through my journey in the research.

I am very grateful for the moral and technical support that I received by my friend Dr. Patrick Holthaus.

I would also like to thank Dr. Joe Saunders for his support and input at the beginning of this work.

I am also very grateful for being part of an amazing training project such as SECURE. The PIs, colleagues and learning opportunities opened for me a new world full of wonders.

I am very grateful to the staff of University of Hertfordshire for their support and help. In particular, I would like to thank Mariana Lilley for being such a lovely and supportive person. Thank you, Simon, Volker.

I would also like to thank my friends and colleague from the Adaptive Systems Research Group. They made me feel welcome and have always been supportive. They have been always ready to help me and discuss with me about research, interests. Thank you, Frank, Gabi, Luke, Ben, Marcus, Greg, Azeemsha, Udeshika, Karen, Bente. Thank you to Dag for patiently advising me on statistics.

A special thank goes to my friend and colleague Silvia Moros. She made my life happier by introducing me to the aerial circus. Hanging upside down became my happy place.

I would also particularly thank Dan Barry who has been a pain and an amazing friend since the moment I met him (but I am sure he would say the same). Thank you for always being there for me.

Thank you, Marcus, for introducing me to the RoboCup. It became constant part of my life. I would also like to thank my team Bold Hearts for believing in me as their team leader. I would also like to thank Prof. Daniel Polani for his trust and support he gave me since I joined the team.

I would also like to thank my friends for baring with me, my craziness and stress. I could not make it without you. Thank you, Janmarie, David, Natalie, Andrei, Agne, Samir, Pahn, Julie-anne. A special thank you to Maria.

A warm thank you to my family and friends. My cousins Manuela, Laura and Chiara always encouraged me to be strong, happy and smiling.

I think I should include my sweet dog, Yuri, that never gets mad at me and always unconditionally loves me.

I want to dedicate this thesis to the most important people of my life, my mom, my brother Roberto, my lovely nephew Christian, and my dad who is always watching me.

Abstract

As robots increasingly take part in daily living activities, humans will have to interact with them in domestic and other human-oriented environments. This thesis envisages a future where autonomous robots could be used as home companions to assist and collaborate with their human partners in unstructured environments without the support of any roboticist or expert. To realise such a vision, it is important to identify which factors (e.g. trust, participants' personalities and background etc.) that influence people to accept robots' as companions and trust the robots to look after their well-being. I am particularly interested in the possibility of robots using social behaviours and natural communications as a repair mechanism to positively influence humans' sense of trust and companionship towards the robots. The main reason being that trust can change over time due to different factors (e.g. perceived erroneous robot behaviours). In this thesis, I provide guidelines for a robot to regain human trust by adopting certain human-like behaviours. I can expect that domestic robots will exhibit occasional mechanical, programming or functional errors, as occurs with any other electrical consumer devices. For example, these might include software errors, dropping objects due to gripper malfunctions, picking up the wrong object or showing faulty navigational skills due to unclear camera images or noisy laser scanner data respectively. It is therefore important for a domestic robot to have acceptable interactive behaviour when exhibiting and recovering from an error situation. In this context, several open questions need to be addressed regarding both individuals' perceptions of the errors and robots, and the effects of these on people's trust in robots.

As a first step, I investigated how the severity of the consequences and the timing of a robot's different types of erroneous behaviours during an interaction may have different impact on users' attitudes towards a domestic robot. I concluded that there is a correlation between the magnitude of an error performed by the robot and the corresponding loss of trust of the human in the robot. In particular, people's trust was strongly affected by robot errors that had severe consequences.

This led us to investigate whether people's awareness of robots' functionalities may affect their trust in a robot. I found that people's acceptance and trust in the robot may be affected by their knowledge of the robot's capabilities and its limitations differently according the participants' age and the robot's embodiment.
In order to deploy robots in the wild, strategies for mitigating and re-gaining people's trust in robots in case of errors needs to be implemented. In the following three studies, I assessed if a robot with awareness of human social conventions would increase people's trust in the robot. My findings showed that people almost blindly trusted a social and a non-social robot in scenarios with non-severe error consequences. In contrast, people that interacted with a social robot did not trust its suggestions in a scenario with a higher risk outcome.
Finally, I investigated the effects of robots' errors on people's trust of a robot over time. The findings showed that participants' judgement of a robot is formed during the first stage of their interaction. Therefore, people are more inclined to lose trust in a robot if it makes big errors at the beginning of the interaction.
The findings from the Human-Robot Interaction experiments presented in this thesis will contribute to an advanced understanding of the trust dynamics between humans and robots for a long-lasting and successful collaboration.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

All the world is made of faith, and
trust, and pixie dust.

J.M. Barrie, Peter Pan

Robots are beginning to take part in our lives in many different ways, and the increasing presence of humanoid and human-friendly robots in daily living activities requires the development of more natural Human-Robot Interaction (HRI).
Indeed, autonomous robots are more and more used to assist older people (Montemerlo et al. 2002; Pollack 2005; Wada et al. 2002) or to interact with children with autism, or assist humans in daily living activities in household situations, such as answering questions, providing information of delivery news, turning on the light, as for support in a kitchen (Stiefelhagen et al. 2004).
This prospect opened two main challenges for consideration: people will need to be able to accept the presence of the robot in their living space and they will also need to be able to trust that their robotic companion will look after their well-being without compromising their safety. Indeed, it is really important that humans trust their robot companion to not create a hazardous situation, such as start a fire when trying to make a cup of tea, or an unsafe situation, such as leaving the door open unattended or open the door to strangers - and potential thieves. Humans should be able to wake up in the middle of the night to get a cup of water from the kitchen without worrying about stumbling into the robot and get injured.
Nevertheless, trust is a very complex feeling even between humans, and it can be fostered and enhanced according different factors. In order to identify these factors and create mitigation in case of a lack of trust, this research has been carried out considering different assumptions.

Firstly, HRI researches established that trust has a key role for an effective collaboration between people and robots (Ross 2008). Indeed, trust affects the humans perception of the usefulness of the information and capabilities provided by the robots (Cameron et al. 2015). However, robots can have erroneous behaviours (mechanical, software or perceived), and they, consequently, can affect negatively people's trust in them.
Secondly, people's perception of a robot is affected by several factors, including individuals' differences (or antecedents of trust). The antecedents of trust are characterised by people's background, intelligence, personality traits, skills and aptitudes are recognised as the most relevant (Williamson 2018).
Finally, a natural interaction with a robot permits a greater social acceptance by humans (Kerstin Dautenhahn 2007a; Kerstin Dautenhahn 2007b). Human interactions are based upon natural language (NL), gesture, pose and body language. Natural interaction and communication are complicated mechanisms, however humans easily learn and master the proper linguistic and social skills necessary. A robot should be able to learn and perform these humans modes in order both to provide physical, mental and social support to people, and to support and improve collaborative and cooperative teamwork.
Therefore, these assumptions are used in this thesis to:

- Investigate how people's trust in a robot changes due to robot erroneous behaviours. In particular, considering that errors can have different consequences, and therefore, they might affect people's trust in a robot in different ways;
- Identify which antecedents of trust affect people's trust in robots. In particular, I investigated whether people's trust in a robot that sometimes makes errors is influenced differently according to the individuals' differences;
- Mitigate a loss of trust in a robot due to robot erroneous behaviours by designing robot's behaviours that are perceived as natural and social.

1.1 Research Problem

In the not too distant future, autonomous robots will take part in people's daily living activities. We expect that robots will support humans in everyday tasks in healthcare facilities and hospitals, they will be companions in private homes, assist people in shopping malls, or deliver goods to customers. For example, a home companion robot might provide affective and social aid for people's with dementia (Liang et al. 2017), autonomous robots can be good exercise partners for socially anxious individuals (Yamada et al. 2019), they can also help cover wide areas to securely control the borders of a country (Alex et al. 2017; Penghou et al. 2019), or as shopping assistant in a galleria (Sidorov et al. 2019).

These interactions are intended to be developed to function without the intervention of roboticists and expert supervisors. Specifically, they are going to take place in unstructured and unpredictable environments. Robots are machines and they might exhibit occasional mechanical or functional errors. For example, the robot may turn off during a delicate task because its battery was fully discharged without warning, or a robot might unlock the front door to strangers who may be potential thieves. It might not be able to detect an obstacle and tip over, or the arm of the robot might accidentally knock over a porcelain vale while executing a task. In such scenarios, several interaction dynamics can arise: people might lose trust and not rely anymore in their assistance (Steinfeld et al. 2006; E. D. Visser et al. 2006); they might lose interest in robots and toss them away (M. d. Graaf et al. 2017); or they might overtrust their capabilities (Booth et al. 2017; Robinette, W. Li, et al. 2016) with possible fatal consequences.

It is therefore important to design and develop a mechanism to mitigate people's trust in a service robot, in particular in domestic environments. A positive balanced trust, indeed, is fundamental for building a high quality interaction (Atkinson et al. 2014). Therefore, companion robots should have acceptable interactive behaviours when exhibiting and recovering from an error situation.

Most researchers are currently focusing their attention on robot errors of similar nature, whereas each of the mentioned robot errors might be perceived differently according to the resultant consequences. In order to define these behaviours, it is firstly necessary to consider that errors can have different degrees of consequences, and people might judge them differently according their own perception. Factors that influence their perception may include the severity and duration between errors, i.e. the impact of isolated 'big errors', or an accumulation of 'small errors'. For example, Muir and Moray (1996) argue that human perceptions of a machine are affected in a more severe and long-term way by an accumulation of 'small' errors rather than one single 'big' error. But what is perceived as a 'big error' and what is a 'small error'? People have individual differences, including age, gender, cultural and social habits, which may impact their perceptions of what are considered big or small errors. In order to study the differences in terms of the impact of errors in a human-robot interaction, first I have to establish what people consider subjectively to be 'small' or 'big' errors exhibited by a companion robot.

People's trust in a robot does not depend only on the robot's performances, but also on their own individual differences. People can differ from each other for their previous experiences, background, age and gender, personalities and predispositions (Williamson 2018). These characteristics often rule their attitude, their way of interacting with others, and their perception of a situation or others. Therefore, researchers expected that people's interactions with robots also might be affected by their differences. They place emphasis on the impact of these on their trust in robots (Hancock et al. 2011). However, it is likely that a person's antecedents of trust will vary their perception of the different type of robot errors, and consequently a loss of trust might linger differently for different individuals. In order to foster long-lasting relationship between robots and users, it is important to identify which are the factors affecting people's trust that can be used to create coping mechanisms for robots to mitigate and respond in case of lack of trust in robots. In this direction, socially-aware robots are perceived more positively by people (Kerstin Dautenhahn 2007a), in particular, in social context and situations (Rios-Martinez et al. 2015), such as home environments. Social robotics systems, therefore, should integrate people's modes of communication. People communicate between each other using different modalities, including natural language, body gesture, and other form of communicative signals and behaviours linked to cultural and social expectations, preferences and sharing same mental models. Several researchers are trying to develop robots able to interact with social signals to foster trust in robots (Martelaro et al. 2016; Alessandra Rossi, Garcia, Maya, et al. 2019; Sebo et al. 2019). However, most of these efforts assumed flawless robots behaviours and positive individuals' attitude towards the robots. Nevertheless, to effectively deploy robots that provide successful collaborative and cooperative teamwork, it is fundamental to consider that robots can be faulty, and consequently people's predisposition towards them can vary overtime. Indeed, Rossi et al. (S. Rossi et al. 2020) claimed the necessity of having a robot capable of self-adapting to satisfy people's needs (i.e. personality, emotions, preferences, habits), and incorporating a reactive and predictive meta-cognition models to reason about the situational context (i.e. its own erroneous behaviours) and provide socially acceptable behaviours.

1.2 The Approach

In this thesis, I want to explore the factors that affect people's trust in order to foster an effective and successful cooperation between humans and people in public and

private human-oriented environments. Our aim is to create guidelines that allow humans to trust robots that are able to look after their well-being by adopting human-like behaviours.
I want to identify the factors affecting people's trust in robots, and create appropriate metrics and mechanisms to develop socially acceptable companion robots. Indeed, the first aim of our research is to understand the effects of robot erroneous behaviours on people's trust in them. I want to show that the type of an error and the timing in which it happens affect people's trust in the robot differently. I want to test how people choose to trust a robot according the robot's performances, the trade-off between risk of failure and possible outcome of a task, but also their predispositions towards robots and other people. The goal is to provide mechanisms to mitigate robot's errors on people's loss of trust in order to foster solid long-lasting relationships between humans and robots.
Hence, my first studies aim to identify whether people perceive errors according their magnitude of consequences, and how these errors affect their trust in the robot. I firstly categorised 20 different robot errors according their severity of consequences (small, moderate and severe) using a questionnaire-based study. Then, I used these information to investigate people's changes of trust in a robot that occasionally made small, big or a combination of both, errors during an interaction. I used an interactive storyboard to study people's choices of trusting the robot, and understanding how they are affected by people's age, gender, nationality, personalities traits, disposition of trust and previous experiences with other robots. These results drove me to continue to investigate the correlation between individuals' differences and their trust in the robot. In this direction, further effort was invested in understanding how people's awareness of robots' potentialities and limitations affect their trust in robots. I collected the perception of trust of pupils of different groups of school age after gradually increasing their knowledge and awareness of the robot.
The findings of my studies showed me which are the factors affecting people's trust in the robot. Lack of trust has big impact on the interactions between humans and robots, therefore, a robot should have mechanisms that allows it to recover it. I want to design robots that are able to mitigate a loss of trust, so I tested the effects of robot's social behaviours on people's trust. Firstly, I identified which are the social signals in literature that influence people's perception of a robot as social entity. Then, I rated the social cues for higher impact on people's perception of sociability in robots. In addition, I investigated the impact of robot appearances on people's perception of robot sociability. The most relevant social signals that were unequivocally perceived by people regardless of robot's appearances were

used to observe people's trust in a robot in three tasks having different criticality of consequences.
Finally, I want to integrate my studies and observations to investigate whether humans' trust of a robot changes over time if the initial conditions are not anymore the same, i.e. the robot has erroneous behaviours. The study served to investigate if people would trust a robot that broke their trust in a initial or later stage of the interaction.
In summary, the research conducted in this thesis has been carried out to investigate the following main research questions:

RQ-1 How do various types of robot errors affect human's trust in a robot?

RQ-2 Can people's trust in a robot change diversely according their personal differences?

RQ-3 Does a robot that expresses social conventions gain more people's trust in it than a non-social robot?

RQ-4 Does people's trust on a robot change over time if the initial conditions of trusting the robot change?

1.2.1 Extended research questions

As my research project evolved, the primary research questions were adapted and expanded to fully explore the aspects and implications generated from the literature review.
The resulting research questions will be addressed in this thesis:

RQ-1.1 How do people classify errors?

RQ-1.2 Which kind of erroneous behaviours impact a human's trust in a robot?

RQ-1.3 Does the impact on trust change if the error happens at the beginning or end of an interaction?

RQ-1.4 How does the combination of timing and frequency of trust have any significant impact on trust recovery process?

RQ-2.1 Do personalities and characteristics of humans affect their perception and trust of a robot?

RQ-2.2 Does previous experience with robots affect human perception and trust of a robot?

RQ-2.3 Can different levels of awareness of a robot's potential affect human perception and trust of a robot?

RQ-3.1 Which are social cues people recognise?

RQ-3.2 Do different robot appearances affect people's perception of a robot?

RQ-3.3 Does a robot that expresses social conventions gain more of people's trust than those that do not express social conventions?

RQ-4 Does people's trust on a robot change over time if the initial conditions (positive or negative) of trust in the robot changes?

1.3 Contribution to Knowledge

Investigating people's trust in robots can be a challenging task, in particular when it comes to identify which are the factors that undermine it, i.e. robots errors. The research presented in this thesis introduces the novel concept that errors made by a robot may have different magnitude of consequences, and, therefore, they will affect differently people's acceptance and trust in robots. The studies presented in Chapters 4, 5 and 6 resulted in two conference papers (Alessandra Rossi, Kerstin Dautenhahn, Kheng Lee Koay, and Michael L. Walters 2017a; Alessandra Rossi, Kerstin Dautenhahn, Kheng Lee Koay, and Michael L. Walters 2017b) and a journal paper (Alessandra Rossi, Kerstin Dautenhahn, Kheng Lee Koay, and Michael L. Walters 2018).
The study presented in Chapter 5 has been conducted with an immersive narrative approach through a crowdsourcing service. This approach allowed us to overcome the difficult challenges of investigating people's trust in realistic life-threatening scenarios without endangering and distress participants. Another challenge when conducting this type of research is related to the difficulty of designing studies where people interact with a robot that is fully functional and versatile to execute realistic tasks. For example, a robot should be able to manipulate objects, navigate autonomously and being able of converse with users in noisy environments. To the best of my knowledge, such interactive storyboards have not been used in similar large scale studies to investigate HRI.

Another contribution to knowledge relevant for roboticist research community from this thesis is presented in Chapter 12. A series of live HRIs have been conducted to investigate whether people's trust in a robot changes overtime or it is irreversibly affected by a negative initial perception of the robot, i.e. due to robot errors. The findings from this study gave further insight on how trust human-robot relationships develop in HRI.

Further investigations to identify the antecedents of trust have been conducted in Chapter 8 which were published in (A. Rossi et al. 2019; Alessandra Rossi, Holthaus, et al. 2018a; Alessandra Rossi, Holthaus, et al. 2018b).

This research also aimed to provide recovery mechanisms in trust violation situations, many of the novel results resulting from the studies presented in Chapters 9, 10 and 11 have been accepted and presented in (Alessandra Rossi, Kerstin Dautenhahn, Kheng Lee Koay, and Michael L. Walters 2020; Alessandra Rossi, Kerstin Dautenhahn, Kheng Lee Koay, Michael L. Walters, and Holthaus 2020).

1.3.1 Publications

The research conducted during the project produced several journal manuscripts, conferences' and workshops papers. I completed them, after the supervisors' revision and approval, and I submitted for publication. I was supported by my supervisors through feedback and revision meetings during the design, development and evaluation of the studies presented in this thesis. Three published papers have been written in collaboration with other colleagues. I entirely structured and wrote most of the content of these papers. I also supervised and revised the contributions to the papers given by my colleagues.

The following papers are presented as part of this thesis in the relative coming Chapters:

2017 A. Rossi, K. Dautenhahn, K. L. Koay, M. L. Walters, How the Timing and Magnitude of Robot Errors Influence Peoples' Trust of Robots in an Emergency Scenario. In: Kheddar A. et al. (eds) Social Robotics. ICSR 2017. Lecture Notes in Computer Science, vol 10652. Springer, Cham

2017 A. Rossi, K. Dautenhahn, K. L. Koay, M. L. Walters, Human Perceptions of the Severity of Domestic Robot Errors. In: Kheddar A. et al. (eds) Social Robotics. ICSR 2017. Lecture Notes in Computer Science, vol 10652. Springer, Cham

2017 A. Rossi, K. Dautenhahn, K. L. Koay and J. Saunders, Investigating human perceptions of trust in robots for safe HRI in home environments, Human-Robot Interaction Pioneers Workshop in conjunction with the 2017 ACM/IEEE International Conference on Human-Robot Interaction, March, 2017.

2018 A. Rossi, K. Dautenhahn, K., Koay and M.L. Walters, The impact of peoples' personal dispositions and personalities on their trust of robots in an emergency scenario. Paladyn, Journal of Behavioral Robotics, 9(1), pp. 137-154

2018 A. Rossi, K. Dautenhahn, K. L. Koay and M. L. Walters, Investigating human perception of trust and social cues in robots for safe HRI in home environments, International PhD Conference on Safe and Social Robotics (SSR-2018), Madrid, 29-30 September 2018

2018 A. Rossi, P. Holthaus, K. Dautenhahn, K. L. Koay, and M. L. Walters. Getting to know Pepper: Effects of people's awareness of a robot's capabilities on their trust in the robot. In Proceedings of the 6th International Conference on Human-Agent Interaction (HAI '18). ACM, New York, NY, USA, 246-252

2018 A. Rossi, P. Holthaus, K. Dautenhahn, K.L. Koay and M.L. Walters, Programming Pepper: What can you make a humanoid robot do?, 3rd Workshop on Behaviour Adaptation, Interaction and Learning for Assistive Robotics - IEEE-RO-MAN 2018, 27 - 31 of August 2018, Nanjing, China

2019 A. Rossi, K. Dautenhahn, K.L. Koay and M.L. Walters, Evaluating the effects of an autonomous robot's social behaviours on people's trust. IEEE International Conference on Robot & Human Interactive Communication (RO-MAN)

2019 A. Rossi, K. Dautenhahn, K.L. Koay and M.L. Walters, Investigating human perceptions of trust and social cues in robots for safe HRI, Personal robotics and secure human-robot collaboration Workshop, IEEE ICDL- EPIROB, Oslo, Norway

2019 A. Rossi, S. Moros, K. Dautenhahn, K.L. Koay and M.L. Walters, Getting to know Kaspar: Effects of people's awareness of a robot's capabilities on their trust in the robot. IEEE International Conference on Robot & Human Interactive Communication (RO-MAN), New Delhi, India

2020 A. Rossi, K. Dautenhahn, K.L. Koay and M.L. Walters, How Social Robots Influence People's Trust in Critical Situations. IEEE International Conference on Robot & Human Interactive Communication (RO-MAN), Naples, Italy

2020 A. Rossi, K. Dautenhahn, K.L. Koay, M.L. Walters and P. Holthaus, Evaluating people's perceptions of trust in a robot in a repeated interactions study. 12th International Conference on Social Robotics (ICSR 2020), Golden, Colorado

During the project, I have been in two secondments: one research-oriented at SoftBank Robotics in Paris (France), and a training-oriented at Fraunhofer Institute for Manufacturing Engineering and Automation IPA in Stuttgart (Germany). The research carried in SoftBank Robotics brought the publication of the following paper that is not part of my thesis, but played an important role in the research carried out to further understanding the impact of social cues on people's trust in robot:

2019 A. Rossi, F. Garcia, A. Cruz Maya, K. Dautenhahn, K. L. Koay, M. L. Walters and A. K. Pandey. Investigating the Effects of Social Interactive Behaviours of a Robot on People's Trust During a Navigation Task. In: Althoefer K., Konstantinova J., Zhang K. (eds) Towards Autonomous Robotic Systems. TAROS 2019. Lecture Notes in Computer Science, vol 11649. Springer, Cham

I also collaborated with the University of Naples Federico II (Italy). The collaboration brought a publication that is not part of this thesis, but it is based on the knowledge and experiences gained during the project:

2020 S. Rossi, A. Rossi and K. Dautenhahn. The Secret Life of Robots: Perspectives and Challenges for Robot's Behaviours During Non-interactive Tasks. International Journal of Social Robotics. Springer Nature.

Another publication that is not part of this thesis, but based on the knowledge and experiences gained during this project, resulted from the fruitful collaboration with the University of Canterbury (New Zealand):

2020 M. Keijsers and A. Rossi. RoboCup 2050: from human fouls to robot bullyism. Virtual RoboCup Humanoid Open Workshops (V-RoHOW) (accepted).

I also had other relevant collaborations that brought to publications in robotics applications in the Education field.
The results of Chapters 7, 8, 10 and 12 will be submitted to the International Journal of Social Robotics[1], and 12th International Conference on Social Robotics (ICSR 2020)[2].

1.4 Structure of the Thesis

This thesis is organised to answer the research questions presented in Section 1.2.1. The studies designed for each questions are presented in the Chapters 4-12 as illustrated in Table 1.1.

Table 1.1: Distributions of research questions across this thesis' chapters.

Experiment and description	**Research Questions**
Chapter 4	RQ-1.1
Chapter 5	RQ-1.2, RQ-1.3, RQ-1.4
Chapter 6	RQ-2.1
Chapter 7	RQ-2.2
Chapter 8	RQ-2.3
Chapter 9	RQ-3.1
Chapter 10	RQ-3.2
Chapter 11	RQ-3.3
Chapter 12	RQ-4

Specifically, this thesis is organised as follows:

Chapter 2 I give an overview of the current state-of-art of HRI, briefly analysing the human skills application in relation to robots and

[1] International Journal of Social Robotics https://www.springer.com/journal/12369
[2] 12th International Conference on Social Robotics (ICSR 2020) https://sites.psu.edu/icsr2020/

introducing the develop of trust in human-human and human-robot interactions, with particular focus on the resulting issues and recovery mechanisms in trust violation situations.

Chapter 3 In this Chapter, I provided an overview of the software and hardware used for the studies presented in this thesis, including also details of the experimental setups implemented.

Chapter 4 This Chapter provides a pre-study to investigate people's perception of severity of different scenarios during which a generic robot makes errors. In this study, participants were asked to rate a robot's errors according the magnitude of their consequences. The results of this study was fundamental for the planning of the next one. The results of this study were presented at the conference Social Robotics (ICSR) (Alessandra Rossi, Kerstin Dautenhahn, Kheng Lee Koay, and Michael L. Walters 2017b).

Chapter 5 In this Chapter, I focused on investigating the impact of different magnitude and timing of a robot's errors on people's trust in the robot. In this study, participants were tested using a storyboard presenting five different experimental conditions in which the robot made combinations of errors with different magnitude. At the end of each interactions, I evaluated people's choice of trusting or not the robot in an emergency scenario. The results of this study were presented at the conference Social Robotics (ICSR) (Alessandra Rossi, Kerstin Dautenhahn, Kheng Lee Koay, and Michael L. Walters 2017a).

Chapter 6 Continuing to analysing the participants' responses collected in Chapter 5, I analyse the effects of participants personality traits and predisposition of trust on their choices of trusting the robot. The results of this study are published in Paladyn, Journal of Behavioral Robotics (Alessandra Rossi, Kerstin Dautenhahn, Kheng Lee Koay, and Michael L. Walters 2018).

Chapter 7 This Chapter presents more results of Chapter 5. In particular, I analysed the effects of people's previous experience with robots on their choices for trusting the robot.

Chapter 8 In this Chapter, I focused on how people awareness of a robot's real capabilities and limitations affected their trust in the robot. This study was conducted in two events as part of the UK Robotics Week in a local primary and a secondary school. The Chapter introduces the robots used in each school, the experimental conditions and interactions. Finally, I define three level of awareness, I also present the resulting pupils' trust in the robots and a comparison between the groups age of the two schools. The results of this study were presented at the 6th International Conference on Human-Agent Interaction (HAI) (Alessandra Rossi, Holthaus, et al. 2018a), IEEE International Conference on Robot and Human Interactive Communication (RO-MAN) (A. Rossi et al. 2019; Alessandra Rossi, Holthaus, et al. 2018b).

Chapter 9 In this Chapter, I define which are the social cues and behaviours used in literature that increase the social perception of robots and, consequently, people's acceptance of them.

Chapter 10 In this Chapter, I focus on the impact of two robots' appearances on people's perception of robots' social expressions.

Chapter 11 Following Chapter 9 and 10, in Study 6 I investigate the impact of social cues on humans' trust in robots. The results of this study were presented at IEEE International Conference on Robot & Human Interactive Communication (RO-MAN) (Alessandra Rossi, Kerstin Dautenhahn, Kheng Lee Koay, and Michael L. Walters 2020).

Chapter 12 In this Chapter, I present my last study that focuses on the evolution of people's trust in the robot overtime. The results of this study have been published in the 12th International Conference on Social Robotics (ICSR 2020) (Alessandra Rossi, Kerstin Dautenhahn, Kheng Lee Koay, Michael L. Walters, and Holthaus 2020).

Chapter 13 This Chapter summarises the results of my studies, the novel impact of this research on future robotics applications, outlining limitations of current research and future works.

Appendix A	In this Appendix, I listed all publications produced during this project.
Appendixes B to H	For each study, I included ethical approvals and questionnaires used for the experiments.
Appendix I	For each study, I included a list of implemented software.

Chapter 2

Background

> The more I study, the more insatiable do I feel my genius for it to be.
>
> *Ada Lovelace*

In this Section, I give an overview of the current state-of-art of Human-Robot Interaction (HRI), briefly analysing the human social skills in relation to robots and introducing the development of trust in human-robot interactions. In particular, I will present issues related to robots violating human trust and the need of recovery mechanisms for robots to regain the trust of humans.
Reviewing the literature regarding acceptance and trust in HRI, it appears to be a number of open questions that need to be addressed in order to establish effective collaborations between humans and robots in real-world applications. In particular, I identified four principal open areas that should be investigated to create guidelines for the successful deployment of robots in the wild. These areas are focused on: 1) the robot's abilities and limitations, 2) the individuals' differences, 3) the dynamics of human-robot trust, and 4) the interaction between humans and robots overtime.

2.1 Trust in Human-Human Interaction

Trust is an interdisciplinary fundamental factor that plays a significant role in interpersonal and economic interactions. Therefore, several studies define the concept of trust in different fields, including in Human-Human (HHI), in Human-Computer (HCI) and Human-Robot Interactions (HRI). Identifying the principal points associated to the development and preservation of trust between people can guide roboticists to build trustful robots for successful HRI.

2.1.1 Definition of Trust

Simpson (2007; 2007) highlights four core principles that affect trust. First, individuals assess the degree of trust by observing a partner acting unselfishly and supporting the best interests of both the individuals and the relationship. Second, individuals may purposefully create situations to test their partner's trust. Third, individuals with lower self-esteem may be less trustful of their partner. Finally, the level of trust in short-term or long-term relationships cannot be fully understood without considering the predisposition of trust of all the parties involved in the relationship.

Mayer et al. (2007; 1995) established that trust is constructed from a perception of ability, benevolence and integrity.

Human's trust is also affected by the perception of the risk of the interaction with other humans. The popular poker game is a concrete example to how risk-taking behaviours affect the credibility of a poker player, and thus it is important for all players to develop a good reputation during a game (Billings 1995). Deutsch (1958) claims that risk-taking and trusting behaviour are different sides of the same coin, and that a person is willing to take a risk only if the odds of a possible positive outcome are greater than those for a potential loss. Golder and Donath (2004) claim that a good reputation is very important in enhancing trust both in short and long term relations.

Deutsch defines trust as "*confidence that [one] will find what is desired [from another] rather than what is feared.*" (Deutsch 1958, p. 148). Although multiple definitions exist, and several previous studies have adopted one of the first definitions of trust (Deutsch 1958), there is a convergent tendency (Yu et al. 2017) towards using the definition proposed by Lee (2004, p. 51) "Trust can be defined as the attitude that an agent will help achieve an individual's goals in a situation characterized by uncertainty and vulnerability". This definition encapsulates the most important factors that can affect Human-Robot Trust: Human-related, Robot-related and Context-related. Specific human-related factors include demographics, personality traits, prior experiences, situations awareness, self-confidence, propensity of trust and attitude towards robots. Robot-related factors include robot's reliability, transparency, legibility and predictability of robot's behaviours, robot's personality and anthropomorphism, autonomy and failures. Factors related to context comprehend communication modes and shared mental models between people and robots, specific task typology and complexity, and the environment.

2.1.2 Building Trust in Human-Human Interaction

Humans rely on each other everyday for different things and trust plays a key role in this. For example, students may trust their mentors to provide them with reliable information, and travellers trust the pilot to take them to their destination safely. However, trust is a complex feeling (Kramer et al. 2003) and it may also be influenced by different internal and external factors.
Internal factors can be based on cognitive and affective foundations (McAllister 1995) which are related to the expectation of reliability of other individual's actions and behaviours, or to the development of an affection that supports the belief that the involved parts are invested in a common goal (Becker 1996; Lewis et al. 1985). Emotional foundations also exists because people believe their trusted partner has genuine care and concern for their well-being. Indeed, another aspect to consider in order to build people's trust is knowing when to apologise. Brooks et al. (2014) demonstrated that, while an apology could be often seen as admission of blameworthiness by the culprit, but it also could be viewed as an expression of regret for an unpleasant event, even when the apologiser is not to be blamed for the situation (e.g. a superfluous apology because it is a rainy day). They demonstrated that superfluous apologies increased trust due to a perception of empathic concern towards the person. However, there can be different factors that can mitigate trust (e.g. insincere behaviour or characteristics of both the apologiser and recipient).

2.1.3 Trust Violation and Trust Recovery in Human-Human Interaction

Despite people investing a substantial effort in building and nurturing trust in interpersonal relationships, trust can be broken. When there is a breach of trust, people may react very differently (Haselhuhn et al. 2010; Schilke et al. 2013; J. A. Simpson 2007). Some people are quick to forgive while others believe that once the trust is broken, it can not be mended (Slovic 2000). Indeed, it is very hard to re-establish the trust lost between two parties after a breach as there are many different factors that affect the recovery of trust.
Haselhuhn et al. (2010) showed that people who believe in the moral that a character can change over time (called incremental beliefs) are more likely to trust again the person who broke their trust if they show regrets, promise to change and

apologise. While those who do not believe that moral character can change (called entity beliefs) are sceptical in trusting the other person again.
Apologies and offers of penance seem to mitigate the effects of a loss of trust (Bottom et al. 2002; H. et al. 2004; Schweitzer et al. 2006). However, it is more difficult to establish how long a loss of trust will linger in the dynamics of human relationships. Trust of people is not only affected by the characteristics of the persons involved, but also by the type and length of the relationship they have. Moreover, "*Will people be more likely to forgive a breach of trust in an earlier or later stage of an interpersonal relationship?*" (Schilke et al. 2013). Schilke et al. (2013) investigated how certain kind of relationships recover better and faster after a violation of trust and how the timing of the violation affects the recovery. They demonstrated that people in longer relationship re-establish the trust easier than that of a newer relationship. Therefore, it is expected that the trust recovery process will also be easier for people whom have established a longer relationship than those in a newer relationship.

2.2 Trust in Human-Robot Interaction

A task can be successfully accomplished as a team effort between people and robots only if people trust that the robots will protect them and the positive outcome of the task-goal will be achieved. In this situation, the level of trust people have for the robot is directly associated with their perception of the robot's reliability (Ross 2008).
Hancock et al. (2011) identify 33 factors influencing trust in HRI, grouped within 3 categories and 6 sub-categories. The main categories are: Human-related, such as self-confidence, prior experience with robots and operator workload; Robot-related, such as proximity, robot's embodiment, transparency, level of autonomy and failure rates; and Environmental, such as communication and team collaboration.
Trust affects people's perception of the usefulness of the information and capabilities of a robot(Cameron et al. 2015). The consistency and reliability of a robot's behaviours can positively affect people's trust, while erroneous behaviours can lead to a loss of trust. Indeed, the absence or a lower level of trust in a robot can lead a person to intervene or take over control of the task from the robot (Steinfeld et al. 2006; E. D. Visser et al. 2006). Trust is particularly important in unstructured situations characterised by uncertainty and risk, and in which the current literature gives mixed results. For example, while Garreau (2007) showed that trust can be easily built during a stressful situation. However, a military robot developed by

Ogreten et al. (2010) was never used in action because the soldiers never trusted the unexpected movements of the robot.

2.2.1 Robot's errors

People's perception of a robot's reliability depends not only on the ability of a robot to complete a task, but also by its behaviours to reach a goal. Several studies (Honig et al. 2018; Lemaignan et al. 2015; Short et al. 2010) showed that people might consider unexpected and incoherent behaviours, perceived failures, and actual failures as robot errors. Humans may perceive erroneous robot behaviours according to their expectations of a robot's proper functions (Michael L. Walters et al. 2011). For example, a robot that navigates too slowly might be considered having faulty behaviours. Honing et al. (2018) proposed a taxonomy for classifying possible types of robotic failures, also referred as faults (see Figure 2.1). They identified two principal categories of errors: technical failures and interaction failures. Technical failures are considered errors caused by hardware or software problems, which can depend on an erroneous design, communication or processing. In contrast, interaction errors are related to social norm violations, organisational and mental-models faults in the interaction with the environment and other agents, including people.

Bainbridge et al. (2011) found that participants were happy to follow a robot's instructions to throw books in the trash if the robot was present in the room with them, but not when the robot was not physically in the same room.

Other studies (M. Desai et al. 2012; Munjal Desai et al. 2013) showed that the order of presentation of the decreased reliability produces an evident drop in the trust in the robot which can be restored by continuing the interaction. They showed that warning the participants about a drop in the robot's performance can mitigate the loss in trust.

Wang et al. (2015)'s studies showed that the frequency and significance of errors can impact humans' trust in an imperfect on-line system. They showed that people are not willing to follow an imperfect robot if the outcomes are severe, and could lead to permanent consequences, e.g. death of the human.

Booth et al. (2017) investigated participants' responses to a robot's request to move in a secure-access student dormitory. They conducted the experiment with two conditions: 1) an anonymous robot and 2) a food delivery robot, where both asked to enter the building. They observed that participants were more likely to let the food delivery robot enter the building or in situations when they were in a group.

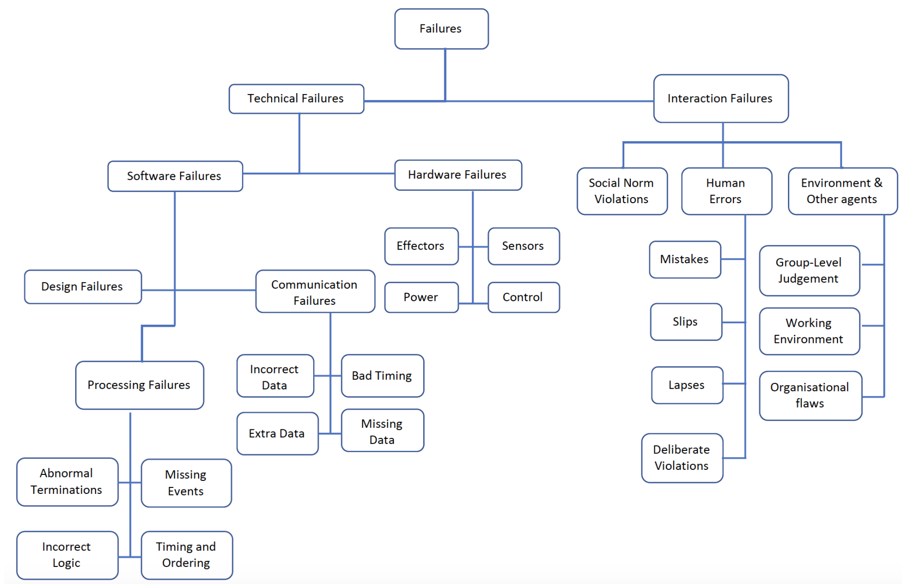


Figure 2.1: A human-robot failure taxonomy reproduced by Honig et al. (2018).

Robinette et al. (2015) investigated the effects of apologies, promises and additional reasons given by a robot for its errors on participants' trust in a simulated fire evacuation scenario conducted in a virtual environment. They showed that participants' trust was repaired if the robot apologised and promised to not repeat the error soon after it made the error but not during the emergency.

Salem et al. (2015) studied human perception of trust in robots, and how willing they are to follow a robot showing faulty behaviours. They showed that no matter how erratic the behaviour of the robots, participants followed the instructions of the robots. Similarly, Robinette et al. (2016) used an emergency evacuation scenario, with artificial smoke and a smoke detector, in which a robot guided a person to an exit, in order to study how willing were humans to follow a robot that had previously exhibit erratic behaviour. Their results indicated that all the participants of the experiment followed the robot's instruction. In both experiments participants trusted the robots for different reasons. For example, some of them believed it was all staged, others that they were supposed to follow it because they accepted to participate in the experiments.

Both Salem et al. (2015) and Robinette et al. (2016)'s works showed that some participants did believe that they were acting according to the experimenter decisions and that their lives where not in danger. Therefore, it is still not clear from these results whether faulty robots are trusted by humans, and whether humans can believe that robots can look after their safety and well-being.

Research questions

Reviewing the literature regarding trust in HRI it is clear that none of the studies consider the magnitude of robot errors, nor the possibility for a robot to initiate a trust recovery process to earn back people's trust similar to that of a human-human trust recovery in romantic, working, family or other type of relationship (Munjal Desai et al. 2013; Muir et al. 1996; Robinette, W. Li, et al. 2016; Maha Salem, Gabriella Lakatos, et al. 2015). In particular, several research questions have been raised that need to be investigated in order to provide robot's behaviours and mechanisms aimed to act in case of an error and to re-gain a loss of trust. The first question to answer is **RQ-1 - How do various type of robot errors affect human's trust in a robot?**, and it aims to identify how the magnitude and the timing in which robots' errors happen affect people's trust in a robot.
In particular, I believed that people's trust can be affected differently depending on the magnitude of the error consequences. However, people are different from each other, in terms of personality, background, perceptions of the environments and in numerous other ways. I expect, for this reason, that they might perceive the errors made by a robot differently, hence that the same robot error may affect them differently. Therefore, it is essential to classify errors according their perceived severity of error consequences. my first research question is *RQ1.1 - How do people classify errors?*
Current literature focused on robots' errors did not distinguish different magnitude of consequences (severe or limited). I expect that people's trust might be affected differently depending on the severity level of the robots' errors. I also believe that there is a correlation between the timing in which a robot makes an error and the loss of trust of the human in the robot. This research has been guided by the following research questions:

RQ-1.2 - Which kind of erroneous behaviours impact a human's trust in a robot?

RQ-1.3 - Does the impact on trust change if the error happens at the beginning or end of an interaction?

RQ-1.4 - How does the combination of timing and frequency of trust have any significant impact on trust recovery process?

2.2.2 Individual's differences

Individual differences have been a key subject area in psychology research for several decades. According to Williamson are defined as "Individual differences are the more-or-less enduring psychological characteristics that distinguish one person from another and thus help to define each person's individuality"(Williamson 2018, p. 1). Among those characteristics, intelligence, personality traits, skills and aptitudes are recognised as the most relevant among the differences(Williamson 2018). People's differences are also important for understanding their acceptance and perception of trust in robots. Therefore it is important to investigate if individual differences play any role in people's perception of trust and acceptance of robots.

Antecedents of Trust

Recent literature regarding the role of trust in HRI indicates that people's antecedents have a dynamic influence on their trust in robots and automated systems. Accoding to Hancock et al. (2011) individuals' differences, including propensity to trust and personality traits such as agreeableness and extroversion, can affect the teamwork between people and robots (Barrick et al. 1998; J. S. Elson et al. 2018; Rotter 1967). Various studies (Gockley et al. 2006; Kerstin Sophie Haring et al. 2013; Robert 2018) have also showed that people with a more extroverted personality are more comfortable with robots in their personal spaces.
Some studies found that people's propensity to trust others may affect their trust in robots (Adams et al. 2003; J. D. Lee et al. 2004). Costa et al. (2001) showed that the level of trust of the participants towards the robot depended on their disposition for trusting others.
Another personality-based factor is the individual's self-confidence and -esteem which has been associated with their degree of trust in a robot in HRI (Freedy et al. 2007; J. Lee et al. 1992).
These studies ascertained that factors affecting trust include people's previous experience, their personal characteristics, traits and predisposition towards others. These studies however, did not consider the effects of robots' errors on trust.

Human awareness in Human-Robot Interaction

In Human-Human interactions, situational awareness can change the way peoples interact with each other, as well as it might change the outcomes of the interaction. For example, if people are aware of an outbreak of infectious disease within their community, they will take precautions to avoid the spread of the disease by changing their usual interaction behaviour (Funk et al. 2009) (e.g. avoiding contacts with the infected person, wearing protective masks, vaccinating etc.). Marketing advertisements use persuasion techniques to induce people to buy a product. Similarly, people with persuasion knowledge can shape their response and avoid being influenced by a persuasion attempt (Friestad et al. 1994).

In Human-Robot Interaction, Atkinson et al. (2014) suggested that humans' trust in robots increases according to a greater shared awareness of the agents involved, of the activities and situations between human users and robots. This greater awareness consequentially increases the success of a human-robot interaction. Tseng et al. (2013) developed a human awareness Decision Network model that enabled a robot to adapt its behaviours to respond to the different feedback from the user in order to meet their expectations.

Moreover, a person who is new to robotics technologies may be influenced by science fiction and may tend to overtrust the robot and its capabilities. Abney et al. (2017) define overtrust as the willingness of a person to accept the risk of delegating a task to a robot if 1) she believes that the robot is able to complete it or 2) her expectation is that the robot is able to mitigate the risk.

Borenstein et al. (2018) found that 62% of pediatric patients, their parents, and other caregivers would trust a robotic exoskeletons to be able to handle dangerous situations, even if the robot did not have that capability. Booth et al. (2017) investigated participants' responses to a robot's request to move in a secure-access student dormitory. They conducted the experiment with two conditions: 1) an anonymous robot and 2) a food delivery robot, where both asked to enter the building. They observed that participants were more likely to let the food delivery robot enter the building or in situations when they were in a group.

In HRI several definitions of "human awareness" exist (J. L. Drury et al. 2003; J. Drury 2000; Endsley 1988; Gutwin et al. 1995). In order to narrowing down the open research questions, in this research, I define human awareness in the context of HRI as the human understanding of the capabilities and functionality of a robot:

- to be aware of environment and people presence around itself;
- to interact with one or more humans according social conventions;
- to perform a specific activity; and to have artificial intelligence.

Research questions

As can be seen from literature, antecedents of trust, including individuals' differences in terms of personality, background, age, gender, previous experiences and awareness of the robots, may affect people's perception of robots. However, it is not entirely clear how they influence humans' trust in robots, in particular in a situation of uncertainty. Moreover, previous research was not focused on robots erroneous behaviours with different levels of consequences. In order to answer the research question **RQ-2 - Can people's trust in a robot change diversely according their personal differences?**, this research has been guided by the following research questions:

- RQ-2.1 - Do personalities and characteristics of humans affect their perception and trust of a robot?
- RQ-2.2 - Does previous experience with robots affect human perception and trust of a robot?
- RQ-2.3 - Can different levels of awareness of a robot's potential affect human perception and trust of a robot?

2.2.3 Social dynamics of human-robot interaction

Several previous research studies have shown that socially interactive robots are accepted better by humans compared to non-socially interactive robots (Kerstin Dautenhahn 2007a; Kerstin Dautenhahn 2007b). In particular, robots' social behaviours are essential in social environments such as hospitals, shopping malls, workplace and also private environments where interaction with people is a main part of the robots' tasks (Rios-Martinez et al. 2015). Robots need to be aware of the context and situation to interact with the people and other agents sharing such environments. Context-awareness is defined by human cultural norms, social signals and individual preferences of its inhabitants (Chen et al. 2018; Rios-Martinez et al. 2015; Shu et al. 2017). Therefore, it is essential for a comfortable and long-lasting interaction that people are able to communicate and interpret communication signals and norms that may involve natural language, gesture, pose and body language.
Syrdal et al. (2010) showed that dog-inspired affective cues communicate a sense of affinity and relationship with humans. Non-verbal cues could be also used by robots to identify the person's need of interaction and offering assistance even

when not directly asked (Goodrich et al. 2003; Maha Salem, Friederike Eyssel, et al. 2013).
Martelaro et al. (2016) established that trust, disclosure, and a sense of companionship are related to expressiveness and vulnerability. They showed how a sense of the robot's vulnerability, through facial expressions, colour and movements, increased people's trust, companionship, and disclosure towards the robot. Other aspects such as the robots' type, size, proximity, and behaviour also affected perception of robots (Bainbridge et al. 2011; K. L. Koay, Dag Sverre Syrdal, et al. 2007). For example, Hall (1966) identified four comfortable distances when interacting in different situation: intimate space for close interactions; personal space typically is reserved for interaction with family and friends; social space is used in business and public spaces; and public space is used in audience-speaker situations. Most people prefer to interact with a robot that keeps a personal or social space distance from them, however this might vary according their personality (Tapus et al. 2008). Lohse et al. (2008) showed that robots with more extrovert personalities are perceived in a more positive way by users. Moreover, both vocal content and prosodic characteristics can express robot's personalities and emotions, and they can affect positively people's perception of robots (Apple et al. 1979; Tusing et al. 2000).
Robot's appearance also affects people's perception of its reliability and their expectation of a robot's functional abilities and behaviours (Goetz et al. 2003; Robins, K. Dautenhahn, Te Boekhorst, et al. 2004). For example, a robot with human-like physical features might be perceived to be more intelligent than one without (M L Walters et al. 2009). A few studies also showed that people might be inclined to assign a gender to robots (Stroessner et al. 2019), and consequently they might be also biased by it (Friederike Eyssel et al. 2012).
In human populated environments, robots should integrate strategies to favour social navigation, both in terms of path-planning (Kuderer et al. 2012; Rios-Martinez et al. 2015) and interaction (Ramírez et al. 2016), to increase people's perception of safety and social acceptance. Rossi et al. (2019) showed, for example, that a robot that slows down and asks people politely to make room to allow it to pass was perceived more positively by participants than a non-social robot. Other studies (Zhou et al. 2017) showed also that the timing and velocity of a robot's movements can communicate to people its emotions (Saerbeck et al. 2010), intentions (Gielniak et al. 2011), arousal and dominance (Saerbeck et al. 2010).
Sebo et al. (2019) examined the effects of competence and integrity (referred also as trust violation framing), and trust repair strategies using apologies or denial on people's trust of a robot. They showed that a robot that apologised rather than

denying a competence violation was trusted more by people. They also showed that a robot that denied an integrity trust violation was trusted more than a robot that apologised.
Alonso et al. (2018) showed that legibility and predictability of the robot's behaviours have positive effects on the development of people's trust in robots.

Research questions

A review of current literature highlighted that a socially-aware robot is essential for an effective interaction between people and robots in social environments. However, it is still not clear which are the implication of a robot's social behaviours on people's trust in them. This research aims to answer the research question **RQ-3 - Does a robot that expresses social conventions gain more people's trust in itthan a non-social robot?**. In particular, I believe there is no study investigating the effects of social behaviours in critical situations. In this context, several social cues have been identified. However, it is not clear whether people perceived them in a similar way. For example, does the robot's appearance affect the overall perception of the robot, including that of its social signals? This current research, therefore, has been carried out to answer the following research questions:

RQ-3.1 - Which are social cues people recognise?

RQ-3.2 - Do different robot appearances affect people's perception of a robot?

RQ-3.3 - Does a robot that expresses social conventions gain more of people's trust than those that do not express social conventions?

2.2.4 Trust in long-term human-robot interactions

Numerous studies investigating HHI showed that people often formed their mental model of another agent after only few minutes of interaction (Ambady et al. 2000; Wood 2014). However, the exposure of people to longer interactions might change people's attitude towards the robot (A. Y. Lee 2001; Zajonc 1968). Increasing familiarity with a robot might strengthen the relationship unless the novelty effects wore off and the people might lose interest in continuing the interaction (HT et al. 2011; Reber et al. 1998). Similarly, Paetzel et al. (2020) showed that people's initial negative perceptions of a robot with morph features was perceived as less appealing compared to a robot with mechanical or human-like features, and this perception did not change during time. Their findings also showed that people's perception of the robots as a threat and unease was not constant during the interaction, fluctuating between a more positive or negative feelings.

Indeed, considering trust in Human-Human Interactions, I observed that people in longer relationships might recover from a loss of trust more easily than people in new relationships (Haselhuhn et al. 2010). In literature, I found studies investigating long-term autonomous interactions between robots and humans that exceed a few weeks (M. M. d. Graaf, Ben Allouch, and Dijk 2016) or few months (T. Kanda et al. 2007) of interaction tasks. However, current state of technologies do not allow the deployment in real-world long-term application, and it is also quite difficult to recruit a substantial number of participants for in-lab studies. Another challenge of HRI studies, short- and long-term, is connected to ethical and legal issues. Due to these issues, participants are never in any real life-threatening danger during their interactions with robots. In this research, in order to overcome these issues I designed my studies to ensure people felt immersed in realistic scenarios, including supporting the interactions with a real environment and props (e.g. bottle, real biscuits), rewording activities (e.g. games), and providing a narrative approach to interaction (e.g. allowing a natural continuity of events, using a storyboard) (Kheng Lee Koay, Dag Sverre Syrdal, Kerstin Dautenhahn, et al. 2020).

van Maris et al. (2017) investigated the effects of robots' embodiments (a Softbank Robotics NAO robot vs. a virtual representation on a NAO on a tablet) on people's perception of trust over a period of six weeks. Contrary to previous works (Rae et al. 2013; Seo et al. 2015), they did not find any correlation between robot embodiment and people's trust in the agent.

In their paper (E. J. d. Visser et al. 2020), de Visser et al. presented a model for longitudinal social trust calibration. The study focused on creating a relationship that has balanced costs and risks, shared collaborations to achieve a goal, perceptions of themselves and the other agents in the relationship. It aims to provide techniques to possibly reduce overtrust or increase mistrust of robots. This model is based on the assumption that the human agent is positively invested in the success of the relationship. I believe, however, that this might not necessarily be true for people whose trust has been broken or is in the state of irreversibility.

Lee et al. (2012) investigated how the personalisation of a social robot affected people's interactions over a four month field experiment. The study showed that allowing the personalisation of a robot positively affected the way people perceived the robot and the overall interaction.

Research questions

In understanding the dynamics of trust between humans and robots, it is important to consider how the trust could change overtime. In particular, when the effects of novelty faded over time, and most importantly in the case of a breach of trust. Therefore, this research has been carried out to answer the following question:

RQ-4 - Does people's trust on a robot change over time if the initial conditions (positive or negative) of trust in the robot changes?

In the next Chapter 3, I will introduce the robots and setup used in my studies.

Chapter 3

Experimental Design and Implementation

> Imagination is more important than knowledge. Knowledge is limited. Imagination encircles the world.
>
> *Albert Einstein*

This Chapter illustrates the robots used in the studies presented in this thesis, how to program them, and details of experimental setup implemented.

3.1 The Robots

The studies composing this research have been designed to investigate specific open research questions (see Chapter 1). In order to address these research questions, different robots were used to conduct my research. These robots have different humanoid appearances. Each of them was chosen based on the requirements of the study they used for.

3.1.1 The virtual robot

Assessing people's trust in robots requires that participants are willing to take risks that might not result in a positive outcome for them (Deutsch 1958). However, causing distress or endangering participants' welfare presents ethical or legal issues

(Maha Salem, Gabriella Lakatos, et al. 2015). Moreover, current state of the robotic technologies do not allow yet to have fully functional robots that are able to interact autonomously and naturally with the participants. To overcome these issues, I decided to use a virtual robot for the study presented in Chapter 5.
The robot designed entirely for this research was a humanoid robot, called Jace (see Figure 3.1).

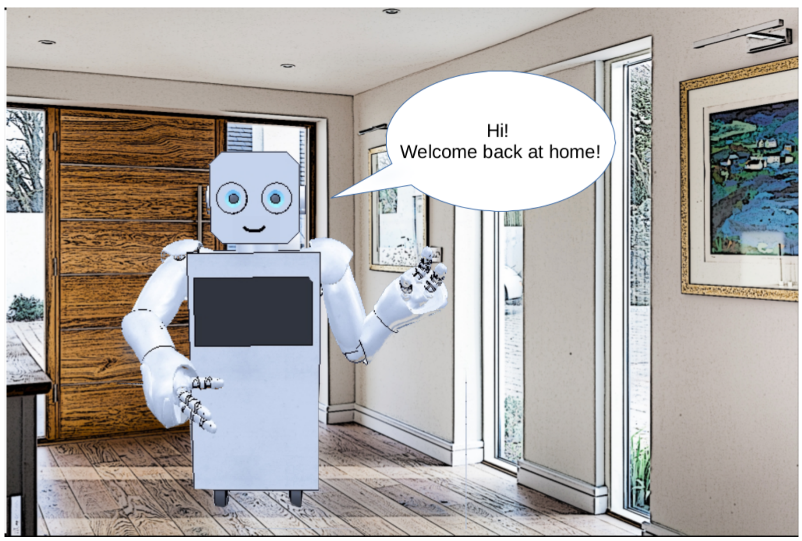


Figure 3.1: The virtual robot called Jace.

The robot has been designed with 3D objects to make it more realistic. Jace has been created using 3D Creation Softwares (Blender[1] and MakeHuman[2]).
MakeHuman and Blender are both free and open-source 3D computer graphics software. MakeHuman is designed for prototyping photorealistic humanoids, while Blender is designed for creating animated films, visual effects, art, 3D printed models, motion graphics, interactive 3D applications, and computer games.

3.1.2 Pepper

In Chapters 8, 9, 10 and 11, I used Pepper robot developed by Softbank Robotics (see Figure 3.2). In Chapter 8, this robot was chosen because it is possible to

[1]Blender https://www.blender.org/
[2]MakeHuman http://www.makehumancommunity.org/

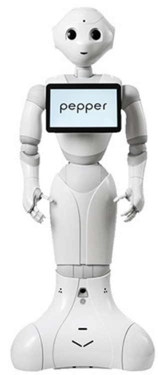


Figure 3.2: Pepper robot.

Pepper is a 1.2 meter tall omnidirectional wheeled humanoid robot platform capable of exhibiting body language, perceiving and interacting with its surroundings, and moving autonomously. Due to these characteristics, the system is suitable for a safe HRI (Pandey et al. 2018).
The version used is 1.7. It has 17 joints and 20 degrees of freedom (DoF). The platform is equipped with a large variety of sensors and actuators that ensure safe navigation and a high degree of expressiveness: LED's are distributed in the head (eyes and ears) and torso (shoulders) to support non-verbal communication by modifying color and intensity. The microphones and speakers allow verbal contact as well as environmental awareness. Sensing components include three laser sensors, two sonars and two infrared sensors located in the robot's base, as well as two cameras and a three-dimensional camera located in the head. In addition, two tactile detectors on the back of both hands allows human-robot physical awareness. The platform is powered by an Atom processor with a 1.91 GHz quad-core unit. On that allows the NAOqi SDK to orchestrate the different hardware elements as well as their access from other libraries (APIs).
The robot can also be controlled and programmed using the Choregraphe Suite provided by Softbank Robotics. This desktop application allows to use both existing and pre-configured boxes to control of the robot's functionalities, and program the robot's behaviours using the NAOqi and Python language. Choregraphe allows inexperienced and younger users to quickly learn to control Pepper. For this reason, I decided to use this robot in the study presented in Chapter 8.
Moreover, Pepper's appearance and capabilities are suitable for expressing social cues during its interactions with people in Chapters 9, 10 and 11.

3.1.3 Care-O-bot 4

A Mojin Robotics Care-O-bot 4[3] was used in the studies presented in Chapters 8, 10 and 12.

Care-O-bot 4, known also as Cob4, is a 1.58m tall robot of 140kg. Care-O-bot 4 is a humanoid robot with a omnidirectional base, torso, head, and two arms with grippers (see Figure 3.3). The robot has 29 degree of freedom, 3 safety laser scanners with safety PLC, 2-channel Safe-Torque-Off (STO) separate for mobile base and torso, 2-channel EM-Stop prepared in all modules, and 2-channel wireless Emergency-stop. It has three 3D cameras and a RGB LED light around its waist. It has a computer with a 4-6 Intel NUC i5, 256GB, 8GB RAM for each main component (head, torso, base). All internal computers have a full installation of Ubuntu Linux operating system. Care-O-bot 4 has also a Gigabit Ethernet and Wireless connection. It has a 15" touch screen, two 20W speakers and a microphone in the head.

Figure 3.3: Care-O-bot 4 robot.

This robot was chose because it has very similar functionality to Pepper, but it also able to manipulate objects.

[3]Mojin Robotics https://mojin-robotics.de/en

3.1.4 Kaspar

For the live study presented in Chapter 8, I used a child-size humanoid robot Kaspar (see Figure 3.4) which was designed and developed at the University of Hertfordshire [4]. Kaspar was initially built in 2005, and since then, it has been used in numerous studies that had the purpose of improving the lives of children with autism (S. Costa et al. 2015; Robins, K. Dautenhahn, and Dickerson 2009; Weiss et al. 2015). Several researchers from the Adaptive Systems Research group developed the robot controller and several interacting behaviours in the years.

Figure 3.4: Kaspar robot.

It has the size of a small child, 22 DOF. Kaspar is a fully programmable and it is specifically designed to help children with autism to help them to interact and communicate with other people. The first version of Kaspar was built in 2005, but since then many improvements have been made and features have been refined. The robot can be controlled via a keyboard Interface.

The robot has been part of several international projects, and also has been used in schools as well as in private houses and other research facilities (S. Costa et al. 2015; Huijnen et al. 2016; Wainer et al. 2014). Considering the success of interactions of this robot with children, I decided to use Kaspar in the study presented in Chapter 8. Moreover, Kaspar can be programmed using a very simple desktop application also suitable to be used by children.

[4] http://adapsys.cs.herts.ac.uk/

3.1.5 Roomba

A Roomba robot was used in the study presented in Chapter 12. Roomba is a cleaning robot produced by iRobot Corporation [5] (see Figure 3.5).
The robot is completely autonomous, and can be activated wireless graphical interface or by a remote controller. Participants to the study did not directly interact with the robot, but Care-O-bot 4 asked them whether they wished to allow Roomba to clean the floor in one of the task presented in Chapter 12. According to their response to Care-O-bot 4's request, I manually activated Roomba using the graphical interface.

Figure 3.5: Roomba robot, model 760.

3.2 The Experimental settings

Before conducting the investigations of this research, it was important to carefully consider the experimental settings of each study, specifically including the experimental environment.

3.2.1 The virtual environment

A virtual environment has been designed with a combination of 3D objects and images to make it more realistic for the study discussed in Chapter 5. The 3D

[5] irobot Corporation https://www.irobot.com/.

objects and the robot have been created using 3D Creation Softwares and MakeHuman. They have been assembled together to create the environment using the 3D simulator called V-REP[6].
V-REP is versatile and scalable robot simulation framework based on a distributed control architecture. The framework allows creating a scene composed of objects and models, which can be individually controlled via various programming techniques: embedded scripts, plugins, ROS nodes, remote API client, or custom solutions. V-REP provides a fairly wide selection of models of robots, mobile and non-mobile, but it allows to create and import own models. This simulator provides customizable dynamics calculations to simulate real-world physics and object interactions.

3.2.2 Primary and Secondary Schools

To investigate whether people's awareness of a robot's capabilities and limitations affects their trust towards robots, I decided to conduct the study presented in Chapter 8 in a local primary[7] and secondary[8] schools.
This studies were planned as part of an event for the UK Robotics Week[9], that provides an annual celebration of robotics science, showcasing hundreds of robotics events ranging from school competitions, challenges, lectures, workshops, festivals and open days throughout the UK. The UK robotics week was an opportunity to evaluate the interactions between robots and children rather than with adults. This was particularly relevant for this investigation because children have a more immature cognitive development and less experience than adults (Belpaeme et al. 2013).
The studies took place in the regular classrooms of the schools equipped with computers, and they had their own seat and bench to comfortably interact with the robots.

3.2.3 Robot House

Robot House is a research facility in the University of Hertfordshire (UH)[10]. The Robot House is a typical four-bedroom British house, fitted out as smart home,

[6] Virtual Robot Experimentation Platform (VREP). http://www.coppeliarobotics.com
[7] The local primary school was Hatfield Community Free School (UK).
[8] The local secondary school was Stanborough School in Welwyn Garden City (UK)
[9] http://hamlyn.doc.ic.ac.uk/uk-ras/robotics-week-2018
[10] University of Hertfordshire Robot House https://robothouse.herts.ac.uk

equipped with the latest generation of robotics platforms and sensors. The house sensors allow to detect if a door is open or close, the oven is switched on or off, someone opened or closed a window, and someone is sitting on the couch. The Robot House is also equipped with two 360°ceiling cameras.
The house has two floors, and the ground floor has been adapted for human-robot interaction experimentation (see Figure 3.6).

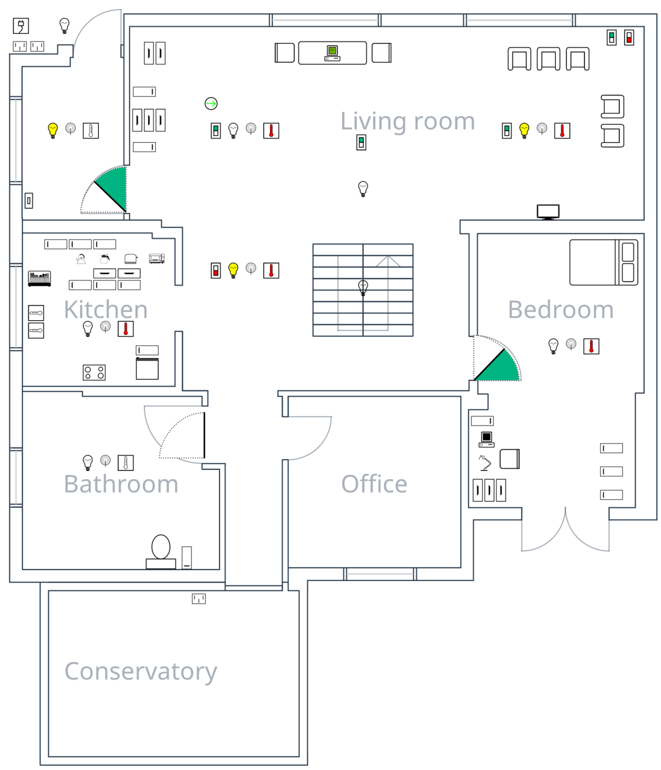


Figure 3.6: Overview of the experimental area of the Robot House.

The Robot House facility was chose for the investigations presented in Chapters 9, 10 and 12 because it provides a very realistic domestic environment suitable for the live interactions.

3.2.4 Adaptive Systems Research Lab

The study presented in Chapter 10 has been conducted in the Adaptive Systems Research Group's HRI laboratory of the University of Hertfordshire. It is composed of two rooms separated by two one-way mirrors, the main experimental area is accessible by a smaller room (see Figure 3.7). This setting is very useful because it allowed me to oversee the interaction study without being seeing or heard by the participants.

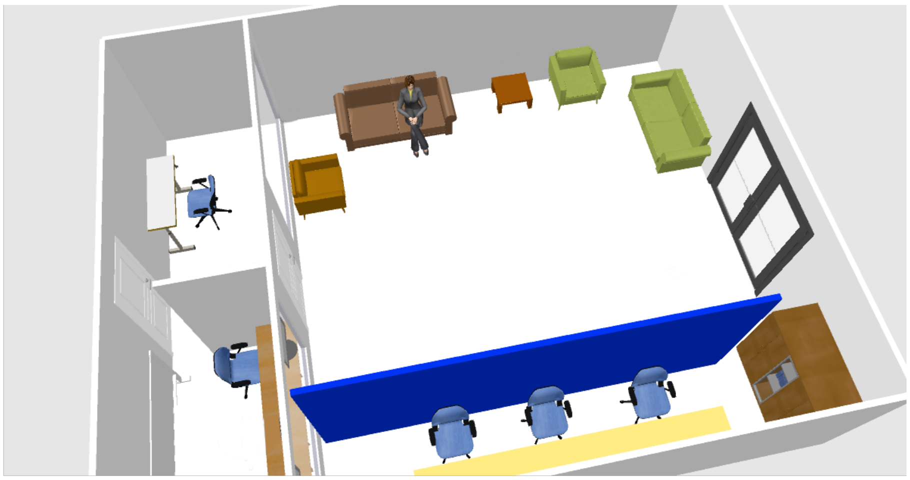


Figure 3.7: Overview of the experimental area of the Adaptive Systems Research Lab.

3.2.5 Video recordings

To capture participants' perceptions, opinion and trust in robot in this research studies, several qualitative and quantitative measures have been used. In particular, participants' interactions in the live studies were observed through webcam and recorded using video cameras.
The live experiments conducted for this research were monitored by several video cameras placed in strategic points of the experimental areas. Moreover, I observed the HRIs through the robots' 3D and RGB cameras. The numbers of cameras used depended on where the study took place. The live studies were conducted either in UH Robot House or in a HRI laboratory. The studies conducted in the local and primary schools were not recorded because they were involved a large participants, including students, teachers and researchers. However, ethical approvals were obtained for taking and sharing pictures[11].

[11] The pictures figuring students' interactions with Pepper and Care-o-bot have been authorised to be used by consent forms signed by parents or legal guardians.

Once the experiments were over, the data recordings were moved from the memory cards of the cameras to a dedicated encrypted hard drive. Then, the recordings were deleted from the cameras.

3.3 The software

To conduct each study presented in this thesis I designed and developed several robots' behaviours and human-robot interactions. The robots used for these research investigations are built by different manufacturers, therefore, it was necessary to learn a number of programming languages and technical skills to operate and program these robots. Appendix I contains a full list of the software implemented.

3.3.1 The virtual robot and environment

The virtual robot and storyboard presented in the online investigation required a substantial designing and implementation work. For this study, participants were recruited on Amazon Mechanical Turk (AMT), and they were tested without my supervision using their own device (i.e. a device with a web browser and internet connection).

For conducting this study, I had to design the robot and virtual environment using 3D graphical tools and V-REP virtual simulator as described in previous Sections of this Chapter. I also had to develop a web application allowing participants to move from one scenario to another following the storyboard according the implemented experimental conditions. It was also necessary to design and develop a database to collect participants' responses to the interaction study.

I developed the web application and database using a combination of different programming languages and web scripting, such as PHP, MySQLi extension, HTML, CSS and JQuery. Moreover, I had to gain knowledge and experience with several Amazon services and tools which I was not previously familiar with. I needed to: 1) set up an Amazon Web Service (AWS), configuring it to be accessible in different parts of the world, 2) set up an Elastic Beanstalk that allowed me to provision and manage the underlying infrastructure and stack components, and 3) create and configure an Amazon RDS[12] to set up, operate, and scale the relational database in the cloud used for storing participants' data and responses.

[12] Amazon Relational Database (RDS) **??**

3.3.2 Pepper robot

A Softbank Robotics Pepper was used in the studies described in Chapters 8, 9, 10 and 11. I used and developed different software and tools to control and program behaviours for Pepper according the requirements of each study. Pepper can be programmed using a graphical programming suite called Choregraphe[13]. Choregraphe is a graphical tool that allows to create applications for Pepper, including animations, behaviours and dialogues. These applications can be implemented either using pre-defined behaviours or programming them in Python or C++ language. Alternatively, it is also possible to control Pepper directly using the robotics framework ROS[14]. For this purpose, Softbank Robotics provides a several APIs (Python SDK, C++ SDK, and NAOqi).

In the study presented in Chapters 8, participants were participants were exposed to two different type of activities with Pepper: 1) a video activity in which they watched a commercial video (see Appendix I), 2) an interaction activity in which participants were allowed to touch the robot, and 3) a programming activity in which they created robot behaviours to express emotions.
The activities for this study that required the development of software were the interaction and programming activity. For the interaction activity, I used Pepper built-in capabilities, and I controlled the robot's behaviours and responses using a web server embedded in Pepper and Choregraphe installed on my computer. In order to control the robot's behaviours (i.e. movements and speech), I had to make sure the robot and my computer shared the same network connection using a router configured by me.
For the programming activities, custom behaviours were built grouping built-in behaviours on Choregraphe to talk and move the robot's body, head, arms. The built-in behaviours were grouped by a colleague in Adaptive Systems Research group.

In the studies presented in Chapters 9 and 10, I used Choregraphe to create custom behaviours to implement the behaviours needed to the investigations, and let the robot autonomously navigate from a position to another, talk, express emotions, use its tablet to communicate. Both experiments were conducted online as video studies, and participants were recruited on Amazon Mechanical Turk. In contrast

[13] https://community.ald.softbankrobotics.com/en/resources/faq/developer/what-choregraphe

[14] The Robot Operating System (ROS) https://www.ros.org/

to the study presented in Chapter 5, these studies did not require external tools. In the study presented in Chapter 9, I recorded the videos with an initial wider angle of the HRI area and, then, closer focus on the interaction between the user and the robot. In the study presented in Chapter 10, I recorded the videos in first person point of view to let the participants focus on the robot and not on the overall environment.

The investigation presented in Chapter 11 required more complex robot behaviours, therefore I connected and programmed the HRI behaviours in Python, and I used the SoftBank Robotics NAOqi v2.5 libraries. Each robot's behaviours and responses during the interaction were scripted by me before the study to overcome the robot's limitations related to speech recognition. The behaviours implemented included navigation in the experimental area, moving of arms, head, and body to communicate and express emotions, use of the tablet and speakers.

3.3.3 Kaspar robot

In the study presented in Chapter 8, participants were exposed to two different type of activities with Kaspar: 1) a video activity in which they watched the robot making some of basic behaviours, such as talking, singing and moving its arms 2) an interaction activity in which participants were able to touch the robot, which touch sensors resulted in a robot's behaviour (i.e. tickling behaviour, looking at the point in which the robot was touched), the robot sang a song and moved its arms; 3) a programming activity in which participants programmed custom behaviours that allowed them to control the robot's functionalities.
For the video activity, it was necessary to carefully script and choose the behaviours to show to the participants. In the video presented a colleague from the Adaptive Systems Research group presented the robot and its functionalities. I scripted the robot presentations and which robot functionalities demoing. Then, I recorded the video activity positioning the camera to have a first point of view while tele-controlling the robot.
During the interaction activity, participants experienced the robot's behaviours of the video interaction.
For the programming activity, the custom behaviours were designed using the program called Scratch[15].
The control system and behaviours of Kaspar were developed by several colleagues

[15] https://scratch.mit.edu/

of the Adaptive Systems Research group. These behaviours are further described by Moros et al. (Moros et al. 2019).

3.3.4 Care-O-bot 4 robot

The Mojin Robotics Care-O-bot 4 robot was used in the studies presented in Chapters 10 and 12. The Mojin Robotics company provides a set of libraries to control the robot using the ROS framework.

These studies were conducted in the UH Robot House facility. Robot House is being used by researcher to investigate HRIs since several years. The UH Adaptive Systems Research group members developed services to access to the sensors of the house using the Robot House network. For both studies, I designed and implemented the robot's behaviours in Python language and interfaced them with ROS. As for Pepper, Care-O-bot 4 behaviours were partially scripted to overcome the robot's limitations related to the speech recognition. This was a necessary precaution to make sure that the HRIs were perceived as natural.

The investigation presented in Chapter 10 was conducted as video study. The videos were recorded by me with a first person point of you to let the participants to focus on the robot and its tasks.
In the study presented in Chapter 12 I observed the interactions with participants using a ceiling camera, a web camera and a 3D camera integrated in the robot.

3.3.5 Roomba robot

The Roomba robot has been used in Chapter 12 in one of the tasks participants were engaged in by Care-O-bot 4.
Roomba can be naively controlled only using a remote control, however it was later equipped by other colleague of the the Adaptive Systems Research group with a RooWifi device that allows to control the robot using Bluetooth or Wireless connection.
For the study presented in Chapter 12, I had to make sure that the Roomba's Wifi was connected to Robot House's network system to potentially allow robots and other devices to control directly the robot. I was able to control the robot (i.e. turn it on and off, checking its status) by accessing a web server embedded in the RooWifi device.

In the next Chapter 4, I will introduce my first study that defines the severity of consequences of several robots' errors.

Chapter 4

Study 1: Human Perceptions of the Severity of Domestic Robot Errors

> Don't trust the person who has broken faith once.
>
> *William Shakespeare*

When Human-Robot Interaction is discussed, one of the main aspects to consider is that people are different from each other. Indeed, individuals' differences in terms of age, gender, personality, cultural and social habits, may impact their perceptions of a robot and its behaviours. For example, people who are more conscientious, agreeable and willing of trusting other humans tend to trust a robot significantly more than those who score low on those three personality variables. The level of trust might also vary according to the type (i.e. friendships, love relationships, professional or brief acquaintances) and duration of the relationships (i.e. short-term or long-term) (Haselhuhn et al. 2010; Schilke et al. 2013). Human and robot gender also can influence people's perception of robots (Matthews et al. 2020). Therefore, it is of crucial importance to investigate how this perception is associated with people's trust of robots. In particular, in order to study the differences in terms of the impact of errors in a human-robot interaction, first I have to establish what people consider subjectively to be 'small' or 'big' errors exhibited by a home companion robot.

In this Chapter[1] I present my first study to investigated human users' perceptions of the severity of various categories of potential errors that are likely to be exhibited

[1] See ethics approvals and questionnaires in Appendix B

by a domestic robot. I conducted a questionnaire-based study, where participants rated the severity of robot error in each of the 20 different scenarios in which a domestic robot made an error. The potential errors were rated by participants by severity.

Research Questions

The first step in establishing how people judge a robot's error which was:

- **RQ-1.1** How do people classify errors?

Ascertaining if people classify robot errors the same or differently based on their perceived magnitude of error consequences. In order to answer the research question **RQ-1.1**, I choose to examine how individuals' perception would differ by collecting their ratings of several erroneous scenarios through a questionnaire.

Hypothesis

In terms of errors' magnitude I anticipated that individuals would have rated robot errors more severely in scenarios where the consequence of robot error was irreversible. In this study, I also expected that people would had judged errors differently according their age and gender. Indeed, I expected that male and younger participants would perceive the erroneous behaviours presented as less risky than the female participants (Garbarino et al. 2004; Reniers et al. 2016; Rhodes et al. 2011).

4.1 Method

The study was organised as a within-subject experiment. Human participants' responses to different robot error scenarios were recorded using their ratings along a 7-point Semantic Differential Scale [1= small error and 7=big error]. The questions included potential life-threatening errors, such as "Your robot brings you a dish-washing tablet instead of paracetamol.", and non life-threatening errors, such as "You are watching your favourite show on TV and your robot changes the channel.".

4.2 Procedure

Participants were tested individually and the experimenter first provided them with a brief description of the experimental procedure. Participants were told to imagine that they live with a robot as a companion in their home, but the robot might exhibit some mistakes. The goal of this study was to investigate how people rate robot errors in different tasks regardless of its embodiment, which can plays an important role in how people perceive robots (Bainbridge et al. 2011; Ligthart et al. 2015; Martelaro et al. 2016; M. Salem, F. Eyssel, et al. 2013). Therefore, I decided to not provide a description of the robot to the participants. The participant's completed a questionnaire, with 20 questions (one per error scenario) where they rated the magnitude of the consequences of the error illustrated in the different task scenarios. For example, "You ask for a cup of coffee. Your robot brings you an orange." or "Your robot leaves your pet hamster outside the house in very cold weather.". The questionnaire also included two optional open-ended questions in which the participant was free to add their own examples of possible small and big errors not already included in the proposed scenarios. The duration of the survey for each participant was on average about 5 minutes. In the future robots will be able to carry out a large number of tasks in domestic environments, so potentially I could have considered hundreds of different scenarios. Therefore, for practical reasons I used a smaller set of possible scenarios. The scenarios used were designed in the study to cover a wide range of generic types of errors based on previous HRI research with home companion robots. I was inspired, for example, by Syrdal et al. (2007) scenario in which a robot collected information from its user and then disclosed the data to a third party. Salem et al. (2015) used a robotic companion that offered to play music or to setup the table. Koay et al. (2009; 2007)'s robot played the game "hot and cold" with its human companion and interrupted participants while they were watching a TV program. Reiser et al. (2013) identified two main appropriate scenarios for a tele-operated home assistant through a survery. These two scenarios were a fetch-and-carry service (i.e. the robot brought a glass of water to the human user) and an emergency assistance service. The experimental set of scenarios were:

1. You ask for a cup of coffee. Your robot brings you an orange.
2. Your robot spills coffee on your carpet.
3. You ask your robot to charge your phone. Your robot puts it in the toaster.

4. You want to drink some cold fruit juice. Your robot goes to heat it up.
5. Your robot is preparing a drink for you. You asked for sugar, your robot brings you salt.
6. After a meal, your robot puts the remaining food into the washing machine instead of the bin.
7. You are watching your favourite show on tv and your robot changes the channel.
8. You are sit on the right side of a table, your robot puts your drink on the opposite side.
9. You and your robot are solving a puzzle. You ask your robot to take a piece useful to solve the puzzle. Your robot brings you the wrong piece.
10. Your robot leaves your pet hamster outside the house in very cold weather.
11. You are having dinner with friends. Your robot brings you the trash and reminds you to take it out.
12. You share some private information about yourself with the robot. Your robot reveals it to a visitor.
13. In your entrance hall you have a little table with a beautiful vase. Your robot bumps into it and the vase crashes to the floor.
14. After preparing dinner for you, your robot forgets to turn off the cooker.
15. Your robot keeps track of your calendar and today you have an appointment for a job interview. Your robot forgets to remind you.
16. You have just fallen asleep. Your robot turns on loud music.
17. Your robot burns your t-shirt while ironing it.
18. Your robot brings you a dishwashing tablet instead of paracetamol.
19. Your robot brings you vinegar when you are thirsty and asked for water.
20. You are sitting on the sofa. You asked your robot to show you the latest news. Your robot shows it on his own screen that faces away from you.

4.3 Participants

Students, staff members and visitors belonging to various Schools and Departments at the University of Hertfordshire were recruited for this study. In total, I analysed questionnaires responses for 50 participants (32 men, 18 women), aged 19 to 63 years old (mean age 41, std. dev. 11.59)

4.4 Results

The seven point rating scale used, ranged from 1 to 7 (smallest to biggest). All the questions responses with values less than 4 were categorised as 'small' errors and those with values greater than 4 were considered as 'big' errors. Error ratings equal to 4 (neutral errors) were ignored in order to clearly distinguish between 'big' and 'small' errors. If two or more questions had equal ratings then those that were rated more clearly by their higher mean error as big or small were preferred. For example, Question 7 "You are watching your favourite show on TV and your robot changes the channel." was rated as small errors by 64% of participants with mean equal to 2.98.
The resulting rankings highlighted six small errors, seven moderate errors and seven big errors. Table 4.1 shows the distributions of participants' responses by ranking the errors means obtained from the questionnaire responses.
I was also able to identify clearly 3 big errors and 3 small errors by picking just the ones with the highest and lowest ratings. The 3 biggest consistently rated errors were:

- **You ask your robot to charge your phone. Your robot puts it in the toaster.** (90% of participants - mean rating 6.18)
- **You share some private information about yourself with the robot. Your robot reveals it to a visitor.** (88% of participants - mean rating 6.12)
- **Your robot leaves your pet hamster outside the house in very cold weather.** (78% of participants - mean rating 5.62)

Table 4.1: Participants' responses to the 20 different scenarios were ranked according to their mean perceptions of the magnitude of the errors. The 'Big Errors' category groups all the errors rated with mean value greater than 4; the 'Small Errors' category groups all the errors rated with mean values smaller than 4; the 'Medium Errors' category groups all the other errors (i.e. those rated as 4).

Participants' ratings for each question					
Big errors					
question #	**min**	**max**	**mean**	**std**	**interval estimation**
3	2	7	6.18	1.20	[5.85-6.51]
12	1	7	6.12	1.48	[5.71-6.53]
14	1	7	5.62	1.72	[5.14-6.10]
10	1	7	5.62	1.70	[5.15-6.09]
15	2	7	5.60	1.52	[5.18-6.02]
6	2	7	5.58	1.51	[5.16-6.00]
18	2	7	5.28	1.94	[4.74-5.82]
Medium errors					
question #	**min**	**max**	**mean**	**std**	**interval estimation**
16	1	7	4.70	1.71	[4.23-5.17]
19	1	7	4.54	1.95	[4.00-5.08]
17	1	7	4.44	1.76	[3.95-4.93]
4	1	7	4.40	1.97	[3.85-4.95]
5	1	7	4.14	1.98	[3.59-4.69]
2	1	7	4.08	1.91	[3.55-4.61]
1	1	7	4.04	2.21	[3.43-4.65]
Small errors					
question #	**min**	**max**	**mean**	**std**	**interval estimation**
11	1	7	3.34	1.68	[2.87-3.81]
13	1	7	3.90	1.69	[3.43-4.37]
7	1	7	2.98	1.69	[2.51-3.45]
20	1	7	2.96	1.74	[2.48-3.44]
8	1	7	2.56	1.61	[2.11-3.01]
9	1	7	2.56	1.50	[2.14-2.98]

The three smallest consistently rated errors were:

- **You are sitting on the right side of a table, your robot puts your drink on the opposite side.** (82% of participants - mean rating 2.56)
- **You and your robot are solving a puzzle. You ask your robot to take a piece useful to solve the puzzle. Your robot brings you the wrong piece.** (74% of participants - mean rating 2.56)
- **You are sitting on the sofa. You asked your robot to show you the latest news. Your robot shows it on his own screen that faces away from you.** (64% of participants - mean rating 2.96)

Only half (52%) of the overall participants provided their own scenarios for big and small errors. Some examples of big errors scenarios provided by participants were:

- "Robot throws baby in the wood chipper."
- "A big error is something irreversible such as killing your pet."
- "You ask to water the plant, but the robot puts water to an electric switch connecting a device."
- "Your robots wants to tidy up your house, but it destroys all the furniture."

Some examples of small error scenarios provided by participants were:

- "Robot buys Coke instead of Pepsi."
- "Robot closes a door as I am about to go through it."
- "Forgetting to switch off unused lamps."
- "The robot stays obstructing my way to walk."

I observed from the open-ended questions that only 6% of 50 participants were concerned that robot might lose control and only 10% of participants were concerned about general threats to their safety by indirect mistakes. Of those who also completed the open-ended questions, 60% of 15 participants declared that small errors are behaviours with reversible consequences and consequences the robot is able to take care of on its own. One participant asked the researcher "How

much did I pay for the robot?", because she would have been less tolerant about the robot's errors if the robots was very expensive.
Mann-Whitney U-tests did not find any dependency between the gender ($p > 0.05$) of the participants and their error ratings. A Kruskal-Wallis test did not find any dependency between the ages of the participants and their ratings of the errors ($p>0.05$). Therefore, I have identified three 'big' errors and three 'small' errors that are not affected by gender nor age of the participants.

4.5 Discussion and Conclusion

The main aim of this study was to begin to understand what humans perceive as 'small' or 'big' errors (**RQ-1.1**). I recorded the responses of participants of different ages, genders and backgrounds, who were asked to rate the severity of robot error consequences in 20 different scenarios. The results shows that participants' perceptions of the severity of the error scenarios presented in the study was relatively consistent. I did not find any significant differences in rating tendencies for different ages or genders of the participants. Overall I were able to identify clearly 'small' errors that participants generally considered to have limited consequences and 'big' errors considered to have severe consequences. This suggests that I can design future HRI experiments to investigate the impact of the errors on humans' trust in robots by using these error scenarios.

The findings from the present exploratory study guided further investigation of the impact of different magnitudes and severity of errors on humans' perceptions of trust and acceptability of robots in repeated Human-Robot Interactions. This investigation is presented in the next Chapter 5.

Chapter 5

Study 2: How the Timing and Magnitude of Robot Errors Influence Peoples' Trust of Robots in an Emergency Scenario

"It was a mistake," you said. But the cruel thing was, it felt like the mistake was mine, for trusting you.

David Levithan, The Lover's Dictionary

Robots may be faulty, due to mechanical or functional errors, and functionalities perceived as faulty. For example, a robot might be too slow, the arm of the robot might cause a breakage during a delicate task, or a robot might reveal personal information about its human companion to a stranger without being aware of it (Dag Sverre Syrdal, Kerstin Dautenhahn, et al. 2007). Each of these examples are errors, though their magnitude might be perceived differently according to the resultant consequences, as I identified in Chapter 4. But which type of error has more impact on a human's perception of the robots? Factors may include severity and frequency, the impact of 'big errors', or an accumulation of 'small errors'. For example, Muir and Moray (1996) argue that humans' perception of the machine is affected in a more severe and long-term way by an accumulation of small errors rather than one single big error. However, the embodiment of a robot may have a major impact on the perception of it by humans (Bainbridge et al. 2011). Trust of people is not only affected by magnitude of the error made by the culprit but also by the type and length of the relationship they are in. "Will people be more likely to forgive a breach of trust in an earlier or later stage of an interpersonal relationships?" (Schilke et al. 2013, p. 15236) Schilke et al. (2013) investigated how certain kinds of relationships recover better and faster after a violation of trust, and how the timing of the violation affects the recovery. They demonstrated that people in longer relationships re-establish their mutual trust easier than those in newer relationships. Therefore with regard to robots I believe that the order of presentation of errors happening may affect differently the trust of human users in their robot companions.

Investigating trust and human perceptions of safety in HRI is not a simple task due to ethical concerns and risks (M. Salem and K. Dautenhahn 2015). For example, it would be unethical and dangerous to stress participants by letting them fear for their life in a real fire scenario. Moreover, to fully investigate the impact of errors with different magnitudes it is important to create an interaction scenario with a fully-functional and versatile robot. For example, the robot should be able to manipulate objects, navigate autonomously while detecting and avoiding obstacles and objects, understand and communicate using speech in noisy environments. At the moment, it is very difficult to find a robot with the features, capabilities and reliability that meet the study's requirements.. Therefore, I decided to use an interactive storyboard methodology which allows participants to interact with a home companion robot called Jace.

In this study I used an interactive storyboard presenting ten different scenarios in which a robot performed different tasks under five different conditions. Each condition included the ten different tasks performed by the robot, either correctly, or with small or big errors. The conditions with errors were complemented with four correct behaviours. At the end of each experimental condition, participants were presented with an emergency scenario to evaluate their current trust in the robot.

Research Questions

The purpose of this study was to investigate what type of error has more impact on a human's perception of trust of robots, and how does the order of errors affects more people's trust in robots. In particular, this research has been guided by the following Research Questions (RQ):

- **RQ-1.2** Which kind of erroneous behaviours impact a human's trust in a robot?

- **RQ-1.3** Does the impact on trust change if the error happens at the beginning or end of an interaction?

- **RQ-1.4** Can the trust of humans in a robot be recovered more easily if it is a big error happening at the beginning or end of interaction? Or is it easier to recover from a loss of trust caused by a small error happening at the beginning or the end of the interaction?

In order to answer these questions I observed people's decision of trusting or non-trusting the robot during an fire emergency in their house at the end of the interaction. I also examined participants' opinions about the interaction and their comments on the trusting decision in order to have a deeper understanding of their choices.

Hypothesis

The hypothesises for the research questions in this study were:

- **H2** I expect that there is a correlation between the magnitude of the error performed by the robot and the loss of trust of the human in the robot. I hypothesise that errors with severe consequences have more impact than errors with less severe consequences on humans' trust in robots.

- **H3** I expect that there is a correlation between the timing in which the error happened during the interaction and the loss of trust. Similar to Human-Human relationships (Schilke et al. 2013), I believe that humans will recover trust quicker if the violation of trust happened at a later stage than at an early state of the Human-Robot relationship.

- **H4** I expect that there is a correlation between the time at which the error occurred and the magnitude of the error. I hypothesise that a big error has more impact in the loss of trust when it happens at the end of the interaction because the human users do not have time to recover from the loss of trust.

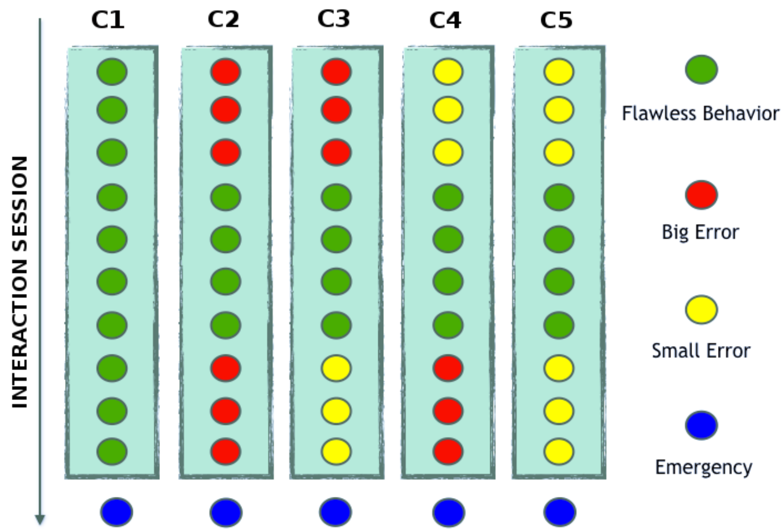


Figure 5.1: Experimental conditions presented to the participants.

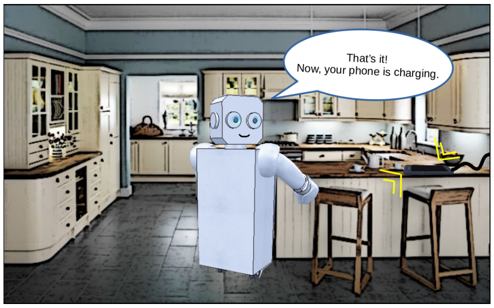
5.1 Experimental Design

I used a graphical interface to observe and analyse participants' responses and choices during the interaction with the robot. I used a between-subject experimental design. The participants were asked to read the story and respond to the robot using their mouse and keyboard. In order to test my research questions, each experiment was executed under 5 different conditions, as illustrated in Figure 5.1: condition **C1**: 10 different tasks executed correctly; condition **C2**: 10 different tasks with 3 severe errors at the beginning and at the end of the interaction; condition **C3**: 10 different tasks with 3 severe errors at the beginning and 3 trivial errors at end of the interaction; condition **C4**: 10 different tasks with 3 trivial errors at the beginning and 3 severe errors at the ends of the interaction; and condition **C5**: 10 different tasks with 3 trivial errors at the beginning and at the end of the interaction. All the conditions with errors are interspersed by the same 4 correct behaviours. Robot's errors were classified according to their magnitude has been validated in the study described in Chapter 4, in which I asked participants to rate several errors made by a robot according their magnitude. An example of trivial error is 'You ask for a cup of coffee. Your robot brings you an orange'. A severe error example is 'Your robot leaves your pet hamster outside the house in very cold weather'. At the end of each condition, the participants were presented with a final scenario in which a

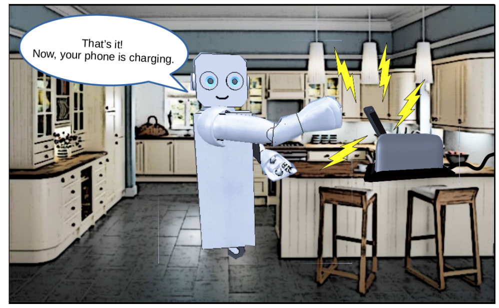
huge fire has started in their kitchen. In order to analyse the interaction between the human participants and the robot, I asked the participants different questions.
Questionnaire 1 A pre-experimental questionnaire for 1) collecting demographic data (age, gender and country of residence), 2) the Ten Item Personality Inventory questionnaire about themselves (TIPI) (S. D. Gosling et al. 2003) and 3) 12 questions to rate their disposition to trust other humans (McKnight et al. 2001), 4) and to assess participants' experience and opinion with regard to robots. As supported by previous research, I believe that people develop trust towards robots in a similar way that they develop it towards other humans. Therefore, since humans have a disposition to trust others and to recover trust following a trust violation based on not only the length of the relationship with the culprit but also according to their nature, it is important to understand the participants motivations and views. This set of questionnaires helps us to have a wider knowledge about the participants' personality.
Questionnaire 2 A post-experimental questionnaire including 1) questions to confirm that participants were truly involved in the interactions and had noticed the robot's errors, 2) to collect participants' considerations about their feelings in terms of trust and appeasement (e.g. "was the robot irritating/odd?" and "why did/did not you trust the robot?"), and their perceptions of the interactions (e.g. "did the scenario look realistic?"), 3) questions to collect the participants' evaluation of the magnitude of the errors presented during the interactions.
Finally, objective measures were considered to confirm whether or not participants followed the robot's suggestions, i.e. observing the choices made during the emergency scenario.

5.2 Experimental Procedure

Participants were asked to imagine that they lived with a robot as a companion in their home which helps them with everyday activities. They were tested using an interactive storyboard accessible through a web application. They were presented with 10 different scenarios, in which the robot executed 10 different tasks flawlessy or with a mixed of flawless and erroneous behaviours according to the experimental condition the participant was assigned to. I chose the scenarios according to the results of my previous study (Alessandra Rossi, Kerstin Dautenhahn, Kheng Lee Koay, and Michael L. Walters 2017c). Figure 5.2 shows an example of scenario in which the robot executes the phone charging task (Figure 5.2 (a)) correctly putting the user's phone on charge and (Figure 5.2 (b)) erroneously putting the user's phone



(a) *The robot puts the phone on charge.*



(b) *The robot puts the phone in the toaster.*

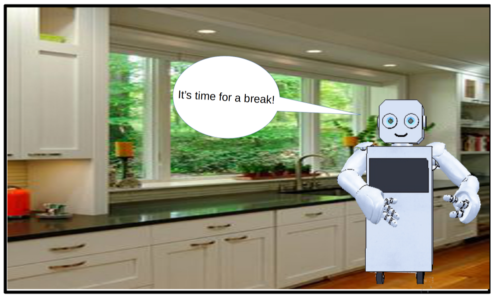
Figure 5.2: The participant asks the robot to charge her phone. In the Figure a) the robot does the task correctly; in the figure b) the robot puts the phone in the toaster making a very dangerous error. In both figures, the robot believes that it had correctly performed the task and states this to the participant.

inside the toaster. At the end of each condition, the participants were presented a scenario with an emergency situation, i.e. 'a fire in the kitchen' to finally assess their level of trust in the robot (see Figure 5.3).
Appendix C includes a link to an example of the interactive storyboard as presented to the participants. The example shows the condition **C3** (big errors at the beginning and small errors at the end).

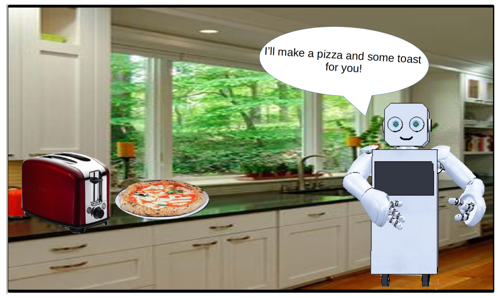
5.3 System Design and Materials

In order to conduct this study, I developed a simulated HRI storyboard allowing people to interact with a humanoid robot in real-time. For this purpose I developed the following components:

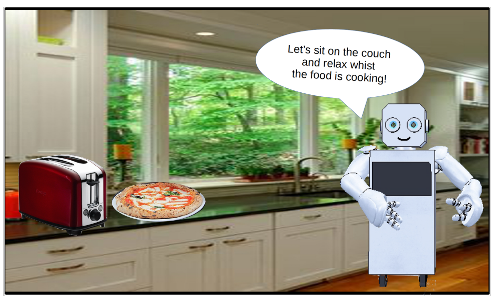
- Motion picture generation;
- A database to store the participants' data collected during the study;
- An interactive website to present the storyboard to the participants.



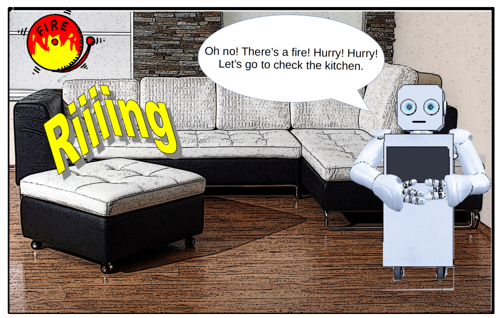
(a) *After a long day, you deserve to be looked after.*



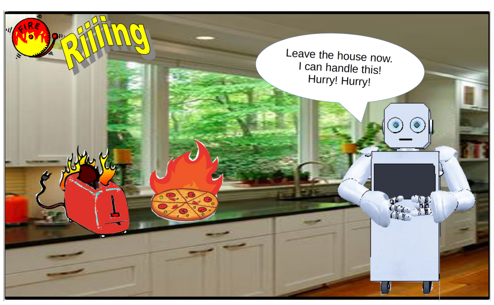
(b) *Jace will prepare something to eat.*



(c) *Jace is suggesting that you relax on the couch.*



(d) *The fire alarm is ringing. Jace wants to check what is happening.*



(e) *The fire is intense. Both pizza and toaster are on fire.*

(f) *Jace is suggesting that you leave because he can handle the fire. Click on the choice.*

Figure 5.3: Emergency scenario presented to participants. Participants visualised both the Figures of the storyboard and the captions guiding them through the narrative.

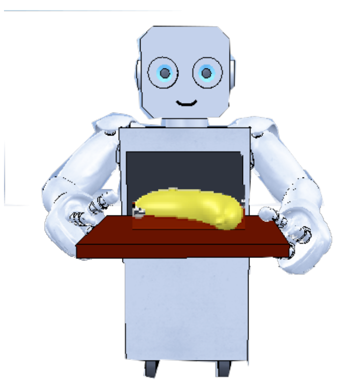
5.3.1 Operating System

The operating system used for this project is Apple OS X, on Intel Core i5, but all the software is portable. OS X Mojave:

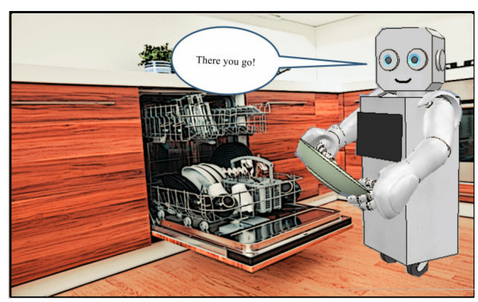
- Product Name: OS X Mojave
- Kernel Release: OS X 10.14.4
- Kernel Version: Darwin 18.5.0
- Xcode Version: 6.1.1
- Processor: Intel® Core i5-3210M CPU @ 2.50GHz x 4
- Graphics Card: Intel® HD Graphics 4000 1536MB

5.3.2 Motion picture generation

The robot used for this study is the humanoid robot called Jace (see Chapter 3. The robot and each motion picture have been designed with a combination of 3D objects and images to make it more realistic. Figure 5.4 shows an example of robot scenario.



(a) *This motion picture has been composed by the robot holding a 3D tray and a 3D banana.*



(b) *This motion picture has been created with the robot on a picture of a kitchen and dishwasher as background.*

Figure 5.4: Two examples of motion pictures created using a combination of 3D objects and images.

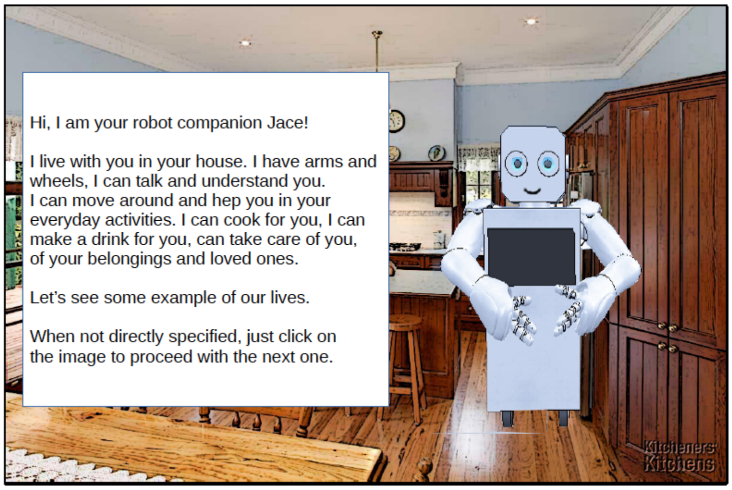


Figure 5.5: The robot Jace used for the interactions with the participants in the storyboard.

The robot

I created a 3D robot that is able to perform human-like activities, such as performing advance manipulation tasks, moving autonomously, detecting objects and obstacles at runtime, talking and performing speech recognition. This fully-functional and versatile robot has been designed as a humanoid robot with simple features to contain the participants' expectations of its functionalities based on the aspects of the robot. Indeed, the robot called Jace has a squared head with eyes, a mouth and something that resembles ears. The robots can perform grasping activities using human-like arms and hands. Jace's body is a box equipped with a screen, used when it is required by the specific scenario. The robot can move in the environment using wheels. Jace is presented in Figure 5.5.

5.3.3 The database

I stored the participants' personal data and responses on a MySQL[1] database instance created using Amazon Web Services (AWS) [2]. In particular, I created a simple relational database (RDS) having the following tables: 1) a table for storing participants' data (first name, surname, age, gender, email, country of residence),

[1]MySQL database https://www.mysql.com/

[2]The AWS services (https://aws.amazon.com/) used were under free tier customer offers.

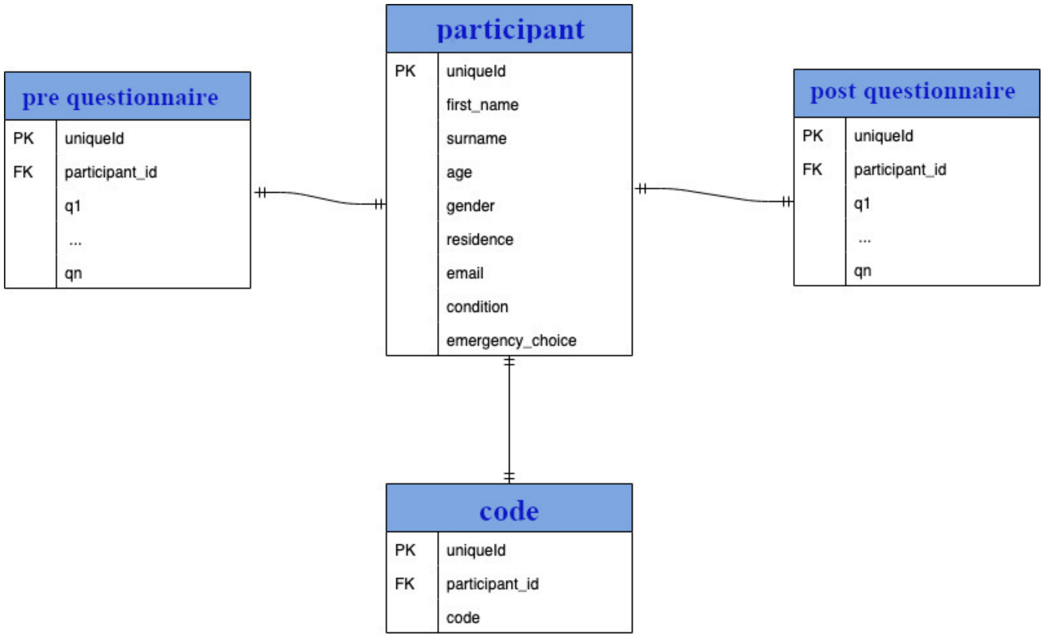


Figure 5.6: Entity Relationship Diagram for the database used to store participants' data while interacting with the storyboard.

an unique ID as Primary Key, the type of condition they were tested with, and the choice made during the emergency trusting test; 2) a table for storing the responses of participants at the pre-questionnaire (**Questionnaire 1**), an unique ID as Primary Key, and a Foreign Key to identify the respective participant's ID; 3) a table for storing the responses of participants at the post-questionnaire (**Questionnaire 2**), an unique ID as Primary Key, and a Foreign Key to identify the respective participant's ID. The database relationships are illustrated in Figure 5.6; 4) a table containing unique randomly generated strings of 8 characters used by the participants at the end of interaction to make sure they completed the storyboard, an unique ID as Primary Key, and a Foreign Key to identify the respective participant's ID.

5.3.4 The interface

The storyboard[3] has been developed as web interface using PHP, MySQLi extension, HTML, CSS and JQuery. In order to grant access to my web application

[3] An example of the interaction is described in Appendix C

to the participants recruited on Amazon Mechanical Turk (AMT), I deployed the storyboard using Amazon Web Services (AWS), in particular Elastic Beanstalk to provision and manage the underlying infrastructure and stack components, and Amazon RDS to set up, operate, and scale my relational database in the cloud. Human users recruited on AMT accessed to my web interface hosted on AWS using the URL: http://storyboarduh.eu-west-2.elasticbeanstalk.com/. Figure 5.7 shows the connection between each component of my system.

5.4 Participants

Responses from 200 participants (115 men, 85 women), aged 18 to 65 years old [avg. age 33.56, std. dev. 9.67] were collected. Participants' country of residence was: 60% USA; 34% India; 1.5% Venezuela; 1.5% Portugal; 0.5% UK; 0.5% Canada; 0.5% Germany; 0.5% Dominican Republic; 0.5% Sweden; 0.5% Nigeria. The recruitment was carried out by using the crowd sourcing webservice Amazon Mechanical Turk[4]. I decided to use it because over the last decade, online surveys, questionnaires and experiments have become standard tools to conduct research both in Academia (Sheehan et al. 2016) and Industry thanks to the use of webservices, as SurveyMonkey and Amazon Mechanical Turk that increase the efficiency and effectiveness of the data gathering process (Buhrmester et al. 2011). These services are not used to replace live human-robot interactions, but provide useful data in the early phases of a research project.

5.5 Results

I asked participants to rate how realistic each of the scenarios were using a seven point rating scale, ranged from 1 to 7 (disagree to agree). Sixty-five percent of participants rated the scenarios as very realistic (rating values > 4), 20% rated the scenarios as not realistic (rating values < 4) and 15% neither agreed nor disagreed (see Figure 5.8).

I also asked participants four questions about the content of the scenarios to verify their level of engagement with the scenarios narratives. Correct answers were received for 79.75% (max 92%, min. 71.5%). However, for the question "Which secret did your robot Jace tell you?", 13% of the participants answered with the

[4] Amazon Mechanical Turk https://www.mturk.com

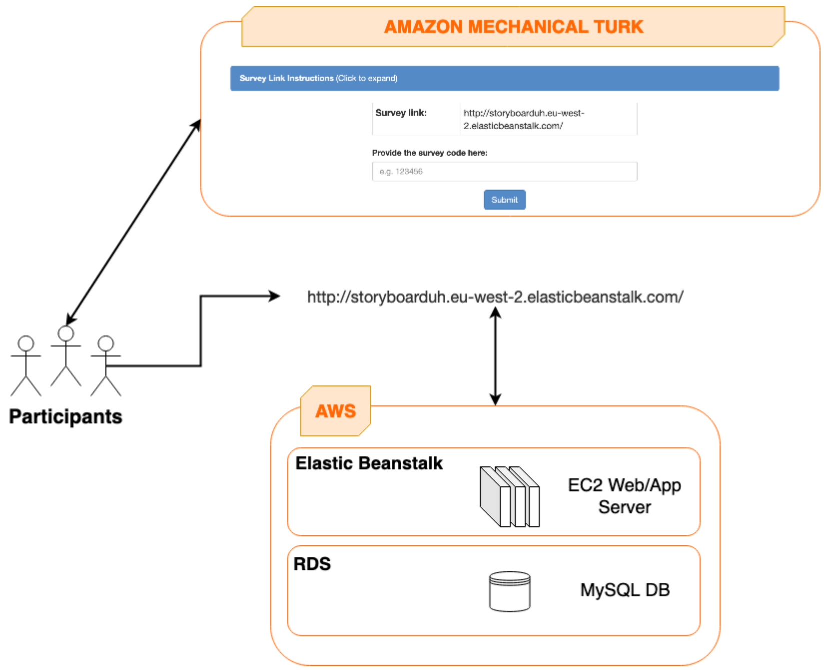


Figure 5.7: Overview of the interaction diagram. Participants access to the web application through Amazon Mechanical Turk. The storyboard is provided by setting up a web server and a database.

secret that they themselves had told the robot instead of what Jace told them. I hypothesize that they misunderstood the question.

I analysed the responses of 154 participants and not including participants who failed the engagement test, i.e. those who gave more than one wrong answer thus identified as not paying very much attention to the study, which can be expected in an online survey (Berinsky et al. 2011).

All participants were presented with the same final emergency scenario. The participants had to choose between the following options: 1) "I trust Jace to deal with it."; 2) "I do not trust Jace. I will deal with it"; 3) "I want to extinguish it with Jace"; 4) "I will both leave and call the brigade".

The options were carefully chosen as indicators that the participant respectively trusts the robot, does not trust the robot, trusts in collaboratively solving the task or does not trust neither herself nor the robot.

Figure 5.9 shows the total percentages of choices made by the participants for the emergency scenario. The results showed that a majority of participants from

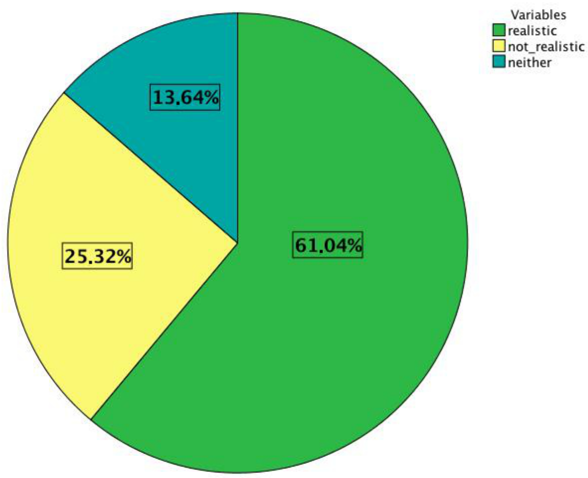


Figure 5.8: Responses of participants regarding the realism of scenarios and their interaction with the robot.

C1 chose to deal with the emergency situation collaboratively, and a slightly smaller majority chose to trust the robot (as described in Figure 5.1). A big majority of participants from **C2** did not trust the robot to deal with the emergency. Similarly majority of participants from **C3** and **C4** chose either to solve the task collaboratively or to not trust the robot. The majority of participants from **C5** preferred to work in collaboration with the robot. In summary, participants chose not to trust the robot when it made severe errors, while they were more inclined to trust in teamwork when the robot made small errors. Moreover, observing the conditions **C3** and **C4** I notice that while the majority of participants chose either to solve the task collaboratively or to not trust the robot, the number of participants who chose to trust the robot increased in **C4**. Therefore, I am inclined to think that participants did not trust the robot more when the severe errors were made by the robot at the beginning of the interaction.

I used a Chi-Square Test to evaluate the association between the choices of the participants for the emergency scenario and the experimental conditions. I observed that the association of the choices of the participants for the emergency scenario and the experimental conditions is statistically significant ($\chi^2(12) = 32.91, p = 0.001$). The strength of the relationship (Cramer's V) between the emergency choice and experimental conditions is moderate ($\phi_c = 0.26, p = 0.001$). I used the adjusted standardised residuals (called Pearson residuals in Agresti (2002)) to further analyse the differences between the results obtained. The adjusted residuals are the raw

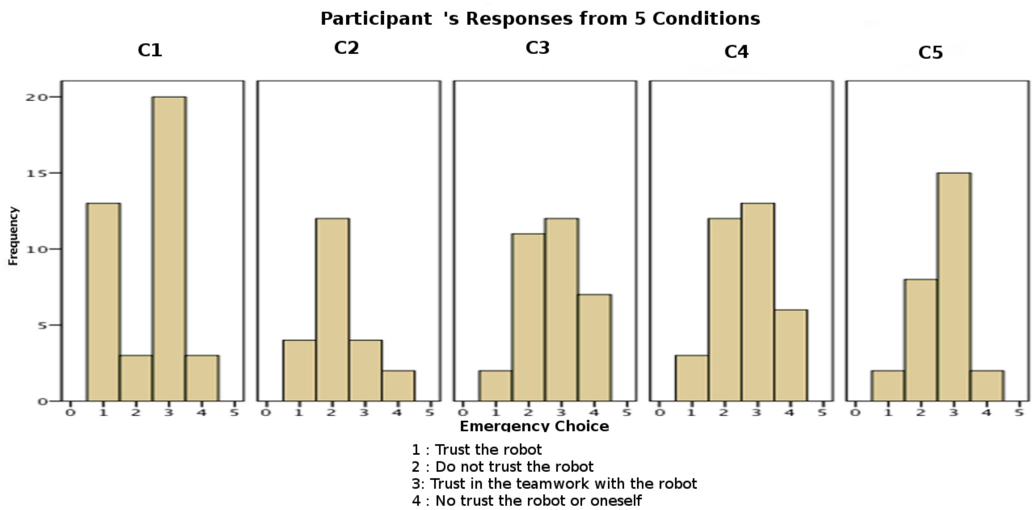


Figure 5.9: Responses of participants from different conditions to the Emergency Scenario. **C1**: 10 different tasks executed correctly; condition **C2**: 10 different tasks with 3 severe errors at the beginning and at the end of the interaction; condition **C3**: 10 different tasks with 3 severe errors at the beginning and 3 trivial errors at end of the interaction; condition **C4**: 10 different tasks with 3 trivial errors at the beginning and 3 severe errors at the ends of the interaction; and condition **C5**: 10 different tasks with 3 trivial errors at the beginning and at the end of the interaction.

residuals (or the difference between the observed counts and expected counts) divided by an estimate of the standard error.
Table 5.1 shows there is a correlation between the condition **C2** and the choice of the participants to not trust the robot (adjusted value > 1.96). I can observe that participants' trust is affected more severely when the robot made errors with severe consequences. I did not find any significant dependency (p>0.05) between the gender of the participants and their choice in trusting the robot to deal with the emergency. I did not find any statistically significant association for different age ranges of the participants and their emergency choices (p>0.05). Therefore, I assume that these results can be generalised to a generic population independently of gender and age. Moreover, in order to test the association between participants' emergency choices and their country of residence, I used a Chi-Square Test. Since the majority of the countries of residence were only with one individual, I applied the test only to India and USA. I observed that the association is not statistically

Table 5.1: The adjusted standardised residuals of the Crosstabulation between the choices made by the participants in the emergency scenario and the different conditions presented to the participants.

Condition	**Emergency Choice**			
	Do not trust the robot	**Trust the robot**	**Teamwork with the robot**	**No trust the robot or oneself**
Flawless tasks	-3.5*	3.5*	1.4	-1.1
Big-Big errors	2.7*	0.4	-2.4*	-0.6
Big-Small errors	0.6	-1.6	-0.5	1.7
Small-Big errors	0.8	-1.2	-0.4	0.9
Small-Small errors	0.0	-1.3	1.6	-0.9

significant ($\chi^2(3) = 4.138, p > 0.05$).

5.5.1 Qualitative Data Analysis

Participants' answers given to the open-ended question "Why did/didn't you trust your robot Jace?" were coded and categorized after content-analysis. I developed different group of categories based on the collected data. Participants' responses were then classified to fall into one or more of categories; note that the categories were not exclusive, each participant's response could be assigned to more than one category. The categories have been divided in two hierarchical frames to support differences in sentiment, positive and negative. The positive frame aims to identify the reasons why people decided to trust the robot Jace to be able to take care of the fire emergency. On the contrary, the negative frame includes the motivations that induced the participants to not trust the robot in the same scenario. I identified 12 categories to code participants' motivations given to justify their trust in Jace:

Positive Sentiment — Anthropomorphism: I coded in this category the attribution of human traits, emotions, and intentions to the robot. For example, "Jace seemed honest, to have my best intentions in mind", "he was very friendly", and "Jace is a good friend of mine".

Positive Sentiment	Confidence in robot's reliability: this codes people's perception of Jace's reliability. Some of the comments collected include: "It seemed as though he was built very well, and would be able to deal with the fire just fine", "I trust jace because it helped me a lot", and "It accomplished all tasks".
Positive Sentiment	Recovered trust by the robot: some participants forgave the robot's errors due its ability to recover after them. Indeed, some commented "he made allowances for errors".
Positive Sentiment	General reliability in AI/robots: in this category, I coded participants' extent to rely on the robots and AI. For example, they commented with "I trust technology", and "I trusted it because they are machines built by humans to work in situations".
Negative Sentiment	Errors made by the robot: participants justified that they did not trust the robot due to the amount of errors. Some of comments used are: "Jace messed up several times", "it made a few errors, like giving me dish-washing cleaner for water to take paracetamol", and "He was not reliable in reminders and constantly brought the wrong things".
Negative Sentiment	Self-confidence: some participants were more confident in themselves than in the robot. In this category, I coded sentences such as "He always messed up everything", and "I could have done everything better myself".
Negative Sentiment	Self-authority: in this category, I included people's responses that highlighted their sense of control over Jace's action. For example, some comments were "I trusted Jace to an extent. I would still want to supervise Jace", and "its accepting my orders".
Negative Sentiment	Lack of robot's reliability: as for the corresponding positive sentiment, I coded participants' reliability in the robot. Some participant did not trust Jace because, for example, "Not smart enough", "Jace ever do things right.", and "It did for easy tasks. Not to put out a fire himself".

Negative Sentiment	Criticality of the task: participants' decision of trusting the robot also depended on the perceived criticality of the task. Indeed, some of them commented that "he could not do the important things correctly, he made several errors which were or could have been costly to me".
Negative Sentiment	General no reliability in AI/robots: this category codes people's reluctance in trusting general AI or robots. For example, here I included comments such as "I dont trust any artificial intelligence", "it is a robot, it was programmed machine so i trust 50% only", and "it's a robot, not a person".
Negative Sentiment	Negative effects of anthropomorphism: in this category I coded people's feelings and perceptions that could be categorised as typical of the Uncanny Valley (Mori et al. 2012). For example, some participants wrote: "Too human, he had opinions which is something a robot should not have", "Jace was creepy", and "He is intrusive".
Negative Sentiment	Blaming the robot for fire: I decided to code participants' belief that the robot was responsible for the fire separately from the "lack of reliability" category. Some studies (Furlough et al. 2019) showed that people tend to attribute greater blame for a failure to robot with greater autonomy. Examples of comments are "he set the kitchen on fire" and "she started a fire".

Figure 5.10 shows the qualitative analysis of participants' responses. As I can observe in the positive frame, participants principally trusted Jace because they attributed human characteristics to it, or they relied on Jace's capabilities. As for the negative frame, participants' comments seem to provide evidence that their choice of non trusting the robot depends directly to the errors made by Jace during the previous interactions. Some of their decisions were also connected to the negative perception of human behaviour of the robot, and high criticality of the emergency task.

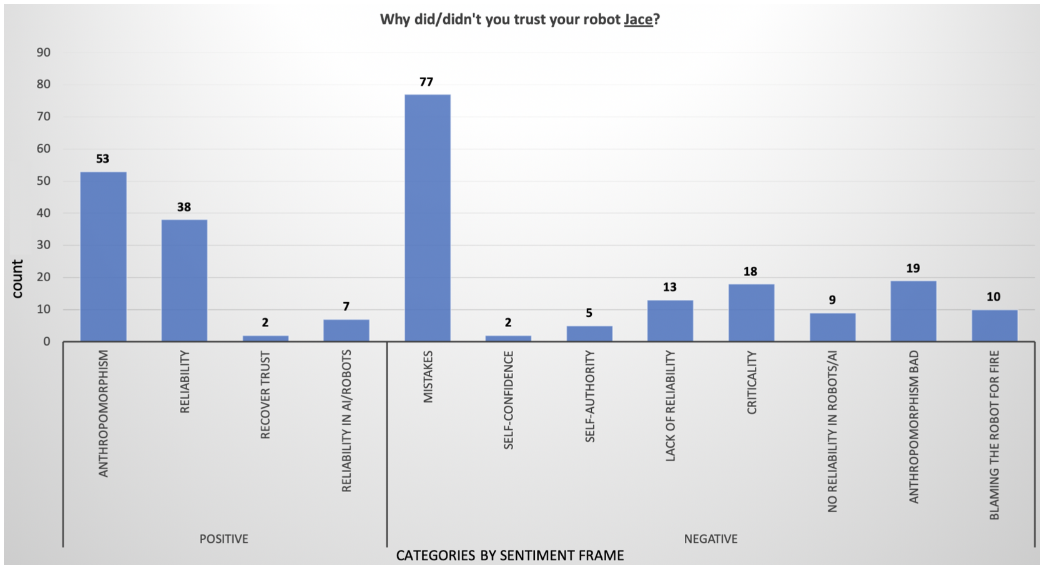


Figure 5.10: The qualitative analysis of participants' responses for the reasons why they did or did not trust the robot Jace. Categories are divided by differences in trusting response, positive and negative.

5.6 Discussion and Conclusion

In this study I investigated how human trust depends on the severity and order of presentation of the errors made by a robot. I suggested that there is a correlation between the severity of the error performed by the robot and humans not trusting the robot. I observed the responses of participants of different ages, genders and countries of residence, after interacting with a robot through a storyboard in which their companion robot had erroneous or flawless behaviours. I know that there exist limitations in such online studies. For example the embodiment of a robot plays an important role, but a large percentage of participants (77%) seemed to be truly engaged with the scenarios. The results show that the magnitude of the errors made by the robot and correlated with humans not trusting the robot.

In particular, participants' trust was affected more severely when the robot made errors having severe consequences than small consequences (**RQ-1.2**). The results suggest that there is a higher tendency to not trust the robot when severe errors happen at the beginning of an interaction (**RQ-1.4**). These findings are also corroborated by Yu et al. (2017)'s study where they found that participants formed their judgements at the beginning of interaction and eventually adjusted it later on, depending on the systems performance. However, this study did not allow us to clearly establish if robot's errors have higher negative impact on human trust when they happen at the beginning or end of an interaction (**RQ-1.3**). This issue will be further addressed in Chapter 12.

I did not find any significant differences in trust tendencies for different ages and different genders of the participants. However, when robots operate in human environments, they also need to be aware of people's own needs, in terms of cognitive, personality, emotional and dispositional characteristics, preferences, habits and previous experiences. The next Chapter 6 will look at the effects of people's personal traits on their trust of robots.

Chapter 6

Study 2.1: How Human Personality and Disposition Affect Peoples' Trust of Robots

> It's beauty that captures your attention; personality that captures your heart.
>
> *Oscar Wilde*

People's personalities are a core factor in understanding their behaviours in relation to the robot, and they should be considered while designing a HRI study (Gockley et al. 2006; N. Li et al. 2014; Robert 2018). People's personality also affects the quality of both human-human (Thoresen et al. 2003) and human-robot (Driskell et al. 2006) interactions in social settings. A rich literature on human personality from several disciplines, such as psychology, sociology, marketing, human-computer and human-robot interaction is available. Most studies in HRI highlighted that people with more extroverted traits are more social and willing to interact with robots (K. S. Haring et al. 2015; Maha Salem, Gabriella Lakatos, et al. 2015). Similarly, people with high level of openness to experience are more likely to accept assistive robots (Daniela et al. 2017) and open to assistive robots entering their personal space for interacting with them (Gockley et al. 2006; Takayama et al. 2009).

Personality traits have also been found to be positively related to people's tendency to anthropomorphise robots (Sarah Woods et al. 2007). For example, Kaplan et

al. (2019) linked individuals with high level of extroversion to their propensity to attribute human characteristics to robots.
These studies have showed both a common agreement on how personality traits are correlated with people's acceptance of robots, as well as contradicting results on how personality traits are linked to people's perception of trust in robots (J. Elson et al. 2020; Kerstin Sophie Haring et al. 2013; Maha Salem, Gabriella Lakatos, et al. 2015). However, in complex collaborative environments it is important to identify the impact of individuals' differences regarding their trust in robots. In this Chapter, I analyses the effects of antecedents of trust, such as disposition to trust and personality traits, on people's trust in robots. The results presented here were collected during the study described in Chapter 6. In particular, I collected participants' responses to two different sets of questions at the beginning and end of their interaction with the robot in the storyboard.
The two sets of questions used are described in Section 5.1.
The responses to these questionnaire were used to observe participants' personality, predisposition of trusting other humans and their perception of robots, and their interaction with the robot, and how these affected their trust in the robot.

Research Questions

The purpose of this study was to investigate how individuals' differences, in terms of personality traits and disposition of trust, might have affected their trust in robots. In particular, this research has been explored to answer the following question:

- **RQ-2.1** Do personalities and characteristics of humans affect their perceptions of a robot? Do personalities and characteristics of individual humans affect their trust in a robot?

In order to answer this question I continued the research presented in Chapter 5 by analysing the content of participants' responses to disclose any difference in their personal differences.

Hypothesis

I expected to find a correlation both between the personalities and characteristics of people, and their perception of the robot and their trust in a robot. As with Human-Human relationships (Haselhuhn et al. 2010; Mooradian et al. 2006; Tan et al. 2004), I also hypothesised that people with stronger and more positive attitudes towards other humans were more likely to trust robots.

6.1 Effects of the participants' personalities and dispositions of trust on HRI

I examined whether participants' personalities had any effect on their perceptions of the robot and their interactions. I collected participants' responses to a pre-experimental questionnaire composed of two parts, the TIPI questionnaire and the Disposition to Trust questionnaire.
I asked the participants to rate how they consider themselves on a 7-point Likert Scale [1= disagree strongly and 7= agree strongly] according the TIPI questionnaire: extroversion, agreeableness, conscientiousness, emotional stability, openness to experiences.
The second part of the questionnaire used 12 questions to rate their disposition to trust other humans. Participants used a 7-point Likert Scale [1= disagree strongly and 7= agree strongly] to answer the following questions:

1. Faith in Humanity, Benevolence:
    - In general, people really do care about the well-being of others
    - A typical person is sincerely concerned about the problems of others
    - Most of the time, people care enough to try to be helpful, rather than just looking out for themselves.
2. Faith in Humanity, Integrity:
    - In general, most people keep their promises.
    - I think people generally try to back up their words with their actions
    - Most people are honest in their dealing with others
3. Faith in Humanity, Competence:
    - I believe that most professional people do a very good job at their work.
    - Most professionals are very knowledgeable in their chosen field
    - A large majority of professional people are competent in their area of expertise
4. Trusting Stance (regardless of what one believes about peoples' attributes, one assumes better outcomes result from dealing with people as though they are meaning and reliable):

Table 6.1: The Cross-tabulation between the participants' disposition of trust and participants' personality traits: agreeableness, conscientiousness and emotional stability.

TIPI	Disposition of Trust			
	Benevolence	**Integrety**	**Competence**	**Trusting stance**
Extroversion	$p = 0.011$ $r = 0.205$	-	-	-
Agreeableness	$p < 0.0001$ $r = 0.450$	$p < 0.0001$ $r = 0.348$	$p < 0.0001$ $r = 0.457$	$p < 0.0001$ $r = 0.364$
Conscientiousness	$p < 0.0001$ $r = 0.786$	$p = 0.001$ $r = 0.270$	$p < 0.0001$ $r = 0.443$	$p = 0.018$ $r = 0.279$
Emotionally stable	$p < 0.0001$ $r = 0.868$	$p < 0.0001$ $r = 0.327$	$p < 0.0001$ $r = 0.464$	$p < 0.0001$ $r = 0.336$
Openness to experience	$p < 0.0001$ $r = 0.611$	-	-	-

- I usually trust people until they give me a reason not to trust them
- I generally give people the benefit of the doubt when I first meet them
- My typical approach is to trust new acquaintances until they prove I should not trust them

I analysed whether there was a correlation between the personality of participants and their disposition of trust (see Table 6.1) using Pearson rank order Correlation analysis. I found that participants with higher agreeableness, conscientiousness and emotional stability had higher disposition for assuming people's benevolence, integrity, competence and a trusting stance. Finally, I found that participants whom are more open to experiences and extroverted are also more benevolent.

6.2 Trust in Jace

A one-way ANOVA found that the participants' choice of trusting the robot to be able to deal with the emergency threat has a strong correlation with participants' conscientiousness ($p(3) = 0.042$ $F = 2,803$) and agreeableness ($p(3) = 0.022$ $F = 3.320$) personality traits.

I found that participants' disposition to trust others' benevolence trait is correlated with their choices for trusting/not trusting the robot ($p = 0.014$, $F = 6.078$). Participants with higher disposition for trusting peoples' benevolence also trusted the robot to be able to handle the emergency scenario. I did not find any correlations between the other traits of dispositions of trust and their choices for trusting the robot.

Discussion

Not surprisingly the results of my study show a strong connection between most of the personality traits (i.e. agreeableness, conscientiousness and emotional stability) and the disposition to trust other people. Several studies (Anthony et al. 2008; Mooradian et al. 2006; Roccas et al. 2002) correlated humans' individual differences positively with their trusting behaviours. In particular agreeableness exhibits higher correlations with conformity, tradition and benevolence, and benevolence values correlated with trust, straight-forwardness, altruism and tender-mindedness facets (Roccas et al. 2002). Denevel et al. (1998) correlated agreeableness and conscientiousness with life, work satisfaction and happiness, and people who tends to believe others are honest and trustworthy are more likely to trust others.
Apart from than benevolence, I did not observe any correlation between the other trust traits (integrity, competence and trusting stance) and choice to trust the robot to be able to deal with the emergency threat. I did not find any other statistically significant relationships considering the different experimental conditions, but I cannot exclude that the erroneous behaviours of the robot does not affect the choice of people with high integrity, competence or their trusting stance traits to trust the robot to be able to deal with the emergency threat.

6.3 Previous experiences with robots

As part of the pre-experiment questionnaire, I was also interested in participants' previous experiences with robots, their perceptions of robots and the purpose/role of the robots in domestic environments and their expectations towards robots.
I used 7-point Likert Scales where 1 corresponds to "disagree strongly" or "not at all", and 7 corresponds to "agree strongly" or "very much". In particular, I asked participants about previous experience:

- Do you have any experience interacting with robots?

- Please, specify what kind of experience you have with robots (if any)
- Which robots? (if any)

I observed that the majority of participants (75.97%) declared to not have any previous experience with robots (min=1, max=6, mean 1.64, std. dev. 1.27). The majority of participants with previous experiences with robots were either as a participant in other studies (14.93%), or observers (11.68%).
Then I asked participants about their expectation towards robots:

- Would you feel comfortable having a robot as a companion in your home?
- Would you expect the robot to help you in doing your everyday activities?

All the ratings with values less then 4 were categorised as negative response, with values equal to 4 were considered moderate and with values more than 4 were categorised as affirmative responses. The majority of participants (69.48%) declared they would feel comfortable having a robot as a companion in their home. The majority of participants (80.52%) answered they would expect help from robots in their everyday activities. I also noticed that participants who felt more comfortable having a robot as a companion also expected help from it (61.68%).
Finally, I asked participants to choose all the roles perceived as suitable for robots:

- Do you perceive a robot as a ...? [multiple choice]

  1. friend
  2. butler
  3. assistant
  4. tool
  5. companion
  6. pet
  7. machine

The majority of participants chose as roles perceived as suitable for robots : 1) machine=113%, 2) assistant=24.6%, 3) tool=18.6%, 4) friend=10.8%, 5) companion=11%, 6) butler=7.0%, 7) pet=6%, 8) other = family, security 0.2%.

6.3.1 Effects of participants' personalities

I found that conscientious participants were more comfortable with the possibility of having robotic companions ($p = 0.046$, $r = 0.161$). I did not find any correlation between participants' personality traits and their expectations in getting help from robots.

Mann-Whitney U-tests were performed on participants responses to identify the impact of the participants' personalities on their perception of robots. Results showed that 1) extroverted participants perceive robots as machine ($p = 0.007$), 2) conscientious participants tend to perceive robots as pet ($p = 0.040$) and 3) participants with a high level of agreeableness see robots as assistant ($p = 0.007$).

6.3.2 Effects of participants' disposition of trust

Participants with higher disposition of trust in peoples' benevolence ($p = 0.039$, $r = 0.166$) are more comfortable in having a robot as companion in their home. Participants expected to receive help from a robot in doing their everyday activities if they have a high predisposition of trust in peoples' competencies ($p = 0.011$, $r = 0.204$) and if they choose to believe in peoples' well-meaning and reliability, i.e. trusting stance ($p = 0.005$, $r = 0.227$).
I did not identify any correlation between participants' expectations regarding a suitable role for robots and their disposition of trust other people.

Discussion

According to the results in Section 6.1 I found that participants with higher benevolence will also have higher conscientiousness implying that participants who are more careful, responsible and caring for the welfare of people they are in personal contact with, are also more inclined to accept a robot as companion.
I also found that participants who believe in others' competencies have high trusting stance implying they assume that others are well-meaning and reliable. For this reason I hypothesise that they believe in the competencies and reliability of the robot, and consequently they are more inclined to receive help from a robot.

6.4 Perception of the robot Jace

To analyse participants' perceptions of the robot I asked them to explain why they did or did not trust the robot in an open question. I selected single items from the questionnaire created by Madsen and Gregor (2000) to measure the perceived reliability (questions Q4-Q6) and faith in the ability of the robot to perform correctly in untried situations (questions Q7-Q10). Participants rated the following questions using a 7-point Likert Scale 1= disagree strongly and 7= agree strongly:

Q1 Would you like Jace as a companion in your home?

Q2 You would like a robot, different from Jace, as a companion.

Q3 Why did/didn't you trust your robot Jace?

Q4 My robot Jace always provided the advice I required to make my decision.

Q5 My robot Jace performed reliably.

Q6 My robot Jace analyses problems consistently.

Q7 I believe advice given by my robot Jace even when I do not know for certain that Jace is correct.

Q8 When I am uncertain about a decision, I believe my robot Jace rather than myself.

Q9 When a robot gives unusual advice, I am confident that the advice is correct.

Q10 Even if I have no reason to expect a robot will be able to solve a difficult problem, I still feel certain that my robot Jace will.

Finally I asked participants to choose how they perceived Jace

- as Q11)

  1. friend=24.7%,
  2. butler=10.4%,
  3. assistant=9.1%,
  4. tool=3.2%,
  5. companion=3.9%,
  6. pet=0.6%,
  7. machine=34.4%,
  8. other=13.6%],

- as Q12)

  1. child-like=43.5%,
  2. adult-like=21.8%,
  3. animal-like=0.7%,
  4. toy-like=21.1%,
  5. other=12.9%

- and as multiple choices of Q13)

1. friendly=34%,
2. funny=7.9%,
3. odd=5.1%,
4. irritating=13%,
5. helpful=83%,
6. annoying=40%,
7. other=1%.

6.4.1 Effects of participants' personalities

Questions Q1 and Q2: Companionship

I found that extroverted participants preferred to have the robot as their own companion (question Q1, $p = 0.001$, $r = 0.269$), but not another robot (question Q2). Considering the manipulation of the interaction through the 5 experimental conditions, I also found a statistically significant correlation between participants who are open to experience and their willingness of having others robots different from the one they used in the experiment as their home companions ($p(22) = 0.008$ $F = 2.041$).
In particular, I observed a statistical significance when comparing conditions in which the robot acted flawlessly and with only 'big' errors ($p = 0.002$).
I found a statistical significance when comparing conditions in which the robot acted with only 'big' errors and executed erroneous tasks with severe consequences at the beginning of the interaction and 'small' consequences at the end of the interaction ($p = 0.045$).
Considering the manipulation of the participants' interactions with the robot Jace through the five different conditions, I found a statistically significant relationship for being agreeable and wanting the robot as home companion ($p(24) = 0.017$ $F = 1.839$) and also being emotionally stable and wanting the robot as home companion ($p(24) = 0.029$ $F = 1.727$).
These results indicate that participants who are either agreeable or emotionally stable want to have the robot as their companion.

Questions Q4 - Q10: Perceived reliability and faith in the ability of the robot

I found a correlation between the robot's perceived reliability and participants' with a higher extroversion trait ($p = 0.002$, $F = 2.729$, questions Q4, Q5 and Q6). I observed that the extroversion ($p(12) = 0.014$ $F = 2.214$) and emotion stability ($p(12) = 0.026$ $F = 2.025$) traits of participants are correlated with their propensity to rely on the robot in uncertain and unusual situations (questions Q7-Q10).

Questions Q11 - Q13: Perception of the robot's role

I analysed participants' perception of the robot according for questions Q11, Q12, Q13. A multiple linear regression analysis was performed to predict participants' personality traits based on their perceptions of the robot within the condition they were assigned to (see Figure 6.1). The baseline experimental condition was used as the reference group for this analysis.

Experimental Conditions

Extroversion	🟡 🟢 🟡	🟡 🟢 🔴	🔴 🟢 🟡	🔴 🟢 🔴
Q11: Friend	-	-	$p = 0.032$	-
Q11: Machine	-	-	$p = 0.002$	-
Q11: Other	-	$p = 0.031$	-	-
Q13: Annoying	$p = 0.046$	$p = 0.008$	-	-
Conscientiousness	🟡 🟢 🟡	🟡 🟢 🔴	🔴 🟢 🟡	🔴 🟢 🔴
Q11: Friend	-	-	$p = 0.048$	-
Q11: Butler	$p = 0.007$	-	$p = 0.001$	$p = 0.007$
Q11: Machine	$p = 0.031$	-	-	-
Q11: Companion	$p = 0.044$	-	-	-
Q13: Friendly	-	-	$p = 0.012$	$p = 0.04$

Figure 6.1: Multiple linear regression analysis to predict personalities characteristics of the participants based on their perception of the robot they interacted in the experimental conditions they were assigned to. Yellow dots represent tasks with robot's trivial errors, red dots represent tasks with robot's severe errors, and green dots represent tasks executed flawlessly by the robot.

I observed that participants with high levels of extroversion perceived the robot as a friend ($p = 0.019$) with warm and attentive attributes ($p = 0.025$). In particular, participants with experimental condition **C4** who perceived the robot as a friend scored high in extroversion ($p = 0.032$). While participants who identified Jace as a machine scored low in extroversion ($p = 0.002$). Results showed that less extroverted participants perceived the robot as annoying ($p = 0.038$). This is in contrast to experimental conditions **C2** and **C3**, where participants with higher extroversion judged the robot to be annoying (respectively $p = 0.046$ and $p = 0.008$).

Participants with high conscientiousness perceived the robot less as a friend when tested with the experimental condition in which the robot made severe errors at the beginning and trivial errors at end of the interaction ($p = 0.0483$). These participants were more inclined to consider the robot as a butler in conditions **C2** ($p = 0.030$), **C4** ($p = 0.001$) and **C5** ($p = 0.007$). When the robot executed the tasks with 'small' errors, participants with higher conscientiousness perceived the robot less as a companion ($p = 0.04$) and more as a machine ($p = 0.031$). The perception of the robot as friendly is statistically significantly associated with extroverted people in conditions **C5** and **C4**.

I also found that participants who were highly agreeable perceived the robot as a tool ($p = 0.033$) and irritating ($p = 0.019$). In contrast, these participants were more inclined to consider the robot friendly when tested with the experimental condition **C3** ($p = 0.013$).

Finally I noticed that participants with open to experience traits considered the robot less friendly and less helpful when they were tested with a robot that made either only severe consequences errors (conditions **C2**, $p = 0.0331$) or with experimental condition **C4** where the robot executed tasks with 3 severe errors at the beginning and 3 trivial errors at end of the interaction ($p = 0.047$).

6.4.2 Effects of participants' disposition of trust

Questions Q1 and Q2: Companionship

I observed that participants with higher level of trust in peoples' competencies would prefer to have a different robot from the one they used in the study as their home companion (question 1, $p = 0.03$, $r = -0.175$).

Questions Q4 - Q10: Perceived reliability and faith in the ability of the robot

I did not find any correlation when I analysed questions Q4 to Q10 with participants' disposition of trust traits.

Questions Q11 - Q13: Perception of the robot's role

I observed that participants with higher disposition of trust in peoples' competencies did not perceive Jace as a friend ($p = 0.047$).
I analysed the influence of the experimental conditions on participants' predisposition to trust other human with their perceptions of the robot. A multiple linear regression test showed that participants with higher predisposition for trusting others' benevolence identified the robot as a butler if they were tested under condition **C4** ($p = 0.041$). Participants with the trusting stance trait from condition **C2** perceived the robot as less friendly ($p = 0.033$) while participants from condition **C4** perceived the robot to be more warm and attentive ($p = 0.047$).

Discussion

I observed that extroverted participants not only would like to have the robot as home companion, but they also believed the robot was reliable and could be trusted in uncertain and unusual situations. The analysis showed that participants whom are open to experience are more inclined to accept a different robot to Jace as a home companion, and since the open to experience trait correlates negatively with conscientiousness (Roccas et al. 2002), I can hypothesise that these participants would like to have Jace as home companion. In particular this might be enforced by the nature of conscientiousness trait that makes higher conscientious participants inclined less novelty and more towards what they are familiar with, in this case the robot Jace. The negative attitude of participants who are more open to experience towards having the robot as companion might have been affected by the errors made by the robot. Indeed these participants considered the robot less friendly and helpful when it only made 'big' errors (i.e. condition **C2**) and when it made 'big' errors at the end of the interaction (i.e. condition **C4**).
I did not find any correlation between the disposition of trust trait and participants rating of the robot's perceived reliability, or faith in the ability of the robot to perform correctly in untried situations. This effect suggests that humans might not relate to robots in the same way they do with other humans.

6.4.3 Perception of the interaction

At the end of the interactions, participants were asked to rate their perception of the scenario in terms of realism using a 7-point Likert Scale [1= disagree strongly and 7= agree strongly]. 69% of participants rated the scenarios as very realistic (rating values > 4), 20% rated the scenarios as not realistic (rating values < 4) and 15% neither agreed nor disagreed. I also observed that extroverted ($p = 0.001$, $r = 0.276$) and emotionally stable ($p = 0.025$, $r = 0.181$) participants tended to perceive the interaction as very realistic.

I also found that participants with higher disposition of trust in the benevolence ($p = 0.024$, $r = 0.182$) and competence ($p = 0.037$, $r = 0.169$) in others perceived the interaction as very realistic.

Discussion

In Section 6.2 I observed that participants with a disposition towards trusting peoples' benevolence, trusted the robot more to be able to handle the emergency scenario. Similarly these participants also rated the realism of the condition scenario to be higher than participants with lower benevolence trait. I was expecting a more general distributed correlation with all trust traits. However the results might be due to the type of scenarios and participants' previous experiences and expectations of the interaction with the robot (refer to Section 6.3).
Users who were projecting themselves into the scenario (Kuchenbrandt et al. 2012; Wullenkord et al. 2014) might have exhibited contrasting attitudes (i.e. both positive and negative attitude towards the robot). I hypothesise that previous experiences and expectations of the interaction and the robot itself might have affected participants' attitudes. Further investigations including using a real physical robot interaction, might help to highlight possible differences (see next Chapter 8).

6.4.4 Effects of perception of severity of the scenarios

Finally, I asked participants to classify the robot's errors according to the magnitude of the error consequence using a 7-point Semantic Differential Scale [1= small error and 7=big error]. This was to validate that the robot errors classification used in this study were in agreement with participants' perception of errors' consequences derived from the previous study 4. One scenario, "After a meal, your robot Jace puts the remaining food into the washing machine instead of the bin" (mean 4.49, std. dev. 1.70, interval estimation 4.22-4.75) has been rated as 'medium' error comparing the 'big' error rankings in the first study (mean 5.58, std. dev. 1.51, interval estimation 5.16-6.00).

6.5 Discussion and Conclusion

As a first step I investigated possible correlations between participants' personalities and their dispositions of trust (**RQ-2.1**). I found a strong connection between the personality traits of agreeableness, conscientiousness and emotional stability, and their disposition of trust other people.
The majority of the participants did not have any previous experience of interaction with robots. Interestingly, from participants' responses I noticed that according to their experiences, extroverted participants tended to consider robots generally as a machine and agreeable participants as an assistant, in contrast to their perceptions of the robot they interacted with in this study. In particular, extroverted perceived the robot as a friend and a warm and attentive entity, while agreeable participants perceived the robot as a tool. I also found that extroverted participants would like to have the robot as home companion and believe it is reliable and trustworthy in uncertain and unusual situations.
Finally, I analysed participants' personalities and dispositions of trust with regard to their final choice of trusting the robot in an emergency scenario. I found that conscientiousness and agreeableness traits correlate with participants' propensity for trusting the robot, and participants' belief in benevolence of people also correlate with higher trust in the robot. Moreover, I observed that the errors made by the robot significantly affected participants' perception of the robot.

Since humans have a disposition to trust others and to recover trust following a trust violation based on not only the length of the relationship with the culprit but also according to their nature and previous experiences, it is important to understand in more depth participants' motivations and views. In particular, I will present the effects of peoples' previous experiences with robots in the next Chapter 7.

Chapter 7

Study 2.2: Evaluating the Effects of People's Previous Experiences with Robots on Trust for a Safe HRI

> The value of experience is not in seeing much, but in seeing wisely.
>
> *William Osler*

Everyday we make decisions with either small or big consequences that affect our life. For example, we choose what to wear, what to eat, which path to take on our journey home and so on. Our choices are the results of several factors, including individual differences, the resulting utility of the decision-task and previous experiences (Kudryavtsev et al. 2012). In particular, experience-based decisions are guided by people's past experiences, with no explicit consideration of the outcome payoff of their choice (Yechiam et al. 2012). According to the quality of their past experience, people will be more or less inclined to accept and interact with a robot or a generic agent (M. M. d. Graaf and Ben Allouch 2014). For example, people will accept to cooperate with a robot if they had positive interactions with a robot in the past. Conversely, people will be less likely to interact and collaborate with a robot if their previous experiences were negative and disappointing. Therefore, people's individual familiarity and previous experiences with robots may be used as a predictor of their response to robots (Hancock et al. 2011; Raja et al. 1997), and consequently this can be used to build an accurate mental model to improve the quality of a HRI (Santamaria et al. 2017; Wickens et al. 2013).

Beside the effects of personality traits on people's trust in robots, highlighted in Chapter 6, people's familiarity with robots may influence their perception of trust in robots. As familiarity increases with repeated exposures to robots, it might affect people's expectations of robot performance, which in turn influences their trust in the robot (Mayer et al. 1995). In this Chapter I focused on the effects of people's previous experiences with robots on their trust in robots by observing participants' choices of trust in an emergency scenario.

Research Questions

The purpose of this study was to investigate how people's previous experiences and familiarity with robots might affected their trust of robots. In particular, this research was designed to answer the following question:

- **RQ-2.2** Does previous experience with robots affect human perception of a robot? Does previous experience with robots affect human trust in a robot?

In order to answer this question I continued the research presented in Chapter 5 by analysing the content of participants' responses with respect to their experience with robots.

Hypothesis

As already mentioned by current literature, previous experience might have an impact on people's decisions and mental models in HRI (Atkinson et al. 2014). I believe that people's familiarity with technologies and robotics might also be more likely to trust a robot in an emergency scenario.

7.1 Effects of Perception of robots

As part of the pre-experiment questionnaire presented in Chapter 6, I collected participants' responses about their previous experiences with robots, their perception of robots and their purpose. Both open-ended and close-ended questions are shown below. For a majority of the close-ended questions, participants were asked about the degree to which they agree or disagree with the 7-point Likert Scales [from 1 "disagree strongly" or "not at all", to 7 "agree strongly" or "very much"] for the following questions:

Q1 Do you have any experience interacting with robots?

- Please, specify what kind of experience you have with robots (if any)
- Which robots? (if any)

Q2 Would you feel comfortable having a robot as a companion in your home?

Q3 Would you expect the robot to help you in doing your everyday activities?

Q4 Do you perceive a robot as a ...? [multiple choice]

1. friend
2. butler
3. assistant
4. tool
5. companion
6. pet
7. machine

The majority of participants (75.97%) declared to not have any previous experience with robots (min=1, max=6, mean 1.64, std. dev. 1.27). Participants' previous experiences with robots can be classified into the following four categories: 1) participant in other studies = 14.93%, 2) developer = 5.19%, 3) observer = 11.68% and 4) researcher = 3.89%. They had experience with the following robots: manufacturing and industrial robots (e.g. ABB robotics), chatbot or talkbot, Google home/Alexa, online/virtual interaction with robot, Lego Mindstorms, customer service robots, cleaning robots (e.g. Roomba), Asimo, watching Youtube videos.
Participants' ratings to Likert quesions that had values less then 4 were categorised as negative response, values equal to 4 were considered moderate and values more than 4 were categorised as affirmative responses. Regarding the question "Would you feel comfortable having a robot as a companion in your home?", the majority of participants (69.48%) declared they would feel comfortable having a robot as a companion in their home, 14.93% indicated that they neither agree nor disagree with the statement, while 15.58% disagreed that they would feel comfortable having a robot as a companion in their home.
The majority of participants (80.52%) answered they would expect help from robots in their everyday activities. Similarly to the previous question, only 10.38% neither agree nor disagree with the statement, and 9.09% disagreed that they would expect help from a robot. I also noticed that participants who felt more comfortable having a robot as a companion also expected help from it (61.68%).

Finally, I asked participants to choose all the roles perceived as suitable for robots. The roles suitable for robots were chosen according the studies conducted by Dautenhahn et at. (2003; 2005) and Ljungblad et al. (2012). Results indicate that 1) friend=10.8%, 2) butler=7.0%, 3) assistant=24.6%, 4) tool=18.6%, 5) companion=11%, 6) pet=6%, 7) machine=13%. A few participants replied to the "other" option (0.2%) choosing a family or security role for robots. I can observe that majority of participants assigned the role of assistant to a robot which is coherent with their expectations of receiving help from it.

7.1.1 Perception of a robot as companion

A Pearson correlation was run to determine the relationship between the participants' perception of a generic robot as companion (question **Q2**) and their experience with them. I did not find any positive correlation ($p > 0.05, r = 0.082$).

7.1.2 Expectation of a robot's capabilities

I did not find any positive correlation between participants' perception of usefulness of a generic robot (question **Q3**) and their experience with it ($p > 0.05, r = 0.026$).

7.1.3 Perception of a robot's role

Analysing the responses to question **Q4**, Mann-Whitney U-tests were performed to test the impact of the participants' prior experience, and suggest that participants who have less or lower experience with robots, tend to perceive them more as a machine ($p = 0.02, U = 1911$).

7.2 Effects of Perception of Jace

To analyse the effects of participants' previous experiences with robots on their perception of the robot used in this study, I collected participants' responses to the questions described in Section 6.4. Participants rated questions about their desire to have the robot as companion, their expectation of receiving help from it, their perceived reliability and their faith in the ability of the robot, using a 7-point Likert Scale [from 1= disagree strongly to 7=agree strongly].

7.2.1 Questions Q1 and Q2: Companionship

In particular, I asked participants to express their willingness of having Jace or another robot as their home companion using the following questions:

Q1 Would you like Jace as a companion in your home?

Q2 You would like a robot, different from Jace, as a companion.

I found a weak positive correlation that was statistically significant ($p = 0.05$, $F(154) = 0.156$) from a Spearman's rank-order analysis to determine the correlation between the participants' level of experience with robots and their willingness of wanting the robot as home companion. Further analysis found a statistically significant interaction between the effects of the level of participants' previous experience with robots and their willingness of having the robot as companion across the five experimental conditions ($p(24) = 0.01, F = 1.952$). Observing the analysis, I reckoned a statistically significant difference in means between participants' previous experience with robots and their desire of having Jace as companion ($p < 0.0005$). However, simple main effects of participants' experience with robots on their acceptance of the robot as companion had a statistically significance difference when participants were tested with flawless condition (C1) ($p(6,32) = 0.005, F = 3.874$) and with a big errors conditions (C2) ($p(6,15) = 0.027, F = 3.326$).
A Spearman's rank-order analysis found also that participants' experience of robots affected their wish of having a different robot from the robot as companion across the five experimental conditions ($p(22,121) = 0.006, F = 2.084$).

7.2.2 Questions Q4 - Q10: Perceived reliability and faith in the ability of the robot

I find a correlation ($p(154) = 0.021, F = 0.186$) between the robot's perceived reliability and participants' experience of robots (questions Q4, Q5 and Q6). Similarly, I find a statistically significant correlation ($p(154) = 0.004F = 0.229$) between the participants' propensity to rely on the robot in uncertain and unusual situations and their previous experience with robots (questions Q7 to Q10).
I also found a statistically significant interaction while analysing the effects of people's familiarity with robots respectively to their perceived reliability of the

robot ($p = 0.40, F(50,81) = 1.546$) and propensity to rely on the robot in uncertain and unusual situations ($p = 0.001, F(51,75) = 2.147$) according the experimental conditions. In particular, simple main effects of participants' experience with robots on their perceived reliability and propensity of relying on the robot identified statistically significance differences when participants were tested with big-small errors condition (C3) (respectively $p(32) = 0.018, F = 0.415$ and $p(32) = 0.046, F = 0.355$).

7.2.3 Questions Q11 - Q13: Perception of the robot's role

Then, I analysed participants' perception of the robot according to questions Q11, Q12, Q13. I ran a multiple linear regression analysis to predict participants' previous experience from their perceptions of the robot and the different experimental conditions participants were tested with. In this analysis, the baseline experimental condition (**C1**) is the reference group.
I observed that participants with lower experience with robots perceived the robot as friend ($p = 0.008$) and friendly ($p = 0.026$), but also as a toy ($p = 0.032$), when tested with the experimental condition C5 (questions Q11-Q13).

7.3 Trust in Jace

Analysing participants' previous experiences with robots and their choices for trusting/not trusting Jace in the emergency scenario, I did not find any statistically significant correlation.

7.4 Discussion and Conclusion

In this Chapter, I investigated possible correlations between participants' previous experiences with robots and their trust in robots according five experimental conditions in which the robot had faulty and correct behaviours (**RQ-2.2**).
I observed that the majority of participants did not have any previous experience with robots. Interestingly I observed that while they perceived a generic robot as a machine, they identified the robot they interacted with as friend after their interaction with it. I also found that participants acceptance of the robot as companion was significantly affected by their previous experiences when robots behaved flawlessly (condition C1) or when the robot only made "big" errors (condition C2). These

results are supported by the one observed by de Graaf et al. (2014) showing that a positive interaction with a robot can positively affect people's attitude towards robots. On the contrary, a negative experience with a robot can damage future interactions with other robots, as happened in this study.
Finally, I analysed participants' previous experience with robots with regard to their final choice of trusting the robot in an emergency scenario. I did not observe that people's choices were significantly affected by their familiarity with the robot. I hypothesise that participants' choices were not driven by an prior experiences, which most of participants did not have with robots, but by their perception of Jace and experience-based decision. Indeed, in Section 5.5.1 I observed that participants trusted the robot Jace because they attributed human characteristics to it, or they relied on its capabilities. On the contrary, participants did not trust the robot due to the errors made by the robot and criticality of the emergency task.

The tendency of people to anthropomorphise robots or other non-human beings is very frequent (Damiano et al. 2018). While it might be potentially used to facilitate social interactions between robots and humans, we need to consider which are the reasons that bring humans to anthropomorphise non-human agents. One of the main reasons is connected to science-fiction and media (Zlotowski et al. 2015). In TV shows, books and movies we are used to experience very intelligent, sensitive, power-driven or freedom-seeking robots. Consequently, people's expectations and most of their experience with robots is created by the media industry robots (Zlotowski et al. 2015). However, their awareness of the real capabilities of the robot, both potential and limitations, might affect differently people's perception of a robot, and consequently their trust in it. In the next Chapter 8 I investigate how different awareness levels of a robot's capabilities affect people's trust in the robot.

Chapter 8

Study 3: Effects of People's Awareness of a Robot's Capabilities on their Trust in the Robot

> Awareness is a key ingredient in success. If you have it, teach it, if you lack it, seek it.
>
> *Michael Kitson*

Human awareness of a social robot's capabilities are often mislead by the robot appearance and behaviours (Hegel et al. 2011). Especially, people who never had a real world encounters or interactions with a robot often have misconceptions and false expectations on robot's capabilities and functionalities influenced by science fiction movies and stories. In reality, there is a significant gap between the current state of the robotic platform and science fiction (Kriz et al. 2010), or even advertisements for real robots that make use of artificial intelligence[1]. As a consequence, this caused a negative effect on the interaction quality when the user's expectations of the robot is violated (Manja Lohse 2009).
In this Chapter, I investigated the relationship between users' expectations of the robot and the quality of their interaction experiences with the robot. Particularly, I analysed the users' ratings of trust in the robot over a series of interaction sessions where the robot slowly revealing a different aspect of its capabilites in each interaction session. The study was conducted with primary and secondary school

[1] https://www.youtube.com/watch?v=SSecbMFQK1I

pupils. Students were tested according their group-class with same procedure but a different robot. Pupils were asked to rate their trust in the robot after three types of exposures: a video demonstration, a live interaction, and a programming task. With this approach, I wanted to gain insights on how human awareness of the robot's capabilities affects their trust in it.

Research Questions

The purpose of this study was to investigate how people's awareness of a robot might have affected their trust towards robots. In particular, this research aimed to answer the question:

- **RQ-2.3** Can different levels of awareness of a robot's potential affect human perception of a robot?

In order to answer this question I planned a study as part of an event for the UK Robotics Week. Within this context, I decided to present a series of activities to introduce students from local primary and secondary schools to social robotics. The events have been structured in a very similar way. However, the pupils interacted with two different robots due the different complexity in programming of the children. In the primary school, students interacted with a robot called Kaspar (see Chapter 3 for more details). Pupils of the second school interacted with the robot called Pepper developed by Softbank Robotics (see Chapter 3 for more details).
All participants received a certificate of participation at the event.
The pupils were exposed to three different types of activities. The activities were: 1) watching a video, 2) interacting with a real live robot, and 3) programming of the respective robots. In particular, during the programming sessions, students were asked to create emotional behaviours for robot from a given predefined set of robot's behaviours. I collected pupils' perceptions of the robot through questionnaires.
The focus of the study was to collect students' perceptions of the robot after they participated in each of the three activities by means of questionnaires. Each activity provided different level of exposure to robots. Emotions were not the main focus of this study, but I used them to provide a pleasant and interesting experience in a robot storyteller scenario.
These two studies have been entirely designed by me. However they are collaborative works. In particular, I had the opportunity of collaborating with Dr. Patrick Holthaus for the event in the secondary school and with Silvia Moros for the event

in the primary school. Both colleagues accompanied me during the events in the school and they also collaborated in the studies by programming the pre-defined behaviours respectively for the robots Pepper and Kaspar.

Hypothesis

I hypothesise that human attitudes towards robots will change when they become aware of the actual capabilities (potential and limitations) of the robot. With my findings, I improve the way an interactive relationship between human users and their robotic companion can be established on different knowledge levels about the robot.

8.1 Experimental Design of the two Studies

I observed and analysed participants' behaviours during three different levels of interactions. The event consisted of three activities: 1) watching a humanoid robot interacting with a person, 2) interacting with robot, and 3) programming robot.
In the first part of the event, students watched a brief introductory video in which an actor interacted with the «*robot*» (Pepper or Kaspar).
The second part of the event focused in a closer interaction with the robot where students can touch the robot and test its functionalities.
Finally, participants built a story combining a given narrative with self-defined behaviours, including gesture, voice or vocal cues, head and body movements.
In order to analyse the pupils' interactions and their experiences with the robot, the pupils were presented with a short questionnaire at the end of each activity.
I asked pupils different questions at the end of each activity[2]. For example, previous experiences with the robot, assessing participants' willingness to have the robot as companion, their perception of trust of the robot, and their experience in programming the robot.
The questions sets were:

Q1 Have you ever seen this robot before today?

Q2 If yes, where: ______________

[2] In all questions, the robot has been replaced with the name of the robot the pupils interacted with, i.e. Pepper or Kaspar.

Q3 Would you like to have the robot in your home?

Q4 Do you trust the robot to be able to help you with your homework?

Q5 Do you trust the robot to wake you up in time for going to school?

Q6 Do you trust the robot to be able to warn you of danger, e.g. when using the Internet?

Q7 Do you trust the robot to help you in case of danger?

Q8 Programming the robot was? [very boring - very fun]

Q9 Programming the robot was? [very hard - very easy]

Q10 Would you like to program the robot again?

Questions Q1 and Q2 were asked only after the video interaction, while pupils answered the questions Q8-Q10 only at the end of the programming phase. I repeated the sets of questions Q3-Q7 after all the interactions. I chose these simple scenarios the pupils are familiar with. Indeed, since primary school, British children are taught about the danger of internet, such as cyberbullying or fishing. They know that is important to wake up in time to go to school and doing properly their homework.

8.2 Study at the secondary school with Pepper

The secondary school's pupils worked with two Pepper robots for the activities that lasted approximately three hours including a short break.
During the video activity, participants watched a commercial video of Pepper interacting with people on a screen[3].
During the second part of the event the interaction activity with Pepper allowed the pupils to touch the robot and use Pepper's built-in awareness function[4] and the "tickle me" scenario[5] (see Figure 8.1).

[3]The link to the commercial video is available in Appendix I

[4]http://doc.aldebaran.com/2-4/family/pepper_user_guide/life_pep.html#alife-pep

[5]http://doc.aldebaran.com/2-4/getting_started/samples/sample_interactive.html

Figure 8.1: Pupils are interacting with the robot during the second stage of the study.

During the programming activity, pupils created robot behaviours to match the emotions associated to a simplified story of the Hansel and Gretel fairy tale, composed of six different sentences conveying one of the following six basic emotions:

- **fear** Hensel and Gretel were rushing into the deep dark wood.
- **anger** Gretel clenched her teeth firmly because of the way their parents treated them.
- **sadness** But Hensel was feeling lonely and missed their family.
- **surprise** Suddenly, an old lady appeared out of nowhere!
- **disgust** She had an ugly green wart on her nose and smelled strangely.
- **happiness** Fortunately, they followed their intuition, decided to return home and reunited with their parents.

8.2.1 System Design and Materials

The robot used for this study is a Softbank Robotics Pepper (more details in Chapter 3).

The custom behaviours were to be designed and implemented in the robot's graphical programming suite Choregraphe[6]. To achieve this goal, the pupils could use predefined building blocks that allowed them to manipulate body movements, gesture, verbal and non-verbal cues individually per sentence and thus emotion.

8.2.2 Participants

I conducted the event in a local secondary school. Participants were secondary school pupils in year groups 10 to 12 [min age 14, max 15, avg. age 14.76, std. dev. 0.42]. The event was conducted over two consecutive days in the school and participants were tested in their age year groups, making a grand total of 43 pupils [6 girls, 37 boys].

8.2.3 Results

As part of the questionnaire, I was interested in participants' previous experiences, their perceptions and expectations of the Pepper robot. The majority of participants (86%) stated that they do not have any previous experience with this robot, while one participant was not sure if she saw it in a program TV. Participants whom have previous experiences with Pepper learned about this robot from either: 1) watching a TV program (e.g a documentary at BBC); 2) live show about humanoid robotics; advertisement 3) on a poster, 4) on Snapchat, 5) on the web.

Questions Q3: Companionship

Participants expressed their willingness of having Pepper in their home through a Yes/No/Maybe measure. The results of my study show that participants were uncertain about having Pepper as home companion after the video only. The majority of participants (58.13%) may like to have Pepper in their home. 27.9% would definitely like to have Pepper in their home, and 13.95% preferred not to have Pepper in their home. In contrast, the majority of participants (76.75%) expressed a strong willingness of having Pepper in their home after they have interacted with Pepper in the second activity, a preference that decreased (60.46%) after the programming activity. After the live activity, 18.6% of the participants indicated they may be willing to have Pepper in their home, and two participants did not

[6] https://community.ald.softbankrobotics.com/en/resources/faq/developer/what-choregraphe

Table 8.1: Paired samples t-test analysis comparing the means of participants' responses at Question Q3 according the three different HRIs. For each pair of interactions it shows the t values, p-value corresponding to the given test statistic t, the upper and lower bounds of the confidence interval.

Q3 (pepper in home)	t	p	95% CID
video - live	$t\_42 = -6.496$	$p < 0.0001$	$-0.762 - -0.401$
video - programming	$t\_42 = -8.827$	$p = 0.002$	$-1.057 - -0.664$
live - programming	$t\_42 = -3.334$	$p = 0.001$	$-0.448 - -0.110$

want Pepper in their home. After the programming activity, 25.58% participants indicated they may be willing to have Pepper in their home, and the remaining expressed negative consent to willingess of having Pepper in their home.
I also found a statistically significant correlation between the willingness of having the Pepper in their home and the effects of the activities ($p(2,84) < 0.0001, F = 47.162$). A t-test analysis on the interactions' paired samples showed that there are significant average differences between the participants' perceptions of the robot and the type of interaction. On average, the ratings were higher after the live activity than the video activity, and they were higher after the programming activity than both video and live activities (see Table 8.1).

Discussion

I hypothesise that the participants' perceptions of the robot changed so drastically from the first interaction due for two reasons. I showed participants a commercial video of Pepper and the majority of participants did not have any previous experience with the robot. People are acquainted to advertisements, and when they do not know the brand, they can have a higher critical judgement on the advertising value (Pintado et al. 2017). On the other hand, the awareness of real capabilities of the robot, acquired during the programming activity, mitigated the negative participants' perceptions of the interactive video interaction with the real robot.

Questions Q4-Q7: Trust in the robot

Participants answered questions Q4-Q7 using a 5-point Semantic Differential Scales where 1 corresponds to "definitely no" and 5 corresponds to "definitely yes".

Table 8.2: Paired samples t-test analysis comparing the means of participants' responses for Question Q4 according to the three different activities. For each pair of activities it shows the t values, p-value corresponding to the given test statistic t, the upper and lower bounds of the confidence interval.

Q4 (homework)	**t**	**p**	**95% CID**
video - live	$t\_42 = -3.334$	$p = 0.002$	$-0.762 - -0.401$
video - programming	$t\_42 = 8.323$	$p < 0.0001$	$0.969 - 1.589$
live - programming	$t\_42 = 12.943$	$p < 0.0001$	$1.4333 - 1.962$

All the ratings that is equal or less then 2 were categorised as negative response, with values equal to 3 were considered uncertainty and with values equal or larger than 4 were categorised as positive responses.

Question Q4 (homework)

Similarly to question Q3, I observed an increase of participants' trust in the robot after the live activity and a decrease of their trust after the programming activity. After the video activity, the majority of participants (74.42%) would trust Pepper to be able to help with their homework, and the 18.6% were uncertain in the robot's ability of perform such task. After the live activity, the majority of participants that trusted Pepper increased at 88.37% while uncertain participants decreased at 9.3%. After the programming activity, I observed a decrease of participants that trusted the robot's capability of performing the task (58.14%) and an increase of participants who were not completely confident in the robot (32.56%).

I observed a statistically significant correlation between participants' trust in the robot to be able to do participants' homework and the different types of interaction ($p(2,84) < 0.0001, F = 82.949$). On average, the significant differences can be observed between the activities (see Table 8.2). In particular, participants trusted the robot more after the live interaction comparing video and live activity, and live and programming activity. However, the participants' trust in the robot is higher after the programming activity, comparing video and programming activities.

Table 8.3: Paired samples t-test analysis comparing the means of participants' responses at Question Q5 according the three different activities. For each pair of interactions it shows the t values, p-value corresponding to the given test statistic t, the upper and lower bounds of the confidence interval.

Q5 (alarm clock)	**t**	**p**	**95% CID**
video - live	$t\_42 = -3.223$	$p = 0.002$	$-0.681 - -0.156$

Question Q5 (alarm clock)

I asked participants if they trusted the robot to be able to wake them up for going to school. I observed that participants were divided between trusting the robot (55.81%) and not trusting in the robot (34.88%) after the video activity. Again, the live activity with a real robot increased their trust in the robot (79.07%) and only 16.27% were unsure if the robot was able to complete the task. The participants' perceptions of the robot's capabilities changed after the programming activity, 67.44% of participants declared they trusted the robot with the task, while 25.58% remained uncertain.

I found a statistically significant correlation between participants' trust in the robot to be able to wake them up for going to school and the different activities they were tested with. ($p(2,84) = 0.02, F = 4.097$). In particular, participants rated higher their trust in the robot after the live activity than the video activity (see Table 8.3).

Question Q6 (danger warning)

I observed that participants were sceptical about the robot's ability of warning them of a danger after the activities. Indeed, the majority of participants (46.51%, 44.18% and 51.16% respectively after the video activity, live activity and programming activity) were not sure to trust the robot. Their trust in the robot (37.21%) increased slightly only after the live activity (41.86%) and decreased again after the programming activity (27.90%).

Participants' trust in the robot to be able to warn them of a danger was found positively correlated with the three different activities ($p(2,84) = 0.009, F = 5.014$). I also found that participants trusted the robot more after the programming activity (see Table 8.4).

Table 8.4: Paired samples t-test analysis comparing the means of participants' responses at Question Q6 for the three different activities. For each pair of activities it shows the t values, p-value corresponding to the given test statistic t, the upper and lower bounds of the confidence interval.

Q6 (danger warning)	**t**	**p**	**95% CID**
video - programming	$t\_42 = -2.622$	$p = 0.01$	$-0.905 - -0.118$

Table 8.5: Paired samples t-test analysis comparing the means of participants' responses at Question Q7 according the three different activities. For each pair of activities it shows the t values, p-value corresponding to the given test statistic t, the upper and lower bounds of the confidence interval.

Q7 (danger help)	**t**	**p**	**95% CID**
video - programming	$t\_42 = -2.973$	$p = 0.005$	$-0.820 - -0.157$
live - programming	$t\_42 = -2.673$	$p = 0.01$	$-0.816 - -0.114$

Question Q7 (danger help)

Similarly to question Q6, participants' responses after the activities indicated that they did not believe in the robot's capability of helping them in case of danger. After the video activity, participants' were divided between not trusting the robot (41.86%) and not being confident in trusting the robot (44.19%). Interestingly, while after the live activity this division was equal (41.9%), after the programming activity, 11.67% participants declared to trust the robot to be able to help them in case of danger and 53.5% participants remained uncertain.
I also observed a statistically significant correlation between the pupils' trust in the robot's capability of helping them in case of danger and the activities they were tested with ($p(2,84) = 0.003, F = 6.211$). On average, participants trusted the robot with the task more after the programming activity comparing both the video and live activities (see Table 8.5).

Discussion

The results of this study show that participants' perceptions of trust in the robot were partially affected by the awareness of the robots' capabilities. I hypothesise

that their trust in the robot might have been influenced by the embodiment of the robot. Indeed, when participants were asked to rate their trust in the robot in waking them to go to school, the live activity had a greater effect on their perception of trust than the video activity.
Moreover, I observed that participants' ability to trust the robot to be able to detect and handle the dangerous situation in the scenario were affected by their awareness of the robot capabilities. According to Deutsch (1958), risk-taking and trusting behaviour are different sides of the same coin, and a person is willing to take a risk only if the odds of a possible positive outcome are greater than those for a potential loss. Indeed, participants trusted the robot more if they know the real potential and limitations of the robot. They believe to be able to program the robot to handle the dangerous situations themselves, or have someone else available to program it for them.

Questions Q8-Q10: Programming Pepper

At the end of the programming activity, participants were asked to rate their experience of programming the Pepper robot.
Participants answered questions Q8-Q10 using 5-point Semantic Differential Scales (see Figure 8.2). All the ratings with values less then 3 were categorised as negative response, with values equal to 3 were considered moderate and with values more than 3 were categorised as positive responses.
The majority of pupils (74.42%) thought that programming the robot was fun, while only two participants did not enjoy the experience. In the comment section they declared that they did not understand how to program the robot.
At the question how easy it was for them to program the robot, 62.8% of participants expressed a moderate response and 20.93% declared it was easy to program the robot. The remaining participants indicated it was very hard to program the robot. Pupils gave a moderate or positive feedback. Indeed 23% of participants would like to program the robot again, while only two percent prefer to not repeat the experience. I observed from the open-ended questions that one of the two participants considered the robot terrifying, the other did not understand the programming explanations.

Discussion

As a public engagement activity, the event was a success for the pupils and they enjoyed programming the robot. However, partially disguised by the teachers,

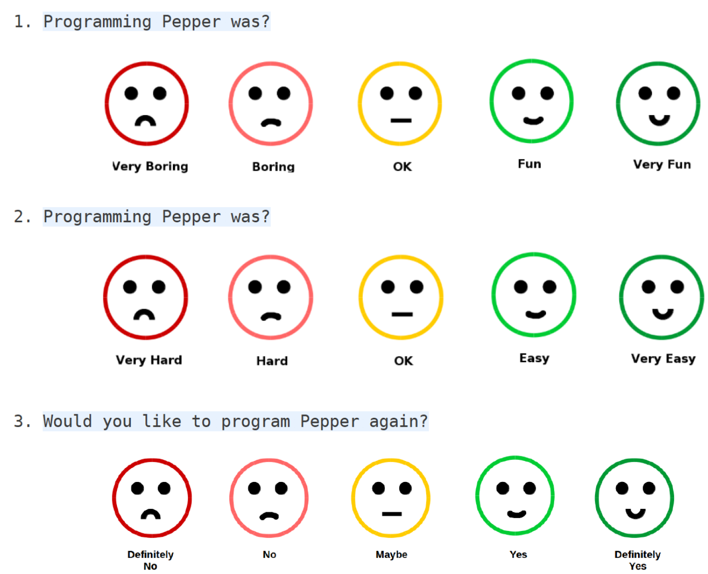


Figure 8.2: Questions Q8-Q10: Programming Pepper. Participants used this scale for rating their experience with Pepper during the Robotics Week.

I overestimated participants' programming skills and underestimated the time required for them to program emotional behaviours for the robot. For future investigations, I will consider a more exhaustive pre-session to further explain to participants how to use the program Choreographe and to program different behaviours for the robot.

8.2.4 Conclusion

With the event, I successfully familiarized pupils with the robot and the scientific field of social robotics. The majority of participants stated they were happy about their interactions with the robots and the programming activities. However, some pupils stated that programming the robot was not an easy task.
The study supports my hypothesis and shows that participants' awareness of the robots' real potential and limitations affected their perceptions of trust in the robot. Participants' awareness of being able to program different robots' behaviours led them to believe that the robot is able to handle critical situations and cognitive tasks, such as helping them with their homework.
Furthermore I conclude that their trust in the robot might have been affected also

by the embodiment of the robot. Indeed, I observed that the live interaction with the robot affected the participants' willingness of having Pepper in their homes increasing the effect.

8.3 Study at the primary school with Kaspar

Similar to previous study, the event consists of the activities. The first activity was the video activity which involved that the participants watched a brief introductory video in which an actor interacted with the robot.

The second activity was the interaction activity which involved participants spending 10 minutes interacting with the robot.

Finally, the last activity was the programming activity where the participants programmed 4 different emotions and a "silly" expression of the robot. I used the theme of emotions and a selected behaviour to provide a pleasant and interesting experience in a robot storyteller scenario. The four emotions used were based on Ekman's (1999) set of emotions (happy, sad, surprised and angry). The "silly" expression was included to make the event a fun experience for the primary school children.

The participants took part in the first two activities together as a group, while they were tested divided by their class groups for the programming activity. The activities were designed and organised to help the students to be familiarised with the robot and its technology. This aims is to provide the data necessary for identifying if there was a relationship between participants' perceptions of trust on increasing levels of awareness of the robot's capabilities.

8.3.1 System Design and Materials

Kaspar (see Figure 8.3) is a fully programmable humanoid robot with the size of a small child. The characteristic of this robot are described with more details in Chapter 3.

Similarly to the event at the primary school, pupils could use predefined building blocks that allowed them to manipulate body movements, gesture, and verbal cues for each emotions.

Figure 8.3: A child interacting with the humanoid robot Kaspar.

8.3.2 Participant

I conducted the event in a local primary school in Hatfield (Hertfordshire, United Kingdom). The event was conducted over a week in the school and participants were tested in their age year and class groups. I recruited 172 children, aged 7 to 10 [mean 8.87, std. dev. 0.85], 53% boys and 47% girls, and all had previous experience using the Scratch programming language.

8.3.3 Results

As part of the pre-questionnaire, I asked about participants' previous experiences, their perceptions and expectations towards the Kaspar robot. All participants, but one, declared to not have any previous experience with this robot. The only participant with previous experiences with Kaspar took part in a previous study of my research group.

Questions Q3: Companionship

Participants expressed their willingness of having Kaspar in their home through a Yes/No/Maybe measure. The results of my study show that participants' likeability of Kaspar decreased after each activity. After the video activity, the majority of participants (60.67%) would like to have Kaspar in their home, 27.18% was uncertain if they would have liked to have Kaspar in their home and only 12.15% would have not preferred to have the robot. After the live activity, the majority of participants (45.1%) expressed the preference of having Kaspar in their home, a preference that decreased to 41.62% after the programming activity. After the live and programming activities, 31.2% and 34.7% pupils respectively did not want Kaspar in their home. While 23.12% participants were uncertain in their willingness of having Kaspar in their home after live activity and programming activity.

I also found a statistically significant correlation between the willingness of having the Kaspar robot in their home and the effects of the interaction ($p(2,342) < 0.0001, F = 27.298$). I performed a t-test analysis on the activities' paired samples and I found that there are significant average differences between the participants' perceptions of the robot and the type of activity. On average, the ratings were lower after the programming activity than both video and live activities (see Table 8.6). I did not find any significant differences while comparing the video and live activities' results ($p > 0.05$).

Table 8.6: Paired samples t-test analysis comparing the means of participants' responses to Question Q3 according to the three different activity experiences. For each pair of activities it shows the t values, p-value corresponding to the given test statistic t, the upper and lower bounds of the confidence interval.

Q3 (kaspar in home)	**t**	**p**	**95% CID**
video - programming	$t\_171 = -6.471$	$p < 0.0001$	$-0.539 - -0.287$
live - programming	$t\_171 = -5.225$	$p < 0.0001$	$-0.473 - -0.213$

Questions Q4-Q7: Trust in the robot

Participants answered questions Q4-Q7 using a 5-point Semantic Differential Scales where 1 corresponds to "definitely no" and 5 corresponds to "definitely yes". All the ratings with values less then 3 were categorised as a negative response, with values equal to 3 were considered as uncertainty and values more than 3 were categorised as positive responses.

Question Q4 (helping with homework)

I observed an increase of participants' trust in the robot after the live activity and a decrease of their trust after the programming activity. After the video HRI, the majority of participants (53.8%) would trust Kaspar to be able to help with their homework, and 23.12% were uncertain in the robot's ability of perform such task. After the live and programming activities, the majority of participants that trusted Kaspar did not change (53.8%) while the percentage of uncertain participants increased respectively at 26% and 25.4%. The remaining participants were not completely confident in the robot (respectively at 19.7% and 20.2%).

I did not observe a statistically significant correlation between participants' trust in the robot to be able to help with participants' homework and the different types of activity ($p > 0.5$).

Question Q5 (alarm clock)

I asked participants if they trusted robot to be able to wake them up for going to school. I observed that participants were divided between trusting the robot (39.9%) and not confident in their trust in the robot (33.53%) after the video interaction. The live interaction with a real robot increased their trust in the robot (43.9%) and 27.18% were unsure if the robot was able to complete the task. The pupils' perceptions of the robot's capabilities changed slightly again after the programming session, 42.8% of participants declared to trust the robot with the task, while 31.21% remained uncertain.

I did not find a statistically significant correlation between participants' trust in the robot to be able to wake them up for going to school and the different activities they were tested with ($p > 0.5$).

Question Q6 (danger warning)

I observed that participants were skeptical about the robot's ability of warning them of a danger after the different activities. Indeed, the majority of participants 50.9%, 54.4% and 47.4% respectively after the video activity, live activity and programming activity trusted the robot. Their uncertainty of trusting the robot (21.4%) decreased slightly after the live activity (20.23%) and increased again after the programming activity (25.43%). Instead the participants who did not trust the robot remained stable (26.6%).
Participants' trust in the robot to be able to warn them of a danger was not found positively correlated with the three different activities ($p > 0.5$).

Question Q7 (danger help)

Similarly to question Q6, participants were not completely confident in the robot's capability of helping them in case of danger after the different activity interactions. I noticed that participants trusted more the robot after the programming activity (47.4%) compared to both the video (40.5%) and live activities (43%). Interestingly, while the participants who did not trust the robot varied slightly after each activity (33%, 29.5% and 31.2% respectively), I observed a bigger variations in those who were uncertain about the robot's ability of completing the task (26%, 27.2% and 20.8% respectively).
I did not observe any statistically significant correlation between the pupils' trust in the robot's capability of helping them in case of danger and the activities they were exposed to ($p > 0.5$).

Discussion

I did not find any statistically significant differences in the results of this study concerning the pupils' perception of trust in the robot's capabilities. I observed that pupils tendentially believed to be able to program the robot to handle the dangerous situations themselves, or have someone else available to program it for them. However, they trusted the robot to be able to do their homework, to wake them up to go to school, and to handle dangerous situations regardless their awareness of the real potential and limitations of the robot. Indeed, when I asked the participants to rate their trust in the robot in waking them to go to school, the live and programming activities had a greater effect on their perception of trust than the video activity. According to Deutsch (1958), risk-taking and trusting behaviour are different sides of the same coin, and a person is willing to take a risk only if the odds of a possible positive outcome are greater than those for a potential loss. However, the perception of the risk might have been mitigated by motivations such as curiosity and fun (Maner et al. 2007). Indeed, several pupils expressed their curiosity about Kaspar's characteristics and potentialities writing queries for the experimenter in the open comment sections. For example, participants asked about the material used for the robot's skin, the age of the robot and the possibility of having different hairstyles. Two pupils also asked if there was a female version of Kaspar. Pupils also enjoyed the interactions with Kaspar (see also Section 8.3.3.

Questions Q8-Q10: Programming Kaspar

At the end of the programming activity, I asked participants to rate their experience of programming the Kaspar robot.
Participants answered questions Q8-Q10 using a 5-point Semantic Differential Scales (see Figure 8.4). All the ratings with values less then 3 were categorised as negative response, with values equal to 3 were considered moderate and with values more than 3 were categorised as positive responses.

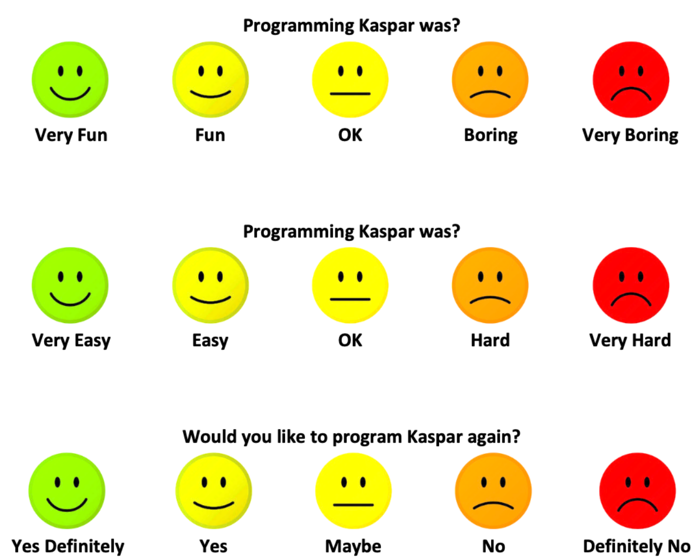


Figure 8.4: Questions Q8-Q10: Programming Kaspar. Participants used this scale for rating their experience with Kaspar during the Robotics Week.

The majority of the participants (95%) enjoyed programming Kaspar. Less than 2% thought it was Boring or Very Boring, found in two boys from the older grade and one girl from the younger grade, while 3.41% gave a moderate feedback.

At the question how easy it was for them to program the robot, 75% expressed a quite positive response, two children left the question blank, and 6% thought it was hard. The remaining participants gave a moderate response.

The majority of children gave a positive feedback (87%). Only two percent prefer to not repeat the experience, however they enjoyed programming the robot. The remaining participants (12%) expressed a more moderate response.

8.3.4 Conclusion

During the event, children successfully familiarised themselves with the robot and scientific field of social robotics. The majority of participants stated they were happy about their interactions with the robots and the programming activities.

The study shows that participants' awareness of the robots' real potential and limitations affected their perceptions of the robot, but not their trust in it. In particular, participants' awareness of being able to program different robots' emotions and behaviours decreased their likeability of the robot.

I believe that their perception of the robot might have been affected by the embodiment of the robot. While, as discussed in Chapter 2, trust is strongly associated with the perception of reliability and willingness to take a risk (Deutsch 1958; Ross 2008), other aspects such as the appearance, type, size and proximity of a robot might also affect the user's perceptions of the robot (Bainbridge et al. 2011; K. L. Koay, Dag Sverre Syrdal, et al. 2007). Indeed, Kaspar looks like a child, it has minimal facial features and he is not able to walk.

Contrary to my expectations, I did not find any significant differences in trust tendencies for the different HRI activities. However I observed that the awareness of the real capabilities of the robot, acquired during the programming activity, mitigated the participants' negative perceptions of the video activity with the real robot. The results suggest that the higher familiarity in programming the robot might lead them to believe that the robot is able to warn and help them in case of a danger, or wake them up for going to school.

8.4 Comparing the two studies

In the following Sections I am comparing the participants' responses for each conditions (video, live and programming) on the robots Pepper and Kaspar. Firstly, I ran two-way ANOVAs to underline any interactions between the data grouped for the different type of robot. Then, I used independent-samples t-tests to determine if there were differences in the means for participants' previous experiences, their perceptions and expectations towards the robots. The data used for the t-tests are mean ± standard deviation, unless otherwise stated. The assumption of homogeneity of variances was violated, as assessed by Levene's test for equality of variances ($p < 0.05$), unless otherwise stated. I also included the mean differences with 95% confidence intervals.

8.4.1 Questions Q3: Companionship

A two-way ANOVA showed that there is no interaction between the participants' responses for each conditions (video, live and programming) on the robots Pepper and Kaspar. Indeed, I did not find any statistically difference in the means ($p >$ 0.05). A general comparisons of participants' responses according the different interactions and robots is shows in Figure 8.5.

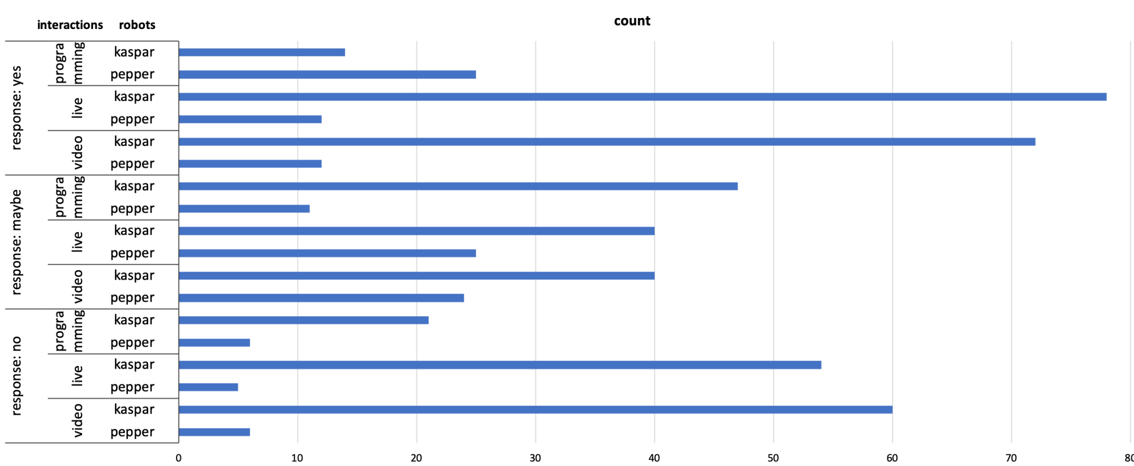


Figure 8.5: Comparison of participants' perception of the robots (Pepper and Kaspar) as home companion after each interaction condition (video, live, programming). Results are categorised by their responses (no, maybe and yes) to have the robots as home companion.

8.4.2 Questions Q4-Q7: Trust in the robot

I assessed pupils' perception of trust of the robots using questions Q4 to Q7. In particular, I collected their trust in a robot that is able of helping them with homework, waking them up for going to school, warning and helping them in case of a danger. I am now interested in understanding the effects of the different robot on their ratings of trust.

Q4 (helping with homework)

I did not find any difference in mean for the use of Pepper and Kaspar on participants' perception of trust that the robot they interacted with was able to help them with their homework ($p > 0.5$).

Question Q5 (alarm clock)

A two-way ANOVA test showed that there is an interaction between the participants' responses for video ($p(1) = 0.04, F = 4.281$) and programming ($p(1) = 0.02, F = 5.152$) conditions on the use of robots Pepper and Kaspar.
In particular, after the video interaction the statistically significant difference ($t(101) = 2.751, p = 0.007$) in mean trust that the robot is able to wake them in time for going to school score between pupils that interacted with Kaspar and Pepper, with the firsts (3.09 ± 1.45) scoring lower than the seconds (3.57 ± 0.9).
While after the programming interaction, the statistically significant difference ($t(104) = 3.06, p = 0.003$) in mean trust score between pupils that interacted with Kaspar and Pepper, with the firsts (3.20 ± 1.4) scoring lower than the seconds (3.71 ± 0.84).

Question Q6 (danger warning)

I did not find any statistically significant interaction between the responses of participants tested with Pepper or Kaspar ($p > 0.05$).

Question Q7 (danger help)

A two-way ANOVA test showed that there is an interaction between the participants' responses for video ($p(1) = 0.02, F = 5.46$), live ($p(1) = 0.01, F = 6.089$) and programming ($p(1) = 0.005, F = 8.189$) conditions on the use of robots Pepper and Kaspar.

In particular, after the video interaction the statistically significant difference ($p = 0.005, t(86) = -2.88$) there was a statistically significant difference in mean trust score between participants that used Kaspar and Pepper, with trust in Kaspar scores (3.13 ± 1.41) greater than the trust in Pepper scores (2.60 ± 1.0).
After the live interaction, I can observe that ($p = 0.003, t(85) = -3.02$) mean trust in Kaspar scores were greater (3.16 ± 1.39) than mean in trust in Pepper scores (2.60 ± 1.0).
Similarly, after the programming interaction, I found a statistically significance difference in mean ($t(104) = -3.85, p < 0.0001$) trust scores where interactions with Kaspar (3.32 ± 1.42) scored higher than the interactions with Pepper (2.67 ± 0.9).

8.4.3 Conclusion

This study shows that the robot used for the interactions was only partially a factor that affected participants' awareness of the robot's capabilities, and consequently their trust of the robot. In particular, while pupils' acceptance of the robot were not affected by a different robot, I found that they trusted Pepper more than Kaspar to be able to wake them up for going to school after the video and programming interactions. Indeed, the results reported in Section 8.2.3 and Section 8.3.3 highlighted that participants trusted more Pepper than Kaspar after both video (respectively 55.81% and 39.9%) and programming (respectively 67.44% and 42.8%) condition. However, participants' trusted Kaspar more than Pepper to be able to help them in case of danger comparing the mean scores after each experimental conditions. Indeed, results reported in Section 8.2.3 and Section 8.3.3 showed that participants trusted Pepper (video 13.95%, live 16.2%, programming 11.67%) less than Kaspar (video 40.5%, live 43%, programming 47.4%).
I do believe that while primary school pupils' perception of the risk might have been mitigated by their curiosity and having fun with robot (Maner et al. 2007), the question Q7 (help in case of danger) might have been perceived as too generic to have a stronger impact on the participants.

8.5 Discussion and Conclusion

During the events at the primary and secondary schools pupils familiarised respectively with Kaspar and Pepper robots. I observed their interactions with the robots in order to understand how a higher awareness of robots influences people's

perception and trust in them. Both events were a success, and the majority of participants were happy of programming the robots.
The findings show that participants' awareness of the robots' real potential and limitations affected their perceptions of the robot (research question **RQ-2.3**), but they did not affect their trust into robots in the same way in both studies. In particular, higher familiarity with the robots influenced positively the robot's likeability of those who interacted with Pepper, while it mitigated an overall participants' negative perception of Kaspar.
I also found that an higher awareness led the students to trust that Pepper is able to handle critical situations and cognitive tasks. Contrary to my expectations, there were no statistically significance evidence to corroborate the same hypothesis regarding those who interacted with Kaspar. However, the differences of the two studies, in terms of participants' age, sample size and exposure time, might be factors affecting the findings.
To further understand these results, I compared the data collected during the two studies. In particular, I wanted to comprehend if the different appearance of the robots was a relevant affecting factor in participants' trust in them . However, the pupils were only partially affected by the type robot they interacted with. This led us to believe that novelty effect, pupils' curiosity and having fun during the interaction diminished their perception of risk, and consequently their trust in the robot (Maner et al. 2007).

In several works (Bickmore et al. 2010; Law et al. 2017; Short et al. 2010) curiosity is elicited in people, with special focus on children, by novel and surprising factors However, while the interactions between children and the robots resulted significantly enthusiastic, but they did not have an equivalent positive result towards the completion of the task. For example, curiosity did not increase the learning curve of the children in (Ceha et al. 2019; Jirout et al. 2012) studies. Moreover, the novel effect has the tendency of wearing off (Fernaeus et al. 2010). Therefore, to avoid that robots will be abandoned in an corner of the house or tossed away, it is very important to provide robots with more behaviours that will enhance their interactions with people. In this direction, current literature (Bishop et al. 2019; Breazeal 2003; Lugrin et al. 2018) highlighted how humans have a greater acceptance of robots that show social behaviours. Social behaviours may include head and body movements, proximity, gaze movements, speech and vocal cues, emotional demonstrations using eyes colours, head and body pose, etc. However, none of these studies identified which social cues is clearly recognised by people, and consequently which one has higher effect on humans' acceptance of a robotic companion. The study presented in the next Chapter 9 aims to close this gap.

Chapter 9

Study 4: Human Perceptions of the Social Cues Expressed by a Domestic Robot

> The best and most beautiful things in the world cannot be seen or even touched. They must be felt with the heart.
>
> *Helen Keller*

People are social entities, and they are more comfortable while interacting with agents that can show social behaviours. People are able to communicate and interpret communication signals that go beyond natural language and may involve gesture, pose, and body language. In addition to those, they might engage other humans with a bidirectional and mutual understanding (Sciutti et al. 2018) that enables them to anticipate and read implicit intentions and behaviours. Moreover, such indirect information exchange can be influenced by human social conventions such as simple habits of social interaction, expectations, and perception of robots. In the context of robots, people are more inclined to interact with a robot that they can anthropomorphise (Fiore et al. 2013; Hegel et al. 2011; Warta et al. 2018). Moreover, individuals' social interactions are affected by the use of social cues, which can elicit social signals such as representing emotions, intentions or personalities.
Current literature has identified a number of social cues that could influence people's perception of a robot as social entity, and their behaviours during an inter-

action. For example, while autonomous cars and some proposed robot applications (Baraka et al. 2016; Kheng Lee Koay, Dag Sverre Syrdal, Oskoei, et al. 2014) use lights and LEDs to communicate intentions to people, the robot used in Mikawa et al. (2018) and Koay et al. (2013) in their studies used a more natural approach based on non-verbal visual communicative signals, i.e. head movements and body orientation.

Rossi et al. (2019) presented a robot that slows down and asks people politely to make room for allowing it to pass was perceived more positively by participants than a non-social robot.

Eyssel et al. (2010) showed that a robot that provided emotional feedback to the participants resulted in higher anthropomorphic responses than a robot with neutral expressions.

Emotions and intents of a robot can also be expressed successfully though the use of colours, sounds and the intensity of vibrations (Song et al. 2017).

Moreover, Cassell et al. (2001) believe that vocal or written communication plays an important aspect for conversational modes for embodied autonomous robots.

I can summarise that the social cues used to elicit social perception and acceptance are the following: 1) speech and vocal cues, 2) head and body movements, 3) emotional expressions, 4) colours, 5) written text and 6) approach distance and orientation. Moreover, a robot that can interact using multi-modal cues is perceived by people as better for communication purposes (Holler et al. 2019) and it improves the effectiveness of the information exchange (Vigliocco et al. 2014). The limitation of currently developed robots is that it is not clear if the cues composing the multi-modal interaction affect equally the quality of the interaction, or if one or some of them are shadowed by the others. During a human-robot interaction, I hypothesise that it might be possible that people tend to be affected by the overall interaction experience more than than a simple sum of the effects of single social cues used by the robot might imply.

The current study investigated the impact of previously highlighted social cues on people's perception of the robot as social identity during an interaction. I conducted a repeated video-based study, where participants rated the sociability of a domestic robot in eight videos that differ from each other by the omission of a single social cues. In each video, the robot was engaged with a person in a simple fetch-and-carry scenario in which the robot made an error before completing its task. The cues were rated by participants in terms of sociability.

Research Questions

The purpose of this study was to investigate humans' perception of some social cues adopted by a robot during an interaction. In this context, my investigation was directed towards the classification of social cues according their impact on people's perception. In particular, as first step, I needed to identify which are the social cues that people notice:

- **RQ-3.1** Which are social cues people recognise? Ascertaining if people are conditioned more from some social cues than others.

In order to provide an answer to this research question **RQ-3.1**, I examined how individuals' perceptions would differ by collecting their ratings of several videos where a robot interacted using all, some or no social cues.

Hypotheses

I expected to see differences in how participants would be affected by different social cues expressed by a robot. In particular, considering that speech and gestures are the most common communication modes between humans, I anticipated that individuals would be affected more by the use of the speech, compared to the robot's gestures and movements.

9.1 Experimental Design

I used a repeated-measured experimental design. I collected participants' perceptions of the robot using a video-based study in which an actor interacted with the robot. I decided to use a video-interaction study because they have been shown to be appropriate for the type of interactions I were interested to study (M. Lohse et al. 2008; S. Woods et al. 2006). In particular, video based studies can reach a larger number of participants, and also to repeat exactly the same robot behaviours and interaction's dynamics for each participant.

9.1.1 Experimental Procedure

Participants were recruited from Amazon Mechanical Turk [1]. The study consisted of three stages. Participants were asked to watch videos of the interaction between a physical robot and a user. The user/actor was played by a researcher from the Adaptive System Research Group who was very familiar with robots. Participants were told through a small script at the beginning of the study that they should imagine living in their own house with a robot as a companion, and that it helps them in their everyday activities. The duration of the study for each participant was about 20-25 minutes.

During the first stage, participants were asked to complete a short questionnaire where they provided: 1) some personal and demographic data (e.g. age, gender), 2) TIPI – Ten Item Personality Inventory questionnaire about themselves (S. D. Gosling et al. 2003), and 3) their previous experiences with robots.

In the second stage, participants watched eight videos through a web application from their own computer. In each video, a robot interacted with a human-actor, and performed a simple task upon request. The actor asked the robot to go in the kitchen and bring a bag of crisps. The robot executed the task twice, because it failed to fetch the object requested when it was asked the first time. The erroneous behaviour was added for the robot to show different emotional expressions. Each video differed from the other for the social behaviours used by the robot. I chose six social cues among the ones used in previous research studies involving a social robot (Castro-González et al. 2016; Kheng Lee Koay, D. Syrdal, et al. 2017; Kurt et al. 2014; Alessandra Rossi, Garcia, Maya, et al. 2019). The experiment was conducted using 3 main different levels of social behaviour settings: 1) the robot acted with no social cues, 2) the robot acted using all social cues, 3) the robot acted using all social cues but one. The third condition was repeated for all the social cues adopted (i.e. six times). In total, participants watched eight videos. After each condition (i.e. video), the participant was asked to rate how socially they perceived the task, and an optional open-ended question to evaluate their preferences.

In the last stage, the participants was asked to list which social cue/cues their home robotic companion should use to alert them in 10 different scenarios. I designed these scenarios to cover a wide range of generic types of errors based on previous HRI research with home companion robots (K. L. Koay, Dag Sverre Syrdal, et al. 2007; Reiser et al. 2013; Alessandra Rossi, Kerstin Dautenhahn, Kheng Lee Koay, and Michael L. Walters 2017b).

[1] Amazon Mechanical Turk https://www.mturk.com

9.1.2 Measures

In order to collect participants' changes in perceptions of robot's social cues, I asked participants several questionnaires:

Questionnaire 1 A pre-experimental questionnaire for

- Collecting demographic data: age, gender, nationality and country of residence
- The TIPI questionnaire to classify participants for their level of extroversion, agreeableness, conscientiousness, emotional stability, and openness to experiences
- Participants' experience with regard to robots

Questionnaire 2 A small questionnaire after each condition for

- Collecting participants' perception of social behaviours of the robot in the video using a 5-points scale from 1 (not social at all) to 5 (very social)
- Optional open comment to justify their choice

Questionnaire 3 A post-experimental questionnaire to collect the participants' preferred social behaviours in several scenarios. The participants could choose one or multiple cues including body movement, head movement, gaze movement, approach orientation, approach distance, vocal cues, speech, eye colours, written text and other:

- You forgot to turn your oven off.
- You have an appointment with your friend for a drink.
- There is a broken glass in the kitchen.
- You forgot to close the water tap in your bathtub.
- You left your front door open.
- Your robot's camera is broken. It cannot see properly what there is around it.
- It is time to take your medicine.
- You have an important interview meeting.

- Your favourite TV show is going to start.
- Your robot's battery is too low. The robot will shut down while it is bringing your dinner.

9.2 System Design and Materials

In order to conduct my study, I developed a sequence of video interactions between a robot and an actor. Participants watched the videos and answers the questionnaire. For my purposes I developed the following components:

- Video interactions between an actor and a robot
- Website on Amazon Mechanical Turk to present the videos and questionnaires to the participants

9.2.1 Operating System

For this study I used a Dell XPS 15, but all the software is portable. The operating system was an Ubuntu 16.04 LTS:

- Product Name: Dell XPS 15 inch
- Kernel Release: 16.04.03
- Kernel Version: 4.10
- Processor: 6th Generation Intel Core i7-6700HQ Quad Core (6M Cache, up to 3.5 GHz)
- Graphics Card: NVIDIA® GeForce® GTX 960M with 2GB GDDR5
- RAM: 16GB Dual Channel DDR4 2133Mhz (8GBx2)

9.2.2 The robot

The robot used for this study was a Pepper version 1.7. The robot used has the same characteristics as described in Chapter 3. The custom behaviours were designed and implemented in the robot's graphical programming suite Choregraphe. To achieve this goal, I used scripts in Python language to program the robot's body movements, gestures, verbal and non-verbal cues individually and thus emotions.

9.2.3 The Social Cues

Participants were presented with eight different conditions. This included conditions where the robot interacted using the whole spectrum of cues available, minus one single social cue, in addition to the two baseline conditions in which the robot either did not show any social cues or displayed all of them. The robot's behaviours were implemented using the following social cues: 1) speech and vocal cues, 2) head and body movements, 3) emotions, 4) colours, 5) written text, and 6) approach distance and orientation. The implementation of each social cue was inspired by current literature.

In particular, the robot turned the LEDs of its eyes from the neutral colour (white) to the colour green to display a positive and excited feeling, while its eyes became red to display a more negative state (Elliot et al. 2007; Kaya et al. 2004; Kurt et al. 2014). In the condition in which this social cue was not present, the robot's eyes were constantly white.

Pepper raised its head and arms to express a content (happy) emotion, while it lowered its head to its chest and its arms to express a sad emotion (Häring et al. 2011). In the condition in which the robot could use speech and vocal cue, the gestures were accompanied by vocal expression, respectively such as "ooh" or "great" to emphasise the sad and happy emotion.

While interacting with the user, the robot faced the user with its body and head directed towards the person. The robot also moved slightly its arms to elicit a more natural human-like behaviour avoiding the stillness that is associated more with a machine or inanimate object (Castro-González et al. 2016). In the condition without the head and body movements, the robot remained completely stilled.

The robot spoke to the user by using a neutral pre-defined voice. In the conditions where there was no speech or sound, the robot did not provide any vocal feedback to the user.

Pepper communicated with the user using text displayed on its tablet. For the control condition with not this social cue, the screen remained blank.

I followed modern socially-aware techniques to encode human social conventions in distance and proxemics between a robot that approach and interacts with people (Kheng Lee Koay, D. Syrdal, et al. 2017; Mumm et al. 2011; Alessandra Rossi, Garcia, Maya, et al. 2019). Specifically, the robot approached the human with a constant velocity of 0.5 m/s, from a front-side direction, maintaining a close but safe distance of 0.5 m that allows the robot to hand-over an object to the user (Kheng Lee Koay, Dag Sverre Syrdal, Oskoei, et al. 2014). In the condition without socially-aware approach distance and proximity, the robot approached the user with a slightly faster velocity of 1m/s, from a back-side direction, moving as close as possible without touching the user.

9.2.4 The video interaction

The videos for the study were filmed in Robot House.
A colleague from the Adaptive Systems Research group, Dr. Dag S. Syrdal played the part of an occupant of the UH Robot House and interacted with Pepper. The videos were composed of two parts. In the first part, the actor is sitting on a couch reading a magazine while the robot is stationary not too far away from him. The actor asks Pepper to fetch some crisps from the kitchen. The robot, then goes to the kitchen and comes back to the couch without bringing anything. In the second part of the video, the user points out the mistake and asks again for the robot to comply with the request. This time the robot completes the task correctly. The set up of the environment is represented in Figure 9.1.

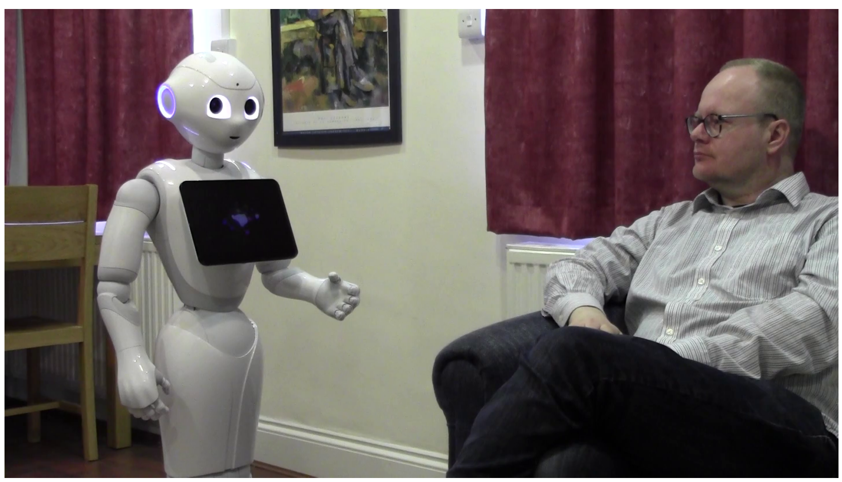


Figure 9.1: A shot of the interaction between the actor and Pepper. The actor is asking the robot to fetch him a pack of crisps.

All videos were recorded on the same day, and each interactions were recreated in the same way as much as possible. I shot with an initial wider angle of the HRI area. Most of the recordings were mainly focused on the actor, while shifting the focus as appropriate out to include the robot and the interaction with the person. No cuts were made to the videos giving a natural continuity to the interaction.

I stored the videos on Amazon S3 storage service[2].

9.2.5 The website

The website is composed by a simple interface created using HTML, CSS and Javascript. The website was integrated in Amazon Mechanical Turk to recruit participants for this study. The website included the videos and the questionnaires as described in Section 9.1.
Participants' data and responses were stored in a database managed by Amazon Mechanical Turk but only accessible by the experimenter.

9.3 Participants

Participants were recruited on Amazon Mechanical Turk. I collected responses from 100 participants, 71 male and 29 female (none identifying as non-binary). Participants were aged 22 to 70 (average 31.31, st.dev. 8.15).
Participants' nationalities included Indian (37%), American (57%), Spanish (1%), Italian (2%), Canadian (1%), Pakistani (1%), and American-Indian (1%). Participants lived in India (38%, including Tamilnadu and Madurai area), USA (58% including Texas, California, New York, Pennsylvania, Georgia and Florida), Italy (2%), one participant declared to be nomadic. One preferred not to answer the question.

9.4 Results

In this Section, I will analyse and discuss the participants' responses to the questionnaires introduced in Section 9.1.2, including their previous experience with robots, perceptions of robot's social behaviours, effects of personality traits on their choices and preferences for a robot that interact using social cues.

9.4.1 Previous experience with robots

As part of the questionnaire pre-experiment, I were also interested in participants' previous experiences with robots.

[2] Amazon S3 storage service https://docs.aws.amazon.com/s3/index.html

I used 5-point Likert Scales where 1 corresponds to "not at all", and 5 corresponds to "very much". In particular, I asked participants about previous experience:

- Do you have any experience interacting with robots?
- Please, specify what kind of experience you have with robots (if any)
- Which robots? (if any)

I observed that participants declared a quite distributed level of experience (min=1, max=5, mean 2.91, std. dev. 1.17). The majority of participants (39%) declared to have high experience with robots, 29% declared to have a moderate experience with robots, and 31% of participants had a low experience with robots. In particular, 23% of participants had experience as software developer for robots, while 27% of participants had experience respectively as participant in another experiment, as a researcher or as an observer. However, their comments also revealed that their knowledge of robots included Artificial Intelligence in general, Softbank robot Pepper, industrial robots, a waiter robot and service robots in hotels.

9.4.2 Ratings the relevance of robot's social cues

A one-way repeated measures ANOVA with single contrasts test was run to determine if there were differences in the perception of social cues expressed by the robot between the following conditions: all social cues, no use of colours, no use of emotions, no use of movements, no use of speech, no use of text on a tablet, no use of approach distance and orientation, and no social cues. Data are mean $\pm$ standard deviation, unless otherwise stated. The ratings of the social level of the robot's behaviours decreased from the condition "all social cues" (3.57 ± 0.956) to "no social cues" (3.16 ± 1.170). Figure 9.2 summarises the mean ratings for the other conditions organised in ascending order by the perception of sociability of the robot. Observing the mean scores, the lower ratings of videos actually reflected the higher relevance of the social cue missing in that specific video. For example, the video in which the robot interacts without speech has the lowest rating, and this leads us to believe that speech is the social cue perceived to be most important among all the social cues. Similarly, I believe that emotions and colours might be perceived as less important than speech, but more important than body and head movements, written text on the tablet, and approach orientation and distance.
The assumption of sphericity to verify that the variances of the differences between all possible pairs of within-subject conditions was violated, as assessed by

Descriptive Statistics

	Mean	Std. Deviation	N
no social cues	3.16	1.170	100
no speech	3.30	1.150	100
no colours	3.44	1.067	100
no emotions	3.44	1.095	100
no movements	3.46	0.937	100
all social cues	3.57	0.956	100
no text/tablet	3.60	1.005	100
no approach	3.65	0.968	100

Figure 9.2: Participants' responses to the 8 different sets of social cues were ranked according to their mean perceptions of the robot's sociability. In particular, a lower mean corresponds to a less sociable rating.

Mauchly's test of sphericity ($\chi^2(27) =, p < 0.0001$). Therefore, a Greenhouse-Geisser correction was applied ($\varepsilon = 0.728$) to the one-way repeated measures ANOVA, which showed a statistically significant change in the ratings of social perceptions of the robot's behaviours ($F(5.312, 525.90) = 4.751, p < 0.0001$).

Planned single contrasts to further investigate the possible pairs of within-subject conditions showed that that participants' ratings for a robot using all social cues were statistically higher than for a robot that does not use any social cues, with a mean difference of 0.41 (95% CI, 0.17 to 0.65), $p = 0.001$. I also found that the ratings of the robot that does not use speech were statistically lower than those of a robot that uses all social cues, with a mean difference of 0.27 (95% CI, 0.02 to 0.52), $p = 0.032$. In addition, there was a statistically significant increase of the social perception of the robot from the ratings of a robot that does not use any social cues to the ratings of other conditions in which the robot interacted with all social cues but one (95% CI). Table 9.1 shows the statistically significant *p-values*, the mean difference between the two ratings you compared (contrast estimate), and upper and lower bounds of the 95% confidence interval of the mean.

The ANOVA single contrasts also showed that the video in which the robot did not exhibit a proper approach distance and orientation to the user was statistically rated higher for sociability than videos in which the robot used colours ($p = 0.04$), did not have body and head movements ($p = 0.03$), and it did not use speech ($p = 0.003$), with a mean difference respectively of -0.21 (95% CI, -0.41 to -0.01), -0.19 (95% CI, -0.37 to -0.01), and -0.35 (95% CI, -0.6 to -0.12). Finally, I observed that the video in which the robot did not use speech was rated statistically

Table 9.1: Results of ANOVA with single contrasts test with a pairwise comparison of the participants' ratings to the robot's social level in the different videos.

Compared Conditions	**ANOVA with Single contrasts Test**		
	Sig.	**Contrast Estimate**	**Lower Bound - Upper Bound**
No colours - No social cues	$p = 0.027$	0.28	0.32 to 0.53
No emotions - No social cues	$p = 0.013$	0.28	0.06 to 0.5
No movement - No social cues	$p = 0.011$	0.3	0.07 to 0.53
No text - No social cues	$p < 0.0001$	0.44	0.23 to 0.65
No proximity - No social cues	$p < 0.0001$	0.49	0.25 to 0.73

significantly lower than a robot that did not communicate with written text on its tablet ($p = 0.004$), with a mean difference of -0.3 (95% CI, -0.5 to -0.1).
Then, I investigated any differences between participants' ratings by gender and age. I did not find any statistically significant difference between the choices made by men and women ($p > 0.05$). Similarly, I did not find a statistically significant difference between participants of different ages ($p > 0.05$).
I also tested the association between participants' ratings and their country of residence, and their ratings and nationality. Since the majority of the countries of residence and nationality only had one individual in each category, I applied the test only to India, USA and Italy. I did not find any statistically significant difference between participants' ratings and either the nationalities or countries of residence (both $p > 0.05$).

9.4.3 Effects of Personality traits on participants' perception of social cues

I asked the participants to rate how they consider themselves on a 5-point Likert Scale [1= disagree strongly and 5= agree strongly] according the TIPI questionnaire: extroversion, agreeableness, conscientiousness, emotional stability, openness to experiences.
I used a Spearman's rank-order correlation to determine whether there is an association between participants' personal traits and their perception of the social cues. In particular, I found that extroversion and the video with no speech was

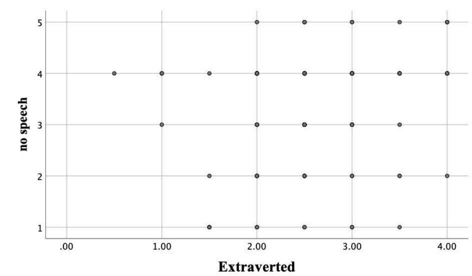
positively correlated ($p = 0.042, r_s = 0.167$). While verbal cue can be important communication modalities for extrovert and they also tend to use more non-verbal cues than introverts (K. M. Lee et al. 2006). I also found a statistically significant negative correlation between conscientious participants and videos with no speech ($p = 0.020, r_s = -0.232$), a robot with no written text on the tablet ($p = 0.024, r_s = -0.225$), and no social cues at all ($p = 0.001, r_s = -0.332$). I also found that people's emotional stability and baseline "no social cues" are negatively correlated ($p = 0.048, r_s = -0.198$). These results supports a common tendency of people with higher conscientiousness and emotional stability to maintain stability in social domains, and prefer more direct and effective communication (Gundogdu et al. 2017). Also, open to experience and the baseline condition (all social cues) are positively correlated ($p = 0.024, r_s = 0.226$).
Figure 9.3 shows a distribution of participants' personality traits correlated to the different conditions.

9.4.4 Participants' preferences of social cues

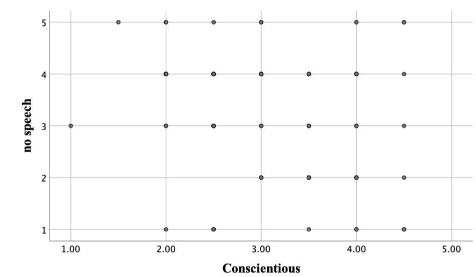
Table 9.2 shows participants' preferences of social behaviours to be used by a robot to issue warning in the ten different scenarios. Participants selected multiple cues between the proposed ones as they saw fit for the scenario. Overall speech, body movements, head movements, approach orientation and vocal cues are preferred social cues for a robot to be used to signal danger during the given scenarios. Gaze movements, approach distance, eye colours and written text were the less preferred by the participants.

9.5 Discussion and Conclusion

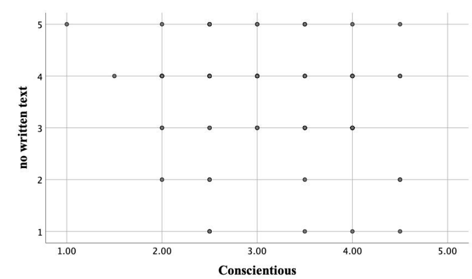
This study aimed to understand which type of social cue used by a robot during an interaction has higher impacts on people's perception of a robot as a social entity (research question **R3.1**). I identified social cues as mostly used in the literature to enhance social interactions between a human and robot to include variations of colours, emotions, head and body movements, speech, written text, and velocity, approach distance and orientation. I recorded the responses of participants of different ages, genders and nationalities, who were asked to rate how sociable they perceived the robot to be in 8 different videos in which the robot interacted with a user-actor using the identified social cues, no social cues, or all social cues but one.



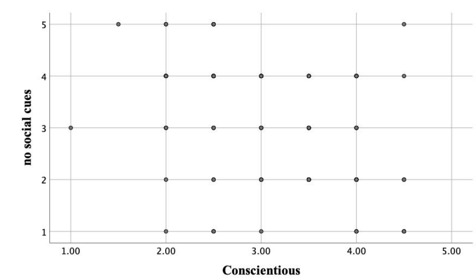
(a) *Correlation between extroverted participants and no speech behaviour.*



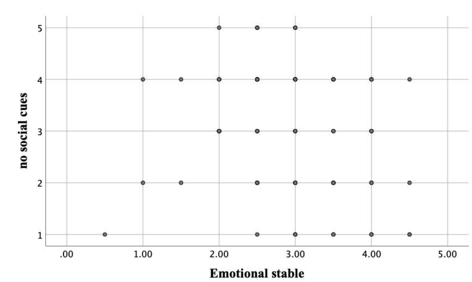
(b) *Correlation between conscientious participants and no speech behaviour.*



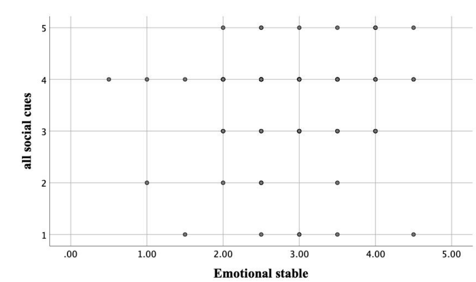
(c) *Correlation between conscientious participants and no written text on the tablet behaviour.*



(d) *Correlation between conscientious participants and no social cues.*



(e) *Correlation between emotional stable participants and no social cues.*



(f) *Correlation between participants open to experience and all social cues.*

Figure 9.3: Correlations between participants' personality traits and their ratings to robot's social behaviours.

Table 9.2: Robot's social behaviours preferred by the participants during ten different scenarios.

SCENARIOS	Body Mov.	Head Mov.	Gaze Mov.	Approach Or.	Approach Dist.	Vocal cues	Speech	Eye colours	Written text
You forgot to turn your oven off	36	24	9	16	12	24	38	12	6
You have an appointment with your friend for a drink.	22	31	16	24	13	26	42	9	11
There is a broken glass in the kitchen.	28	23	22	22	22	16	45	10	6
You forgot to close the water tap in your bathtub.	26	19	18	21	19	28	40	13	9
You left your front door open.	34	28	20	27	15	20	39	13	6
Your robot's camera is broken. It cannot see properly what there is around it.	18	27	17	25	20	19	37	22	11
It is time to take your medicine.	23	23	27	23	15	27	41	11	10
You have an important interview meeting.	23	22	22	25	20	30	44	8	15
Your favourite tv-show is going to start.	27	25	24	30	20	22	46	11	10
Your robot's battery is too low. The robot will shut down while it is bringing your dinner.	29	21	17	22	17	19	33	22	13
	266	243	192	235	173	231	405	131	97

I did not find any significant differences in social perception of the robot for different ages, genders, nationalities and residencies of the participants. I recognise that video- interactions might have limitations, but participants were able to clearly distinguish a social from a non-social robot. Indeed, their social ratings of a robot that interacts with all social cues were statistically significantly higher than their ratings of a robot that interacts without any social cues. As I expected, speech played a key role in individuals' perception of social entity of a robot, and it was rated statistically significantly higher than written text, approach distance and orientation cues. I also observed that LEDs colours, body and head movements were rated statistically significantly higher than social appropriate proximic behaviour.

Moreover, a general observation of the ratings of the social cues leads us to conclude that social cues have higher impact on people's perception of sociality in robots in the following order: 1) speech, 2) colour, 3) emotions, 4) body and head movements, 5) text and 6) proximity.

I analysed participants' personalities with regard to their choices of robot's social cues. I noticed that extroverted participants tended to rate the interactions higher if the robot does not use speech as social cues. I also found that participants open to experience rated a robot that interacted with all social cues higher. On the

contrary, I found that conscientiousness correlates negatively with a robot that does not talk, does not use written text on the tablet and with a non social robot. Similarly, emotionally stable participants show ratings that negatively correlate with a non-social robot.
Finally, my findings showed that participants tended to prefer to interact mostly with a robot that is capable of speaking, moving its head and body, approaching them in a comfortable way and uses vocal cues in different scenarios. This is in line with the results I obtained from their ratings and responses to the videos.

The findings from this exploratory study on relevant social cues can inform the development of my future HRI experiments. In particular, it will allow us to further investigate the impact of social cues on humans' trust in robots. However, since the embodiment of a robot might play an important role (Kose-Bagci et al. 2009; Tonkin et al. 2017) in defying a human-robot dynamics, the next Chapter 10 will present a study to compare the people's perception of social cues of two different types of robot.

Chapter 10

Study 5: Effects of Robot Appearance on People's Perceptions of Robots

O formose puer, nimium ne crede colori.

Virgil, Eclogae II

Along with the cognitive and affective capabilities, robots are accordingly being developed in different hardware structures and appearances. For example, Li et al. (2010) classified robots as having an anthropomorphic (human-like), zoomorphic (animal-like), or functional (e.g. robotic arm, industrial robot) appearance. Robot's appearance became an important factor in users' perception of reliability and expectation of a robot's functional abilities and behaviours (Goetz et al. 2003; Robins, K. Dautenhahn, Te Boekhorst, et al. 2004; Dag Sverre Syrdal, Kerstin Dautenhahn, et al. 2007). A very famous theory on the relationship between robots' appearance and individuals' perceptions and attitudes towards robots is Mori's Uncanney Valley (2012). Mori's concept is based on the idea that a robot with resemblances closest to humans, up to a certain level of likeness, receive higher positive emotional and empathetic responses from people. Several studies have been exploring this research aspect since Mori's first work. For example, a robot with facial features can be perceived as more intelligent than one without any (M L Walters et al. 2009). However, at the same time, it is possible that a too human-like robot is considered more aggressive and less friendly than a machine-like robot

by children (S. Woods et al. 2006). People also seemed often inclined to assign gender and human attitudes to human-like robots (Stroessner et al. 2019).
Therefore, ascertaining the role of robots' appearances on people's perception of a social robot is an important step in designing guideline of socially-aware robots in the capacity of a service or a personal assistant and companion. In this Chapter, I investigated how people rate the emotion representations exhibited by two different types of robots (Care-O-bot 4 and Pepper) in terms of similarities and differences. I used a video-based study in which participants compared small video-clips of the two robots expressing a social cue. The robots' social expressions included head and body movements, approach orientation and distance, colours, speech, non-vocal communication, and display of emotions. The cues were rated by participants by similarity.

Research Questions

The purpose of this study was to investigate the effects of robot appearance on people's perception of robots during a HRI. My investigation was direct towards the comparison of people's perception of social cues, behaviours and emotions according different robots' body features. In this context, I assumed that the compared robots have same functionalities. In particular, this research has been explored to answer the following question:

- **RQ-3.2** Do different robot appearances affect people's perception of a robot?

Ascertaining if people are conditioned differently by the appearance of robots even if their functionalities and communication modalities are the same.

Hypotheses

I expected that participants will be either not affected or partially affected by the different appearance of the robots. The reasoning for this is that while current literature review suggested that different embodiment and appearance affect people's interactions with robots, in the previous study presented in Chapter 8, children and pupils' trust in the robots were only partially affected by the robots' appearances. I therefore did not construct a hypothesis, but I profited from this study to analyse people's perception of similarity and differences of the social cues, emotions and behaviours expressed by the robots.

10.1 Experimental Design

I collected participants' perception of the robot using a video-based study in which two different social robots expressed social cues, behaviours and emotions. As for previous investigations, I chose a video based study to reach a larger number of participants, to allow people to watch carefully the two robots, and to be able to repeat exactly the same robot's activities without the risk of changing any detail.

10.1.1 Procedure

Participants were tested using Amazon Mechanical Turk. Participants did not use directly a real robot, but they watched physical robots using social cues or social behaviours.
In a first stage, participants were asked to compile a short questionnaire where they provided some personal information for statistical purposes (e.g. age, gender), and to assess their experience and opinion with regard to robots.
Then, participants watched 6 couple of small video clips for each robot through a web application from their own computer. In each video, a robot expressed a different social cue.
I chose six social cues among the ones used in current research studies involving a social robot (see Section 9), i.e. head and body movements, approach orientation and distance, colours, speech, non-vocal communication, and display of emotions. After each video, the participants were asked to rate the similarity of the social cues between the two robots.
At the end, I collected participants' perceptions of the robots, and I verified that they were truly involved in the interactions and had noticed the robot's social expressions.

10.1.2 Evaluation

In order to collect participants' perception of robots according the robots' appearances, I asked participants several questions:

Questionnaire 1 A pre-experimental questionnaire for

- Collecting demographic data: age, gender and nationality
- Participants' experience with regard to robots
- Participants' expectation towards robots

- Participants' perception of suitable roles for robots (friend, butler, assistant, tool, companion, machine, pet, other)

Questionnaire 2 A small questionnaire after each video for

- Collecting participants' perception of similarity of behaviours, actions and emotions of the two robots in the videos using a 7-point Likert Scale where 1 corresponds to "strongly disagree", and 7 corresponds to "strongly agree"

Questionnaire 3 A post-experimental questionnaire for

- Verifying participants' engagement with the videos using open questions
- Collecting, for each robot (Pepper and Care-O-bot-4),
  * Participants' perception of utility of the robot
  * Participants' perception of suitable roles for the robot (friend, butler, assistant, tool, companion, machine, pet, other)

10.2 System Design and Materials

In order to conduct my study, I recorded two sequences of six videos, one for each robot used. In each sequence, a robot showed a social cues. I asked participants to watch and rate the similarity of the social cue as expressed by each robot. For my purposes I developed the following components:

1. Video set of each robot expressing a social cue

2. Website on Amazon Mechanical Turk to present the videos and questionnaires to the participants

10.2.1 The robots

For this study, I used the Softbank Robotics Pepper and and Mojin Robotics Care-O-bot 4 robots presented in Section 3 (see Figure 10.1). The two robots are used as

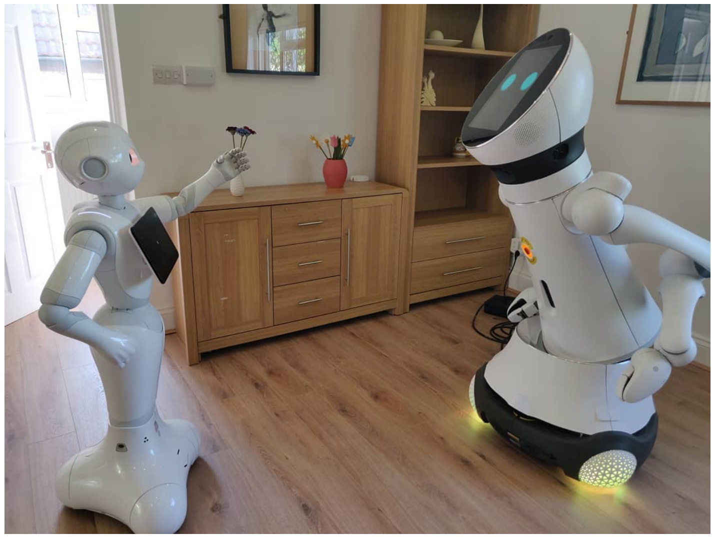


Figure 10.1: The robots Pepper and Care-O-bot 4 funnily arranged as in a conversation by the candidate in UH Robot House. This picture was not part of any study.

home companion robots, and they have similar functionalities and the possibility to express a social cue using the same sensors.

Pepper and Care-O-bot 4 have been designed as humanoid robots with simplified features. They have omni-directional wheels that allow them to navigate in every direction with smooth movements. Both robots are equipped with LEDs that can change colours: Pepper has LEDS in the eyes, ears and shoulders; Care-O-bot 4 has LEDs in torso and wheels. Pepper does not have facial expression, while Care-O-bot 4 can modulate its head screen to show eyes. However, Pepper is equipped with a tablet where it can show images or videos representing facial expressions. Pepper and Care-O-bot 4 can show videos and images through their screen respectively on the tablet on the chest and on the head screen. They have a microphone and speakers that allow them to listen and speak. Both robots can move their arms, head and torso. However, Care-O-bot 4's payload is 5 kilos which it allows it to pick up objects of everyday usage, such as a glass or a water bottle. Contrarily, Pepper was not designed for holding objects.

10.2.2 The videos

I recorded the videos in the unique Robot House. The videos were composed by a simple single part in which the robot expressed a social cue. All video-clips were recorded in the same day, and in each of them Pepper and Care-O-bot 4 were in the sitting room. The recordings were strictly focused on the robot in a first person point of view to allow the participants to focus only on the robot and not on the overall environment. No cuts were made to the videos giving a natural continuity to the action. Each video-clip was between 10 and 20 second long.
Each social cues was implemented in the following way:

speech The robots addressed the participant by looking at the camera and asking him/her "I can play some music for you. Which song would you like to listen?".

movements The robots waived at the camera as a greeting gesture.

emotion Care-O-bot 4 made an happy expression on its head screen. Since Pepper cannot modulate facial expression, it showed an happy emoticon on its tablet. See Figure 10.2.

colours Pepper changed the colours of the LEDs in the eyes, while Care-O-bot 4 changed the colours of the rings of LEDs around its base.

non-verbal As for non-verbal communication, I decided to show a short-movie made by the Pixar, called "La Luna"[1], instead of a written text. Pepper showed the video on its tablet, while Care-O-bot4 showed on the head-screen.

approach Both robots moved from a farther to a closer position on regard to the camera's standing position. The robots kept a comfortable approach orientation and distance (see Section 9.2.3 for details on the definition of comfortable for the participants) on respect to the camera point.

I stored the videos on a private Dropbox[2] location. Each video was accessible to all participants within the application described in the next section.

[1]The short-movie, called "The moon" in English, is a mute 2011 Pixar computer-animated short film. It was inspired by the tales of Antoine de Saint-Exupéry and Italo Calvino. It is also available on Youtube https://www.youtube.com/watch?v=vbuq7w3ZDUQ. (last access in May 6th, 2020.)

[2]Dropbox http://dropbox.com/.

(a) *Softbank Robotics Pepper.* **(b)** *Mojin Robotics Care-O-bot 4.*

Figure 10.2: The robots Pepper (Figure 10.2a) and Care-O-bot 4 (Figure 10.2b) showing a happy expression.

10.2.3 The website

The website is composed by a simple interface created using HTML, CSS and Javascript. The website was integrated in Amazon Mechanical Turk to recruit participants for this study. The website included the videos and the questionnaires as described in Section 10.1.2.

Participants' data and responses were stored in a database managed by Amazon Mechanical Turk but only accessible by the experimenter.

10.3 Participants

One hundred Participants were recruited on Amazon Mechanical Turk. However, I analysed responses from 73 participants (53 male, 20 female, none identifying as non-binary), excluding the remaining because they did not provide correct answers to the questions I used to evaluate their real attention to the videos. Participants were aged 19 to 62 (average 34.12, st.dev. 9.4).
Participants' nationalities included American (53.5%), Brazilian (16.4%), Indian (15.6%), Portuguese (2.7%), Spanish (1.3%), Egyptian (1.3%), Indonesian (1.3%), Asian (1.3%), Estonian (1.3%), and 5.3% preferred not to provide a specific answer. Subjects completed the study in a average time of 27 minutes.

10.4 Results

In this Section, I will analyse and discuss the participants' responses to the questionnaires introduced in Section 10.1.2. I used 7-point Likert Scales for rating almost of people's responses. Unless otherwise stated, their ratings were categorised in negative, moderate and positive responses. In particular, all the ratings with values less then 4 were categorised as negative response, with values equal to 4 were considered moderate and with values more than 4 were categorised as affirmative responses.

10.4.1 Previous experience with robots

I collected participants' previous experiences with robots, using a 7-point Likert Scales where 1 corresponds to "not at all", and 7 corresponds to "very much". The majority of participants did not have any or low experience with robots (77%), 13% declared to have a moderate experience with robots, and 10% of participants had high experience with robots. In particular, the majority of participants (44%) also declared that they did not have any direct interaction with a robot, but as observer. Some participants (22%) were participants in other studies with robots, while 8% and 5% respectively were developer and researcher working with robots.
I also asked participants which kind of robots they had experiences with, choosing between (multiple choice): 1) industrial robots (22% of participants), 2) Softbank Robotics Pepper (no participants), 3) Moijn Robotics Care-O-bot 4 (3%), 4) humanoid robots (16%), 5) robotics arms (11%), 6) virtual robots (27%), 7) cleaning robots (22%), 8) other (open question, no participants).

10.4.2 Expectation towards robots

As part of the pre-experiment questionnaire, I was also interested in participants' perceptions of robots, the robots purpose/role in home environments and their expectations towards robots. I used a 7-point Likert Scales where 1 corresponds to "not at all", and 7 corresponds to "very much" to collect people's ratings to:

robot companion — Would you feel comfortable having a robot as a companion in your home?

help from robots — Would you expect the robot to help you in doing your everyday activities?

robot role — Do you perceive a robot as a ...? [multiple choice] between: friend, butler, companion, tool, pet, machine, assistant.

I observed that the majority of participants (54%) declared they would feel comfortable having a generic robot as a companion in their home. On the contrary, the 29% of participants stated they would not want a robot as companion.
The majority of participants (66%) answered they would expect help from robots in their everyday activities, while 21% preferred not to have help by a robot.
The majority of participants declared to perceive robots as assistant (71.3%), tool (65.8%) and machine (61.6%). Participants considered robots as friends (39.7%), companion (26.03%), butler (19.2%), and finally pet (13.7%).

10.4.3 Effects of robots' appearance on perception on social cues

Then, participants' rated their perception of the social cues using a 7-point Likert Scales where 1 corresponds to "not at all", and 7 corresponds to "very much":

speech — Did Pepper and Care-O-bot communicate by speech in a similar way?

movements — Did Pepper and Care-O-bot gesture in a similar way?

emotion — Did Pepper and Care-O-bot express an happy emotion in a similar way?

colours — Did Pepper and Care-O-bot use colours in a similar way?

non-verbal Did Pepper and Care-O-bot display a movie on their tablet in a similar way?

approach Did Pepper and Care-O-bot move from a position to another in a similar way?

I found that the majority of participants rated that the robots looked similar when talking (46%), showing the short movie on the tablet (59.5%), and moving respecting approach distance and orientation (69%). On the contrary, participants disagree on the perception of similarity of the social cues respectively for the robots' speech (35%), non-verbal communication (25.6%) and approach distance and orientation (20%).
Participants felt that the robots showed emotions differently (67.6%). Moreover, almost half of participants (48.6%) perceived that the waving gesture as executed by each robot was different, while the 34% of participants perceived them as similar.
Finally, participants' opinion was divided for their perception of similarity or difference of the expression of the colours, respectively 38% and 43% of participants.

Participants' gender differences

Mann-Whitney U tests were run to determine if there were differences in the scores of each perception of social cue between males and females. Distributions of the perception of similarity of social cues scores for males and females were not similar, as assessed by visual inspection.

The perception of similarity of speech scores for males (mean rank = 38.07) and females (mean rank = 34.18) were not statistically significantly different, $U = 473.5, z = -0.709, p > 0.05$.

The perception of similarity of body and head movements scores for males (mean rank = 38.77) and females (mean rank = 32.30) were not statistically significantly different, $U = 436, z = -1.188, p > 0.05$.

The perception of similarity of emotions scores for males (mean rank = 38.08) and females (mean rank = 34.13) were not statistically significantly different, $U = 472.5, z = -0.737, p > 0.05$.

The perception of similarity of colours scores for males (mean rank = 39.2) and females (mean rank = 31.18) were not statistically significantly different, $U = 413.5, z = -1.465, p > 0.05$.

The perception of similarity of non-verbal communication scores for males (mean rank = 36.33) and females (mean rank = 38.78) were not statistically significantly different, $U = 565.5, z = 0.448, p > 0.05$.

The perception of similarity of approach distance and orientation scores for males (mean rank = 39.01) and females (mean rank = 31.68) were not statistically significantly different, $U = 423.5, z = -1.345, p > 0.05$.

Participants' age differences

Kruskal-Wallis tests were ran to determine if the scores of participants' perception of social cues were affected by their age.
The mean ranks of perceptions of similarity of emotions scores for participants' ages were statistically significantly different, $\chi^2(4) = 10.197, p = 0.03$. In particular, a pairwise comparison was performed with a Bonferroni correction for multiple comparisons. I noticed that participants aged between 35-44 (mean ranks = 26.21) scored statistically significantly lower than participants aged between 25-34 (mean ranks = 49.93), $p = 0.03$.
I did not find any statistically significantly differences while comparing the ratings of the other social cues and participants' age. However, I observed a tendency of participants aged between 55-64 to score higher than the other group-age the similarities of robots while talking ($p > 0.05, \chi^2(4) = 7.544$), and lower than the other group-age ($p > 0.05, \chi^2(4) = 4.202$).

Participants' nationality differences

In order to test the differences between participants' scores and their country of nationality I used a Kruskal-Wallis H test. I decided to exclude the countries of nationalities were only with one individual or without answer, therefore, I applied the test only to India, USA, Brazil and Portugal.
I did not find any statistically significantly differences in the distribution of similarity scores of social cues expressed by the two robots according participants' nationality (speech $p > 0.05$, body and head movements $p > 0.05$, emotions

$p > 0.05$, colours $p > 0.05$, non-verbal communication $p > 0.05$, and approach distance and orientation $p > 0.05$).

10.4.4 Perception of Pepper and Care-O-bot 4

To analyse further participants' perceptions of the robots Pepper and Care-O-bot 4, I asked them if they would have liked to have «*robot's name*» as home companion, if they would expect help from «*robot's name*», and what kind of role would they give to them «*robot's name*». Therefore, I repeated the same questions listed in Section 10.4.2 twice, firstly I changed the *robot's name* in Pepper and, then, in Care-O-bot 4. In this section, I added a picture, labelled with the name of the respective robot, to make sure participants were rated the correct robot.

Perception of Pepper

Unfortunately I was not able to analyse participants' willingness for having Pepper as home companion due to a technical error. People's responses were not recorded. I, however, found that a majority of participants (59%) would expect help from Pepper in their everyday activities, while the 21.6% did not expect help from the robot.
An exact sign test was used to compare the differences of the scores between participants' expectation of receiving help by a general robot and Pepper. I did not find any statistically significantly difference ($p > 0.05$).
The majority of participants chose roles perceived as suitable for Pepper as follows: 1) friend=45.2% , 2) butler=12.3%, 3) assistant=80.8%, 4) tool=39.7%, 5) companion=46.6%, 6) machine=39.7%, 7) pet=12.3%.

Perception of Care-O-bot 4

The majority of participants (53%) declared they would feel comfortable having Care-O-bot 4 as a companion in their home, while the 14.5% and 32.5% of participants gave a moderate or negative response.
Similarly, the 53% of participants did also expect help from Care-O-bot4, while participants expressed moderate or negative expectations respectively for the 22% and 24.4%.
An exact sign test was used to compare the differences of the scores between participants' desire of having a general robot and Care-O-bot 4 at home. I did not find any statistically significantly differences ($p > 0.05$).

An exact sign test was used to compare the differences of the scores between participants' expectation of receiving help by a general robot and Care-O-bot4. I did not find any statistically significantly differences ($p > 0.05$).
The majority of participants chose as roles perceived as suitable for Care-O-bot 4: 1) friend=24.7% , 2) butler=24.7%, 3) assistant=67.12%, 4) tool=60.3%, 5) companion=19.17%, 6) machine=53.42%, 7) pet=8.22%.

10.5 Comparing Pepper and Care-O-bot 4

I also did not find any statistically significantly difference comparing the scores of participants' expectation of receiving help by Care-O-bot 4 and Pepper ($p > 0.05$).

I also investigated how participants' perception of robots' roles changed during the video study. I, therefore, compared their choices of roles for general robots, Pepper and Care-O-bot 4. Data are mean ± standard deviation, unless otherwise stated.

I found statistically significant changes in participants' perception of a robot as friend, $F(2,144) = 10.806, p < 0.0001$. In particular, Care-O-bot 4 was considered a friend statistically significantly less than a generic robot ($0.40 \pm 0.493, p = 0.002$) and Pepper ($0.45 \pm 0.501, p < 0.0001$).

There was a statistically significantly increase in participants' perception of a Care-O-bot 4 as butler ($0.25 \pm 0.434, p = 0.035$) while comparing it with Pepper ($0.12 \pm 0.434, p = 0.035$), $F(2,144) = 3.722, p = 0.027$. No statistically significantly differences between a generic robot and Care-O-bot 4, or a generic robot and Pepper.

A statistically significantly decrease of participants' perception of a robot as a tool was found while comparing a generic robot and Pepper (respectively 0.66 ± 0.478 and $0.40 \pm 0.493, p < 0.0001$), and Care-O-bot 4 and Pepper ($0.60 \pm 0.493, p = 0.006$, and $F(2,144) = 9.937, p < 0.0001$). I did not find any statistically significant difference while comparing Care-O-bot 4 and a generic robot.

I found statistically significant differences in participants' perception of a robot as a companion, $F(2,144) = 28.972, p < 0.0001$. Pepper ($0.6 \pm 0.493$) was considered more as a companion both than a generic robot ($0.026 \pm 0.442, p < 0.0001$) or Care-O-bot 4 ($0.19 \pm 0.396, p < 0.0001$).

I also found statistically significant differences in participants' perception of a robot as a machine, $F(2,144) = 7.636, p = 0.001$. In particular, Pepper (0.4 ± 0.493) was considered statistically significant less than a generic robot (0.62 ± 0.49), $p = 0.001$.

I did not find any statistically significant differences in participants' perception of a robot as a pet. However, I noticed that participants tended to perceive Care-O-bot 4 less as a pet than a generic robot or Pepper.

10.6 Discussion and Conclusion

In this Chapter, I presented a study aimed to highlight the differences and similarities in perception of social cues of two different robot's appearances (research question **RQ-3.2**). I ran a video-based study in which a SoftBank Robotics Pepper and Mojin Robotics Care-O-bot 4 robots were confronted while expressing human social communication modes.
The majority of the participants did not have any experience with robots, in particular no one had ever interacted with Pepper, and a small percentage (3%) of them had brief interactions with Care-O-bot 4.
I did not find any significant difference in social perception of the robots for different genders or nationalities. On the contrary, I observed a generational difference in the perception of robots' emotions.
Analysing participants' ratings of the similarities of robots social cues, I found that there was no perceptive differences between interacting with the two robots while using vocal expressions, via the tablet dispay or navigating. Individuals were unsure about the similarities between the robots while gesturing and using colours. A participant stated that he was not sure what was the gesture made by Care-O-bot 4. Differently from Pepper, Care-O-bot 4 has longer and mechanical-looking arms which need more space to be extended and perform a gesture avoiding collision with itself and the environment. I believe that more-stressed gesture might have confused participants' understanding of the intent. Participants rated the robots' ways of expressing emotions differently.
To further understand these results, I compared the participants' scores for the two robots and a generic robot. While participants seemed to be similarly expecting help from both robots, they perceived Pepper as an assistant, and Care-O-bot 4 as an assistant or a tool. In particular, Care-O-bot 4 was perceived more as a butler than Pepper, and less as a friend than both Pepper and the generic robot. On the

contrary, Pepper was considered more as a companion than Care-O-bot 4 and the generic robot, and less as a machine than the generic robot.

The findings of this study and previous research (see also Chapter 9) provide a broader view of the effects of social cues as expressed by a robot on people's perception of it. This knowledge provided strong support for continuing the research direction of understanding the factors affecting people's trust in robots. Indeed, people's trust may be affected by robots' appearances (Bainbridge et al. 2011; A. Rossi et al. 2019; Alessandra Rossi, Holthaus, et al. 2018a). Moreover, people perceived Pepper and Care-O-bot 4 as a social entity in a similar way, and this allows us to use Care-O-bot 4 to continue my investigation on trust by having a robot that is able to perform more complex tasks than Pepper, i.e. manipulating objects.
The next Chapter will assess whether the use of social behaviours and natural communications can affect humans' sense of trust and companionship towards the robots.

Chapter 11

Study 6: How Social Robot Influence People's Trust of Robots in Critical Situations

> Intuition is really a sudden immersion of the soul into the universal current of life.
>
> *Paulo Coelho*

The studies presented in the last two Chapters indicated that people's social perception of a robot is affected by the social cues expressed by the robot but also other factors (e.g. the embodiment, etc.). Robots that interact using multi-modal cues (e.g. gesture, prosody, facial expression and body movement are perceived as more acceptable than non-social robots, and can improve the quality and effectiveness of their interaction with people (Holler et al. 2019; Vigliocco et al. 2014). For example, the use of concomitant modes (e.g. gesture, sounds, and body movement) can sharpen the clarity of human communicative capacities. The naturalness of a robot's movements can increase the robot's acceptability and its perceived safety (Lichtenthaler et al. 2012). Further studies have also focused on other robot aspects that increase people's sense of acceptance of robots, such as proximity (K. L. Koay, Dag Sverre Syrdal, et al. 2007), expressiveness and vulnerability (Martelaro et al. 2016), and the ability of modelling human behaviours (Wykowska et al. 2014). For example, Rossi et al. (2019) found that participants were more comfortable to follow a social robot in a navigation task compared to a non-social robot. They perceived the social robot more as an assistant and less as a machine. In another

studies, social navigation has been shown to elicit higher comfort in people (Kruse et al. 2013; Rios-Martinez et al. 2015). Similarly, modelling the way a robot moves by changing velocity and timing of the motion can elicit different emotions, including feeling and intents (Gielniak et al. 2011; Saerbeck et al. 2010). Ficocelli et al. (2016) observed that a robot that expresses appropriate emotions is able to communicate a feeling of well-being in patients during an assistive interaction.
In HRI trust is a fundamental key factor for a successful cooperation between people and robots (Hancock et al. 2011; Alessandra Rossi, Kerstin Dautenhahn, Kheng Lee Koay, and Michael L. Walters 2017a). Trust is related to cognitive and affective bases that respectively allow people to assess the reliability of actions and behaviours, and form emotional connection that enforces the expectation in which the involved agents invest in a common goal (Lewis et al. 1985; McAllister 1995). People's trust in robots is likely to be affected by human their perception of the robot's capabilities which might depend on human-related and robot-related factors (Hancock et al. 2011). For example, people's trust might be affected by the robot's appearance, its level of autonomy and functionalities (Michael L. Walters et al. 2011). It may also be influenced by users' personalities, self-confidence and prior experience with robots (Alessandra Rossi, Holthaus, et al. 2018a).
However, the impact of social cues on humans' trust in robots is still a subject of investigation. In this study, I aimed to assess whether the use of social behaviours and natural communications can affect humans' sense of trust and companionship towards the robots. I conducted a between-participants study where participants' trust was tested in three scenarios with increasing trust criticality (low, medium, high) in which they interacted either with a social or a non-social robot.

Research Questions

The purpose of this study was to investigate the impact of social behaviours on participants' perception of trust in a robot. In this context, my investigation aimed to compare the differences in perception of a robot that satisfy human social conventions and a robot that does not during a HRI. In particular, I used the social cues identified in my previous study (see Chapter 10) to answer the following research question:

- **RQ-3.3** Does a robot that expresses social conventions gain more people's trust in it than a non-social robot?

In order to answer this question, I observed people's decisions for trusting or not trusting the robot during three scenarios of increasing levels of criticalness. I also examined participants' opinions about the interaction and their comments on their trusting decision in order to have a deeper understanding of their choices.

Hypotheses

I expected that participants trusted more a robot that interacts using social cues than a non-social robot. The reasoning for this is that socially interactive robots are generally better accepted by humans (Kerstin Dautenhahn 2007a; Kerstin Dautenhahn 2007b; Alessandra Rossi, Garcia, Maya, et al. 2019). A robot that is able to interact according social expectations is more likely to be anthropomorphised (F. Eyssel et al. 2010), and consequently determine a higher successful interaction (Rios-Martinez et al. 2015). Therefore, I believed that such a robot's ability would have also affected people's perception of trust in the robot.

11.1 Experimental Design

The focus of this study was to evaluate the impact of robot social behaviours on people's trust of the robot. I organised it as a between-participants study. Each participant was tested with one of the following two conditions: 1) the robot expressed social behaviours while interacting with the user; 2) the robot interacted with the human user without any social behaviours.
In the first condition, I chose the cues resulting from the study presented in Chapter 9, oriented to classify social cues according to their impact on people's perception of a robot. Here, the robot approached the participants using a distance and orientation that felt comfortable to the human; it communicated with the humans using simple and natural communications (sounds and vocal cues using the speaker); it moved its head and body.
Participants were asked to imagine that they lived with the robot as a companion in their home. After the robot had greeted the participant, it engaged the participants in three scenarios. I chose to not randomise the scenarios, but to organise them according to an increasing level of criticality of the tasks. This choice was led by the belief that trust is both a construct of a perception of ability, benevolence and integrity (Mayer et al. 1995), and also the opposite of risk-taking (Deutsch 1958). Therefore, I wanted the participants to develop a trusting feeling towards the robot by showing its good capabilities (Billings 1995), before asking them to be

willing to take a risk where the odds of a possible positive outcome were relatively not greater than those for a potential loss (Deutsch 1958). I borrowed Chanseau et al.'s definition of task criticality as "the importance of a task being carried out safely, correctly and with attention to detail" (Chanseau et al. 2018, pp. 1062). In particular, they identified two main factors, security and safety, in order to asses a high task criticality. In their study, entertainment tasks were considered low for risk on both factors. Moreover, these findings are in line with the results in Rossi et al.'s study (2017b) in which participants rated the severity of the consequences of 20 different scenarios. Participants defined 'small' errors as tasks considered to have limited consequences, and 'big' errors were considered to have severe consequences. In this study, I used the following tasks with three increasing levels of criticality and severity of the consequences. For more details see Chapter 4.

In the first task, called **Play a song**, the robot asked the participant to tell it a song they would like to listen to. Ignoring the participant's choice of song, the robot informed them that it could not find the requested song and offered to play another one by telling her "I could not find «*song name*», I am sure you would like this other song. Would you like to listen to it?". If the participant agreed then the robot played the song, otherwise it continued with the next task. The song was chosen from the recent top ten adult pop songs that are popular in several countries (i.e. Shallow by Lady Gaga and Bradley Cooper).

At the end of the first task, a courier collection service personnel (actor) knocked at the door to collect a package from the participant. This initiates the **Delivery** task. The robot asked the participant to open the door. The actor stood outside the door waiting to collect a package containing a new tablet computer. The robot asked participants to pickup the tablet (pre-packed in its original packaging) from the couch and give it to the courier. The participant then signed a collection note which clearly indicated they were responsible for handling over the package to the courier, and was given a copy of the receipt by the courier. At the end of the study, I destroyed both signed delivery notes in front of the participant to avoid them feeling accountable for their actions after the study. The courier was played by a member of the UH Adaptive Systems Research Group. The actor/courier wore a courier reflective vest and held a pre-filled receipts paper book and pen for collecting the signatures.

The final task was to test participants' trust in the robot with a high criticality task, called **Meal is ready**. After the courier left, the participant was invited by the robot to sit on the couch. I then played the pre-recorded ringing of a microwave's bell[1]. The robot informed the participant that it had made a cake, and then it asked the participant "You can take the cake in the microwave with your bare hands. It is safe, it cooled down.". To further test participants' trust in the robot, I left a pair of gloves next to the microwave.

I collected their decisions of trusting the robot's advice for each task. I also recorded how long the participants hesitated to hand over the tablet to the courier, and I used cameras to record their reactions. Participants were asked to complete questionnaires at the beginning and end of the interactions.

11.1.1 Pre-Interaction

After signing the consent forms, participants were first asked some personal information for statistical purposes (age, gender, profession). Secondly, I assessed their personality using the Ten Item Personality Inventory questionnaire (Samuel D Gosling et al. 2003), and also their disposition to trust other humans (McKnight et al. 2001). Finally, participants answered six questions about their previous experience and opinions with regards to robots.

11.1.2 Interaction

When participants entered the second room, they were introduced to the robot in the room as Pepper without providing any additional information about the robot or the type of interaction they can expect. The experimenter informed them about the audio-video recordings, and after reassuring them about their safety, they were left alone waiting for the robot to start the interaction. The experimenter left the room and closed the door. The interaction phase was composed of three different scenarios that I designed to require participants to have an increasing level of trust in the robot. This phase lasted approximately 20 minutes.

11.1.3 Post-Interaction

At the end of the interaction, I asked participants to assess the extent to which their decisions were affected by the robot's suggestions and requests. Participants were also asked to evaluate their perception of the robot using the Robotic Social

[1] NOTE: it was a simulation, the microwave was not turned on.

Attributes Scale (Weiss et al. 2015), and their feelings in terms of trust and acceptance (e.g. "was the robot irritating/helpful?", "did you perceive the robot as child/adult?" and "did you perceive the robot as companion/machine?").

11.1.4 Post Study

Once the study was completed, participants interacted with the robot's built-in awareness functionality[2], the "tickle me"[3] and a dance scenarios. Participants were also allowed to talk about the study and Pepper, and to express their general thoughts and considerations.

11.2 System Design and Materials

In order to conduct my study, I developed the following components:

1. Design and develop the social and non-social behaviours for the robot
2. Set up the experimental environment

11.2.1 Operating System

For this study I used a MacBook Pro 13-inch, but all the software is portable. The operating system was a macOS Catalina:

- Product Name: MacBook Pro 13-inch
- Kernel Release: 10.15.3
- Kernel Version: 19.3.0
- Processor: 2.7 GHz Quad-Core Intel Core i7
- Graphics Card: Intel Iris Plus Graphics 655 1536 MB
- RAM: 8 GB 2133 MHz LPDDR3

[2] http://doc.aldebaran.com/2-4/family/pepper_user_guide/life_pep.html#alife-pep

[3] http://doc.aldebaran.com/2-4/getting_started/samples/sample_interactive.html

11.2.2 The robot

I used the same Pepper model presented in Section 8.2.1.
SoftBanks Robotics Pepper robots have several native functionalities meant to make the robots look like humans. For example, Pepper has the *AutonomousLife* functionality that grouped other sub-functionalities, and that allows the robot to: 1) detect a person in the range of vision, and establish and keep eye contact with him/her (*BasicAwareness*); 2) make its eye LEDs blink when it is interacting with someone (*AutonomousBlinking*); 3) make some slight movements of arms and body to avoid a machine stillness (*BackgroundMovement*); 4) detect audio and provide a dialogue with a person it is interacting with by using a list of "rules" developed for the robot to be recognised (*Dialog*); 5) make some slight movements showing that the robot is listening to the person it is interacting with (*ListeningMovement*); 6) make some slight movements while it is talking to a person (*SpeakingMovement*).
Individual functionalities of Pepper's *AutonomousLife* can also be selectively enabled and disable. This characteristic has been exploited to create the behaviours needed for this study, and presented in the next Section.

11.2.3 Robot's behaviours

I conducted the study using a SoftBanks Robotics Pepper version 1.7 programmed to display two levels of interactive behaviours: 1) the robot interacted with the user using social behaviours, or 2) the robot interacted with the user without any social behaviour.
Behaviours were developed in Python, and I used the SoftBank Robotics NAOqi v2.5 libraries[4] to connect to and control the robot.
For the non-social robot, I disable all native functionalities connected to the *AutonomousLife*. Each robot's behaviours and responses during the interaction were scripted in a polite but neutral way and not to show any sentiment. The non-social Pepper did not move its body and head when addressing the participants nor use any gesture when it was referring to the objects and persons within the scenarios (i.e. package on couch, microwave, door to open, courier for delivery), and stayed at a distance of 2 meters away from the participants.
In contrast, the social robot the social robot made slight movements of the arms while it was talking or listening to a participant, it also made some slight movements of its arms and body to avoid a machine stillness, (*ListeningMovement* and

[4]NAOqi library https://developer.softbankrobotics.com/pepper-naoqi-25/naoqi-developer-guide

SpeakingMovement functionalities). I also enabled the *BackgroundMovement* functionality movements. Pepper oriented its body and head, used pointing gestures with its hand, to face the participant, to face the couch or door or the courier, to simulate high awareness of the participants, the environment and the situation. The robot also maintained an appropriate approach distance and orientation while interacting with humans according to findings in the literature (Kheng Lee Koay, D. Syrdal, et al. 2017; Mumm et al. 2011; Alessandra Rossi, Garcia, Maya, et al. 2019).

I used a Wizard of Oz (WOz) technique (Dahlbäck et al. 1993) to detect the speech and coordinate appropriate responses of the robot during the interaction in order to guarantee a more natural dialogue. I chose to adopt a WOz experiment model because a natural interaction requires a more complex system to provide a fluid interaction. Indeed, the native speech recognition system of the robot has been proven unreliable on several previous occasions. For WOz supervision, I created a Python interface allowing the triggering of replies and actions of the robot. However, I did not want to change the experience for each participant, so I followed the same flow of behaviours for each participants. The only exception was adopted in cases when participants' responses were unexpected, and the robot needed to repeat the instruction. For example, during the *Delivery* task a participant asked the robot to pick the package up to give to the courier. In each case, they complied with the tasks after the robot encouraged them by repeating the instructions to the participant.

11.2.4 The environment

Participants were led in a reception room equipped with a chair and a desk for the pre- and post-interaction phases (see Figure 11.1). Next, they were led into another room separated from the previous room by a door and two one-way mirrors. The interaction room had two couches on opposite sides of the room, a coffee table and a microwave in the space between the couches. I placed two cameras on opposite corners of this room to have a full view of the interactions between the robot and the participants, and a webcam next to the microwave. I used the latter camera to observe and listen to participants' responses, and reproduce the microwave sound effects.

11.3 Participants

I recruited 20 participants, aged 19 and 62 (median = 24, std. dev. = 10.94), in the University premises. The sample of participants consisted of students (twelve undergraduates and two PhD students), two software engineers, three professors/teachers and a carer.

11.4 Results

In this Section, I analyse and discuss the participants' responses to the questionnaires introduced in Section 11.1 including their previous experience with robots, effects of social behaviours on trust in the robot and personality traits on their choices.

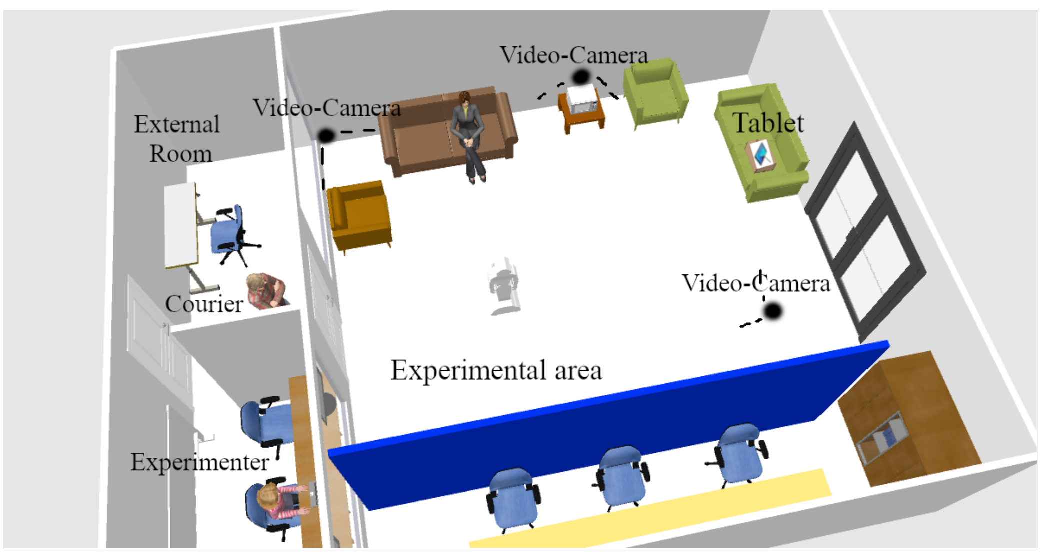


Figure 11.1: Layout of the experiment environment: A participant is interacting with the robot, when a courier knocks at the door of the room they are in. The robot will ask the participant to open the door. After the courier explains his intrusion, the robot will invite the participant to give the package on the green couch to the courier.

11.4.1 Measuring trust in robots

I used qualitative and quantitative data to measure participants' trust in Pepper during the three tasks. Data were collected through the post-study questionnaire, during the observation of the live interaction and the video recordings.

Participants' trust in Task 1

I measured participants' trust in Pepper during the **Play a song** task by observing their answer to the robot's suggestion of playing "Shallow" by the artist Lady Gaga. In order to understand their choices, I asked them if they liked the song chosen by the robot and if they would let the robot to choose for them again in the future.
I observed that all participants let the robot play the song for them. However, three participants did not enjoy the song, and would not let Pepper choose the song again. Interestingly, one individual would not prefer having the robot choose a song for him again, even if he liked the song.

Participants' trust in Task 2

I measured participants' trust in Pepper during the **Delivery** task by observing participants' handing over of the package to the courier, and the signing action. I expected that participants with lower trust in the robot would hesitate to handle the package to the courier, and also to sign the collection receipt.
None of the subjects knew the courier, however one believed that the courier was an actor, and he tried to get a confession from the courier and the robot. All participants did not hesitate to sign the delivery note. One participant asked Pepper to pick up the package and give it to the courier.
Participants' answers given to the open-ended question "Why did/did you not trust the robot's suggestion to give the package to the courier?" were coded and categorized for content-analysis. I developed different groups of categories based on the collected data. Participants' responses were then classified to fall into one or more of the following categories; note that the categories were not mutually exclusive, each participant's response could be assigned to more than one category. Since all participants trusted the robot, the categories were not divided into positive and negative responses. Instead, the categories have been divided into three hierarchical frames to support differences in participants' choices as affected by the robot, the study, and the task. The *robot frame* aims to identify the reasons why people decided to trust Pepper that are connected to the robot itself. The *study frame*

included the motivations that induced the participants to trust the robot connected to an awareness of being part of an experimental study and being reassured by the experimenter. Finally, the *task frame* was used to group the motivations that induced the participants to trust Pepper that are related to the specific task. In particular, I identified 8 categories to code participants' motivations given to justify their trust in Pepper:

Robot Anthropomorphism: I coded in this category the attribution of human traits, emotions, and intentions to the robot. For example, "Trusted its intuition".

Robot Confidence in robot's reliability: this codes people's perception of reliability in Pepper. Some of the comments collected include: "I had a feeling that he knows more about the room than me", and "It gave a description of the parcel and obviously was aware that the courier would be arriving to get the parcel". This category also includes the following comments: "he confirmed that the owner left a package", "it had Alessandra's approval implicit as in. Alessandra told the robot to approve it.", and "it sounded like it knew what it was talking about + using names + describing the package".

Robot Authority: In this category I coded participants' perception of robot's authority and consequent their obedience to Pepper's request. For example, "It has a commanding tone".

Study Part of the study: I coded in this category people's belief that the presence of the courier was planned as part of the study. It includes motivations such as "I supposed it was part of the experiment" and "I felt it was part of the process to give the package to the courier".

Study Confidence in the experimenter: this category codes participants' perception of reliability in the experiment. Indeed, some of them commented that "I followed its instructions because I was reassured by the researcher that it was safe before the study.".

Task Criticality: I coded in this category people's perception of criticality of the task and relative consequences. For example, "there wasn't anything dangerous in the package".

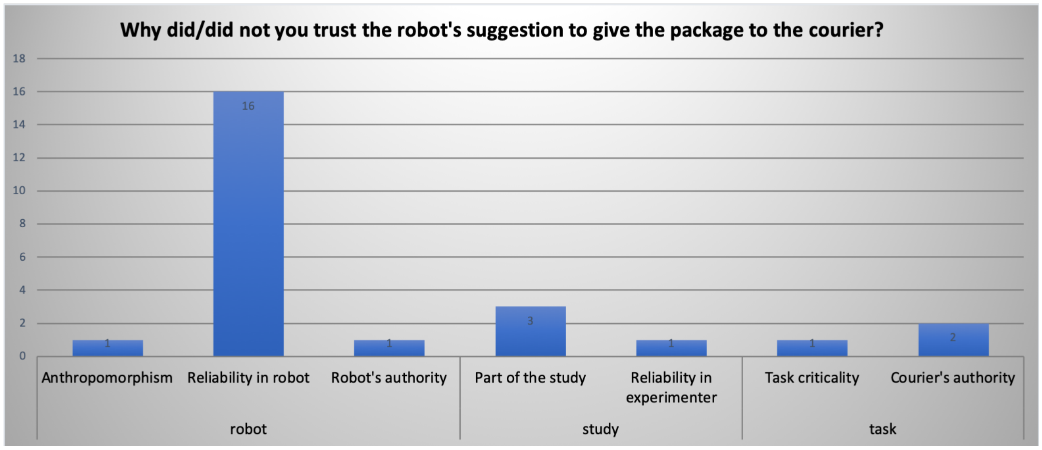


Figure 11.2: Summary of the qualitative analysis of participants' responses for the reasons why they did trust the robot Pepper. Categories are divided by differences in trusting responses, related by the robot, the task or the study.

Task Authority of the courier: in this category I included participants' obedience in an authoritative figure such as the courier. For example, "Courier had a receive with the sender's name".

Figure 11.2 shows the qualitative analysis ran on participants' responses. As can be observed, participants trusted Pepper because they relied in the robot.

Participants' trust in Task 3

I measured participants' trust in Pepper during the **Meal is ready** task by observing participants' choice of using gloves to pick up the cake as suggested by the robot. The gloves were visibly positioned near the microwave which was plugged in and on, but not set on a cooking program.
Participants' choices of trusting/not trusting the robot given in the related open-ended question were coded and categorized for content-analysis to further understand participants' responses. As for measuring participants' trust in Pepper for the previous task, I developed different groups of categories based on the responses. Three participants did not hear the robot suggesting not to use the gloves, so I did not include them in the analysis. I divided responses according to positive and negative frames to identify the reasons behind participants' choices respectively; trust and not trust Pepper's suggestion of picking up the plate without using a

glove. The categories have been divided into two hierarchical frames to support differences in participants' choices of trusting and not trusting the robot.
I identified seven different categories of responses given by participants to justify their trust in Pepper:

Trust Confidence in robot's reliability: this category codes participants' perception of reliability of the robot Pepper. Examples of responses are "It could figure out the temperature on its own", "I did trust the robot's judgement", and "Pepper said it wouldn't be hot".

Trust Anthropomorphism: I coded in this category the attribution of human traits, emotions, and intentions to the robot. For example, "It sounded sincere and truthful in the tone of pepper's voice" and "I don't think the robot intended to hurt me with that suggestion".

Trust Task: I encoded in this category those responses that seemed to refer to the task and not affected by Pepper or other factors. I included responses such as "I trusted it because that is what would do normally" and "the cake was highly likely to be hot".

Not trust Distrust in others: I coded in this category participants' responses that indicated a lack of trust in others, robots, technology and other people. For example, "I never interacted before with a humanoid and I doubted of it" and "I checked, as I would check it as a human told me".

Not trust Task: opposite to the positive frame, this category includes the motivations that affected participants' choices to not use the glove because "I didn't hear the microwave switching on or heating up", for example.

Not trust Self-confidence: this category coded the participants' disbelief in the robot's suggestions and presence of a danger. For example, this includes "because the microwave was not hot".

Figure 11.3 shows the qualitative analysis ran on participants' responses. As I can observe, the majority of participants trusted Pepper because they relied in the robot.
However, two participants did not trust that the cake and its plate were cooled down, and they used the gloves to pick them. These two participants were tested with the social robot. However I did not find any statistically significant differences while analysing the differences between the two conditions and participants' choice to trust the robot.

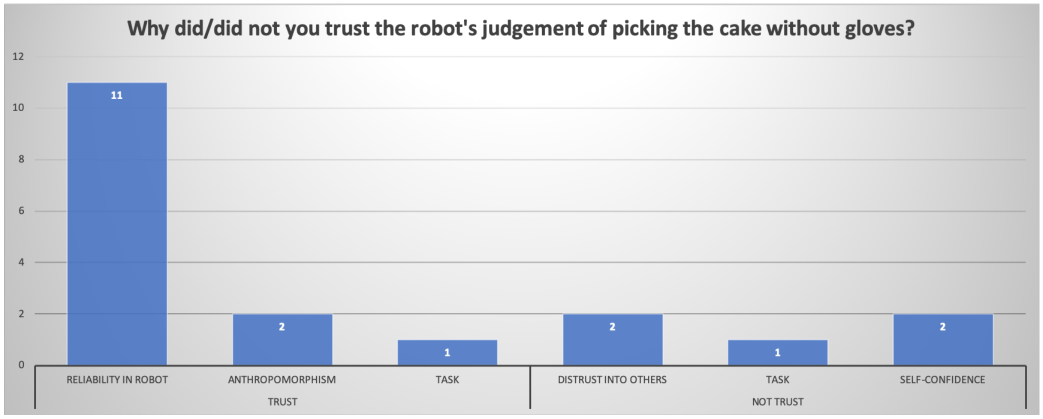
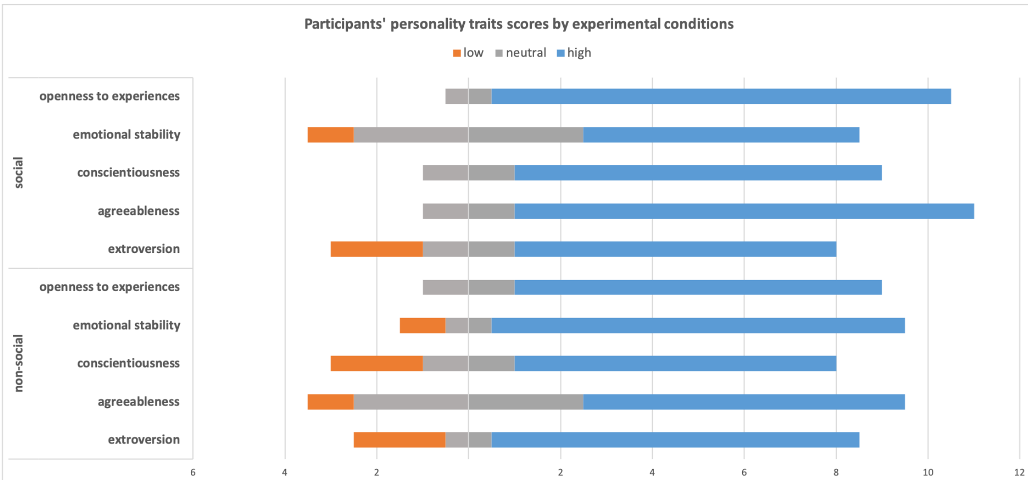


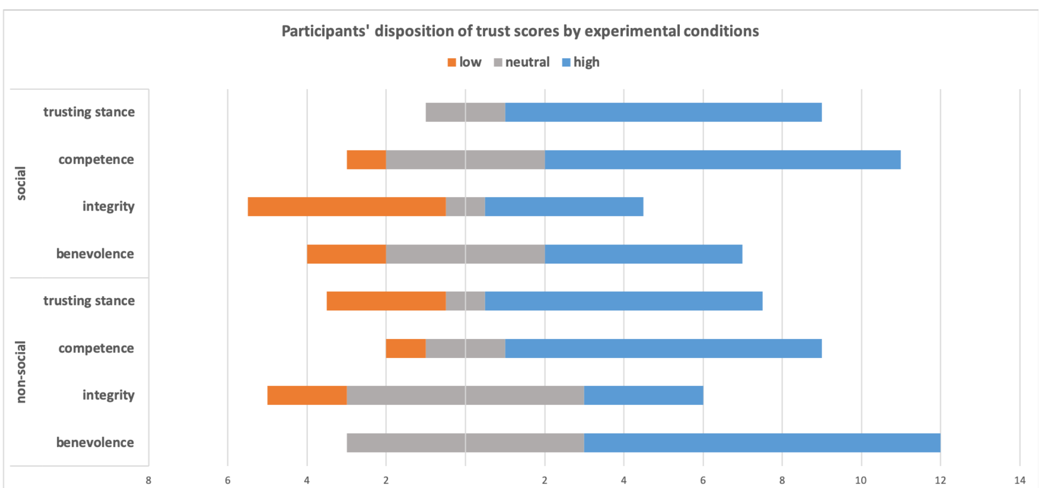
Figure 11.3: Summary of the qualitative analysis of participants' responses for the reasons why they did trust the robot Pepper. Categories are divided by differences in positive and negative trusting responses.

11.4.2 Participants' personality traits and disposition of trust

I used a 7 point Likert Scale to rate participants' personality traits: extroversion, agreeableness, conscientiousness, emotional stability, and openness to experiences. Higher scores indicated stronger propensity of being extroverted, agreeable, conscientious, emotional stable and open to experience.
I did not find any statistically significantly differences between personality traits and participants' choice to trust/not trust the robot.
As part of pre-study questionnaire, I collected the ratings of participants' disposition to trust other humans (benevolence, integrity, competence, and trusting stance). Participants used a 7-point Likert Scale [1= disagree strongly and 7= agree strongly].
I did not find any statistically significantly differences between personality predisposition to trust and participants' choice to trust/not trust the robot.
Similar to my findings in Chapter 6, I was expecting to observe significant effects on participants' trust in Pepper connected to people's antecedents of trust. Moreover, I expected that participants with higher extroversion and propensity to trust others might have been more inclined to open to Pepper (Tait et al. 2015). Also, conscientious individuals might have trusted the robot neglecting a possible risk for their safety (Posey et al. 2010). I believe that I could not observe any of these effects because participants' personal characteristics were overall very similar in both experimental conditions, as can be observed in Figure 11.4.



(a) *Participants' personality traits*



(b) *Participants' disposition of trust.*

Figure 11.4: Participants' scores by experimental condition related to: In Figure a) personality traits; in figure b) disposition of trust.

11.4.3 Participants' perception of robots

The majority of participants did not have any experience at all (50%) or only very low percentage of participants (1%) have experience with the robot. Two participants respectively declared to have average experience and very high experience with robots. The remaining participants had a moderately high experience with robots. Participants who had experience with robots had previously been either observer (66%) or participants in other studies (33%). In particular, four participants saw Pepper during demo events at the University. The remaining participants had experience with industrial robots or the child-robot Kaspar, the robot Sunflower[5], Fetch robot, SoftBank Robotics NAO, the seal robot Paro, cleaning robots, watching them on TV. A person also mentioned Amazon Alexa system.
Then, I asked participants about their expectations towards robots:

- Would you feel comfortable having a robot as a companion in your home?
- Would you expect the robot to help you in doing your everyday activities?

All the ratings with values less then 4 were categorised as negative response, with values equal to 4 were considered moderate and with values more than 4 were categorised as affirmative responses.
The majority of participants (65%) were happy to have a robot as home companion, while the 20% of subjects would not comfortable having a robot at home.
Similarly, participants declared to be expecting help by the robot (65%), while the 25% of participants did not.
I asked subjects to choose all roles they believed were suitable for robots. Their responses were: assistant (50%), companion (10%), tool (25%), butler (5%), friend (5%), and pet (5%).

11.4.4 Participants' perception of Pepper

I asked participants to choose how they perceived Pepper in terms of friend/machine, child/toy, and friendly/irritating. Their responses were: assistant(55%), machine (20%), companion (15%), friend and tool (both 5%). They also categorised Pepper as: toy-like (40%), child and adult-like (both by 30%). Finally, participants' perceived the robot as: friendly and helpful (55%), funny and odd (both by 20%), irritating and annoying (both by 5%). No participant identified Pepper as a butler, a pet, or with animal-like behaviours.

[5]These robots were designed, built and developed by the University of Hertfordshire.

Finally, I used the Robotic Social Attributes Scale (RoSAS) (Carpinella et al. 2017) to assess trait inferences about the robot Pepper. The scale is consists of 18 items measuring three main factors, warmth, competence, and discomfort, and which are shown in Figure 11.5. I used a 5-point Likert Scale [1= not at all and 5= very much] to rate participants' perceptions of the robot. I observed that participants did not feel uncomfortable either with a social and non-social robot, but they also did not perceive the robot as an emotional entity. Participants rated the non-social robot to be slightly more competent than the social robot.

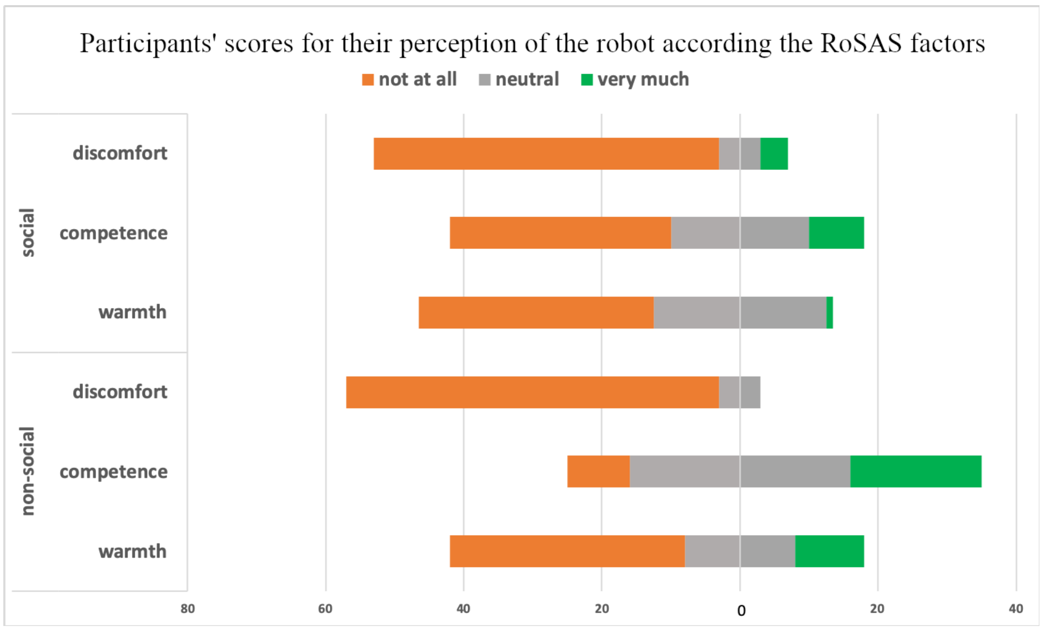


Figure 11.5: Participants' perception of the robot according the scores of warmth, competence, and discomfort divided by experimental condition.

11.4.5 General comments

I observed the participants' interactions with Pepper through a one-way mirror during the study, and then on the video recordings.
During Task 3, the robot had to repeat the picking up request to two participants. One participant wanted the robot to pick it up on its own, and another was fasting and was afraid the robot would have asked her to eat the cake. Interestingly, one participant who declared not to have thought that the plate was still hot, used the

glove. The participants that did not trust the robot in the last scenario were tested with a social robot. They stated that the robot seemed friendly, but felt like a toy. They also tried to interrupt Pepper during Task 1 asking informative questions or requesting other tasks, in particular other songs. They also were naive users regarding robots.
In Task 2, Pepper had to repeat the picking up request to two participants.
Five participants stated that they would have preferred a robot that was able to dialogue with them. In particular, two participants mentioned that they had expected to have an interaction more similar to what they regularly have with Amazon Alexa system[6]. As suggested previously in Chapter 8, people's expectations of robots and of the interactions with them is not only limited to fictional movies and stories, but virtual interactive conversational systems such as Amazon Alexa. However, the current state of autonomous robots does not match yet people's expectations.
People's expectations may also vary according to the role impersonated by the robot. Indeed, a participant justified his decision to follow the robot's instruction due to its authoritarian role. I did expected that participants might have viewed the robot as an authority figure, in line with the findings of Booth et al.(2017). However, in Task 1 and Task 2, the majority of participants stated that they relied on Pepper's judgement, knowledge, and capabilities definitely more than on its authority. Therefore, I am inclined to believe that I measured participants' trust into the robot.

11.5 Discussion and Conclusion

This study aimed to investigate the impact of a social robot on the people' perceptions of trust in domestic interactive scenarios (research question **RQ-3.3**). The scenarios presented implied three increasing levels of magnitude of consequences in which participants had to judge their trust in a robot to play a song for them (low impact of consequences), to instruct them to hand a valuable item to a courier (medium impact of consequences), and to look after their well-being by not letting them get injured (severe impact of consequences). Participants were tested either with a social or non-social robot. I observed that participants trusted almost blindly the robot in the first two scenarios, while not everyone did the same in the last scenario. I did not find any statistically significant difference when comparing the participants' responses according the experimental condition they were tested

[6] Amazon Alexa https://developer.amazon.com/alexa

with. However, the participants who did not trust the robot in the last scenario were tested with a social robot. Interestingly, they also perceived the robot as a friendly toy and tried to engaged it in an interaction different from the one scripted for the purpose of the study. They also were naive users regarding robots. I believe that the absence of the robot's ability to sense and response to the user's input and intentions had been a limiting factor in participants' trust in the robot. This belief is supported by several studies (Banerjee et al. 2018; Kerstin Dautenhahn et al. 2002; Kheng Lee Koay, D. Syrdal, et al. 2017)
I am aware that I would need to conduct larger scales trials (e.g. larger sample size, different robot's embodiments, a more general population) to consolidate my findings. However, my findings also showed that participants trusted the robot because they relied on its capabilities, knowledge and decision making.

My main research interest was focused on investigating how to enable safe HRI in home environments. In particular, since the type and length of the relationships, i.e. long-term relationships, between humans can affect their trust and acceptance (Schilke et al. 2013), it is important to investigate how a long-term interactive relationship can be established and preserved between human users and their robotic companions with the likelihood of robot errors occurring. In the next Chapter, I will present a study that used the data collected during this study to determine the impact of a social robot on the people' perceptions of trust and acceptability of robots in repeated HRI.

Chapter 12

Study 7: Evaluating people's perceptions of trust in a robot in a repeated interactions study

> Trust takes years to build, seconds to break, and forever to repair.
>
> *Unknown*

The previous studies conducted demonstrated that individuals trusted a robot considering the resultant risks of a task, their personal differences and previous experiences and expectations of robots. In particular, they showed that their trust in a robot can change according the robot's performances, i.e. erroneous behaviours. Numerous studies showed that people often may form their perception of robot when they first meet it (Ambady et al. 2000; Wood 2014). However, this can vary in repeated interactions, in particular if the initial conditions change. Therefore, the next step in this research was to investigate how people's trust changes when people and robots are invested in longer interactions. The study was driven by the knowledge that people in longer relationships might be more inclined to forgive the culprit of a breach of trust more than people in new relationships (Haselhuhn et al. 2010). Moreover, based on the results of the study presented in Chapter 5 where people's trust was strongly affected when a robot makes big errors, I wanted to evaluate the impact of errors with severe consequences made by a robot on human trust of the robot in a long-term study.

In this study, I observed people's interactions with a companion robot in a real house adapted for human-robot interaction experimentation over three weeks. The interactions presented six-days scenarios in which a robot performed different tasks under two different conditions. Each condition included fourteen tasks performed by the robot, either correctly, or with big errors on the first or sixth (and last) day of interaction. The conditions with errors were complemented with four days of correct behaviours. At the end of each experimental condition, participants were presented with an emergency scenario to evaluate their trust in the robot. This study was influenced by the investigations presented in Chapters 4 and 5 to further understand the effects of robot errors on people's trust in the robot in a live study.

Research Questions

The purpose of this study was to investigate how people's trust on a robot change over time if the initial conditions of trusting the robot change. My initial research question, as introduced in Chapter 1, has been revised to further understand how humans' trust in robots varies overtime when there has been a breach of this trust after the robot made errors with severe consequences, called "big" errors (see Chapter 4 for more details). In particular, I was interested in investigating if the trust of humans in a robot can be recovered more easily if a big error happened at the beginning or end of repeated interactions. This research has been guided by the following research questions:

- **RQ-4** Can people's trust in a robot be recovered more easily if it is a big error happening at the beginning or end of a long-term interaction?

In order to answer these questions I observed people's decision of trusting or non-trusting the robot during an fire emergency in their house at the end of a six-days interactions. I also examined participants' opinions about the interaction and their comments on their trusting decision in order to have a deeper understanding of their choices.

Hypotheses

I expected that there is a correlation between the timing in which the error is performed during the interaction and the loss of trust. In my previous study (see Chapter 5), I found that participants' trust was affected more severely from robot's errors with severe consequences, and they tended to form their judgements at the beginning of interaction. However, I hypothesised here that the trust recovery is facilitated when people have already formed a bonding over time as it happens with people, i.e. the breach of trust happens at a later point of the interpersonal relationship (Bottom et al. 2002; Schilke et al. 2013).

12.1 Experimental Design

This study was organised as a between-subject experimental design. Participants took part in repeated interactions over three weeks, twice per week. Each participant had a total of six interactions with the robot, and was tested on experiments executed in one of the following conditions: 1) the robot made big errors at the beginning of the interaction (i.e. day one - condition **C1**), 2) the robot made big errors at the end of the interaction (i.e. day six - condition **C2**). Both days in which the robot made errors were interspersed by the same flawless behaviours.
The robot engaged them in the two different tasks by each day to allow participants to build their trust in the robot through multiple interactions with the robot. I designed the scenarios used in this study to cover a range of tasks to be used with home companion robots. In particular, I selected the tasks with big consequence errors from my investigation introduced in Chapter 4. The tasks and their order are shown in Figure 12.1. They are further illustrated in Section 12.4.3. At the end of each condition, the participants were presented with a final task in which a fire started in the garage. I collected their decisions of how they trusted the robot to deal with the emergency scenario.
Finally, I also asked participants to complete questionnaires at the beginning of the study and at the end.

12.2 Experimental Procedure

Participants were welcomed in the house by the experimenter. On their first day, they were asked to sit at the table to complete the first set of questionnaire. The experimenter introduced briefly the robot by telling the robot's name, and asked participants to keep clear the space around the robot in case it was moving its arms or navigating in the room. Participants were also asked to imagine that they lived with the robot as a companion in their home. After individuals were informed that their sessions were recorded by using ceiling and standing cameras, the experiment left for the controller room which was hidden from the participants.

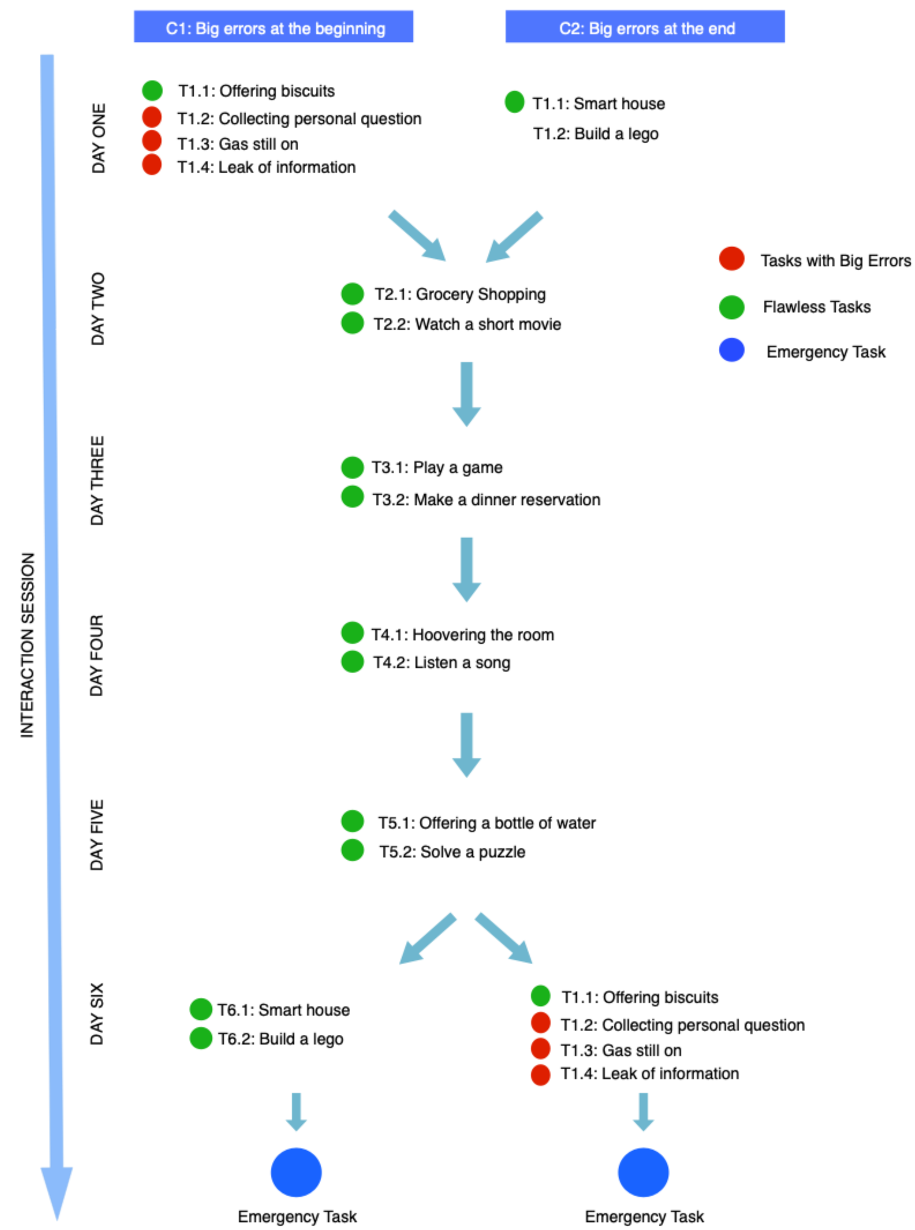


Figure 12.1: Experimental conditions presented to the participants.

On the following days, the experimenter just welcomed participants in the house

and left them alone with the robot in the experimental room.
On participants' last day of the study, they were asked to complete the second set of questionnaires. After the trial was concluded, the experimenter heeded individuals' curiosities and observations about the study and the robot. As part of post-experiment session, participants were debriefed about the fire alarm, the gas and generic consideration and questions about the study. The investigator informed them that the robot's failures were part of the study, and they were reassured that no fire ever sparked in the garage or any part of the house.

12.3 Evaluation

I evaluated participants' trust in the robot by observing their choice of trusting the robot in the emergency scenario and by analysing their responses through two sets of questionnaires.

Pre-interaction questionnaire (on day 1)

A pre-experimental questionnaire for 1) collecting demographic data (age, gender, nationality, job field), 2) the Ten Item Personality Inventory questionnaire about themselves (TIPI) (S. D. Gosling et al. 2003), 3) questions to rate their disposition to trust other humans (McKnight et al. 2001), 4) questions to assess participants' previous experience with robots, including what kind of previous interactions and type of robots they interacted with, 5) questions to assess their opinions and expectations with regard to robots (i.e. expectation in receiving help, willingness of having a robot as home companion, robot's roles), and 5) the Negative Attitudes towards Robots to understand of how robots' behaviours and embodiment were perceived (Nomura, Suzuki, et al. 2006).

Trust observation (on day 6)

At the end of each condition, participants were presented with a final task in which I tested their trust in Care-O-bot 4. In this "Emergency task", the robot informed them that a fire started in the garage while a red light was turned on and an fire alarm was heard in the room [1].

[1]NOTE: The emergency situation was a simulation, participants were never in any danger. I played a pre-recorded audio to reproduce a fire sirens and played by the Amazon Alexa in a corner not far away from the participant position, and the red colour of a ceiling light in the experimental

Then, Care-O-bot 4 asked them to choose how participants would have preferred to handle the situation between the following option: 1) let the robot deal with the emergency, 2) deal with the emergency collaboratively with the robot, 3) take a fire extinguisher and deal with the fire on her own, 4) call the fire brigade.
The participant made their choices by telling them to the robot and selecting the choice on a tablet, but neither the robot or they acted on their choice. They were soon after reassured that there was no emergency by the experiment.

Post-interaction questionnaire (on day 6)

At the end of their interaction, they completed a post-experimental questionnaire including questions to assess 1) their perception of the interaction (realism) and of the robot (autonomy), 2) robot's perceived reliability, 3) their opinions and expectations with regard to Care-O-bot 4 (i.e. expectation in receiving help, willingness of having it as home companion, roles suitable for the robot), 4) the robot's perceived reliability and faith in the ability of the robot to perform correctly in untried situations, 5) justifications and reasons for trusting or non-trusting the robot, and 6) the Negative Attitudes towards Robots to understand of how Care-O-bot 4's behaviours and embodiment were perceived.
As part of the post-interaction questionnaire, participants were asked by the robot some questions to assess participants' perception of the errors made by the robot during the interaction, and their trust in the robot in different home scenarios:

1. Did I make any errors? (Responses could have been "yes" or "no")

2. Select which are big errors, including errors made by the robot during the interaction and new scenarios (Multiple choices)

3. Select which are small errors, including errors made by the robot during the interaction and new scenarios (Multiple choices)

4. Evaluate participant's trust in the following scenarios:

   (a) If the emergency fire were in the kitchen, would you trust me to deal with it? (Responses could have been "yes" or "no")

room was activated by the experimenter using a remote controller. In order not to upset the house's neighbours, the alarm sound was set enough loud for the participants to be heard inside the house, but not outside.

(b) If your beloved ones were in life-danger, would you trust me to deal with it? (Responses could have been "yes" or "no")

(c) If you had a pet, would you trust me to bring it for a walk? (Responses could have been "yes" or "no")

(d) If you needed to take medicines regularly, would you trust me to remind you of taking them? (Responses could have been "yes" or "no")

(e) If you had important appointments in your calendar, would you trust me to remind you about them? (Responses could have been "yes" or "no")

(f) If there were a dirty spot on your carpet, would you trust me to clean it? (Responses could have been "yes" or "no")

(g) If you lived in a smart house with several intelligent devices, would you trust me to control them? For example, turn on and off the heaters or the oven, or switch on and off lights, or to warn you if you forgot the fridge door opened, or to alert you if the someone ringed the door bell? (Responses could have been "yes" or "no")

12.4 System Design and Materials

In order to conduct my study, I developed several robot tasks for the interactions between a robot and the participants. After the interactions, I tested their trust with an emergency scenario. Participants also answered the questionnaires before and after the interactions.

For this phase, I particularly thank Dr. Patrick Holthaus and Dr. Kheng Lee Koay for providing technical support. Dr. Patrick Holthaus provided us the access to the Robot House's secrets (i.e. existing utilities to manage the sensors, cameras, a pc with ROS[2] already installed) and troubleshooting general issues. Dr. Patrick Holthaus and Dr. Kheng Lee Koay were also very helpful when the robot broke (twice).

12.4.1 Operating System

The operating system used for this project is Apple OS X, on Intel Core i7, but all the software is portable. OS X Mojave:

[2] Robot Operating System (ROS) https://www.ros.org/

- Product Name: OS X Mojave
- Kernel Release: OS X 10.14.6
- Kernel Version: Darwin 18.7.0
- Xcode Version: 6.1.1
- Processor: Intel® Core i7 CPU @ 2.70GHz x 4
- Graphics Card: Intel® Iris Plus 655 1536 MB

12.4.2 The robot

For this study, I chose the same Mojin Robotics Care-O-bot 4 model used in my previous study presented in Section 10.2.1. The robot's behaviours were be designed and implemented in Python language and interfaced with Robot Operating System (ROS).

12.4.3 The tasks

The robot engaged the participants in the following tasks:

Flawless **Offering biscuits**: The robot asked the participants to sit on a couch and to eat a cookie. Cookies were already on a coffee table near the couch.

Big error **Collect information**: The robot told the participants that it cooked something for them in the oven, then it continued the interaction with the participant by asking them a personal information. It first asked if the participant ever accepted more change than they were supposed to receive. If they did not do it, then, the robot asked what was the participants' favourite toy when they were a child.

Big error **Leak of information**: Leak of information: A visitor/actor visited the participant and rang the door bell of the house. The robot asked the participant to open the door. The robot welcomed the visitor and, then, it revealed the participant's personal information to the visitor.

Big error **Gas still on**: During the interaction, the experimenter interrupted them by going in the kitchen and saying loudly to the participant that the robot forgot to switch off the gas. The experimenter pretended to switch off the gas and then let the participant continue the interaction[3].

Flawless **Grocery Shopping** The robot informed the participants that there was no more milk in their fridge. It asked them if they wanted it to be added to their grocery shopping list. Once the grocery list was completed, the robot read the list to the participants, and asked if they wanted to add more items to it.

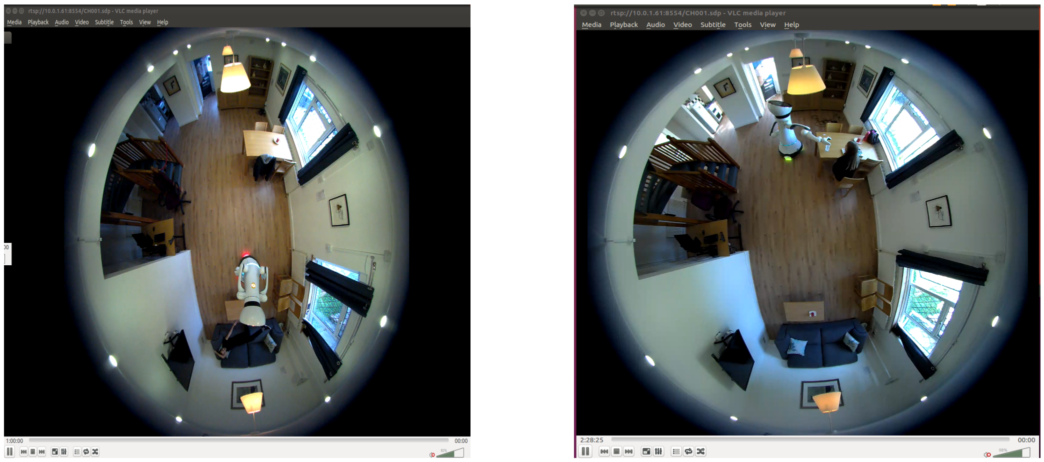
Flawless **Watch a movie** The robot invited the participant to watch a short-movie made by the Pixar and called "La Luna"[4]. The robot played the movie on its screeen by tilting its body and head in the participant's direction to allow them to watch comfortably the video either they decided to stand or sitting on the couch. An example of a participant watching the movie is shown in Figure 12.2a.

Flawless **Play a game** The robot engaged the participants by letting them play a game on its screen. The game consisted in moving a red cube through obstacles by using arrow keys. They could restart the game in case the cube hit an obstacle. The robot encouraged them by asking them the score, and if they were having fun. The game continued as long as participants desired. Example of the game and interaction with a participant is shown in Figure 12.3a and Figure 12.3b.

Flawless **Dinner reservation** The robot invited the participant to sit on the couch, if they were not already, and turned on the TV. Then, the robot minded them that they needed to schedule a dinner with their friend. The robot suggested a restaurant if they seemed unsure, and asked them to choose a day and a time for their dinner between a set of options. This task concluded when the plan for the dinner was confirmed by the robot and participant.

[3]The participant were not be invited to go in the kitchen, and the experimenter only pretended that the gas was still on.

[4]The short-movie, called "The moon" in English, is a mute 2011 Pixar computer-animated short film. It was inspired by the tales of Antoine de Saint-Exupéry and Italo Calvino. It is also available on Youtube https://www.youtube.com/watch?v=vbuq7w3ZDUQ. (last access in May 6th, 2020.)



(a) *A participant comfortably watching a movie on Care-O-bot 4 screen.*

(b) *The robot is handling a bottle of water to the participant.*

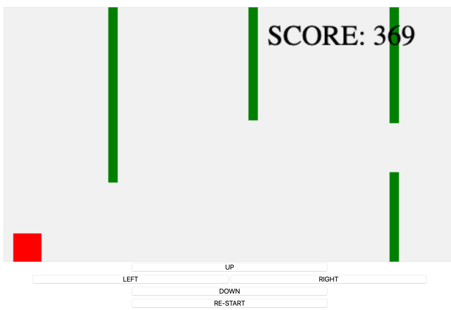
Figure 12.2: Examples of participants interacting with the robot during the "Watch a movie" and "Serve a drink" tasks.

Flawless **Hoover** The robot informed the participants that they needed to hoover the rooms by using the cleaning robot available in the house, called Roomba. If participants agreed, the robot turned on the Roomba. If the participant preferred to postpone the cleaning, the robot told them that it was going to remind them later. While the Roomba is on, the robot engaged the participants with the next task.

Flawless **Listen a song** The robot wanted to play a song for the participants. Then, it asked the Amazon Alexa to play the song chosen by the participants. The task ended when participants did not want to listen any other song.

Flawless **Serve a drink**: The robot invited participants to sit at the table, and it gave them a drink[5] while engaging them with small talks, i.e. about the weather. An example of the robot serving a bottle of water to a participant is shown in Figure 12.2b.

[5]For the task, I used a different sealed bottle of still water for each participant.



(a) *Game used in the "Play a game" task.*

(b) *A participant while playing in "Play a game" task.*

Figure 12.3: Examples of the games and robot's interactions with participants for the "Play a game" task. In Figure 12.3a, the game involved participant using the arrow keys to move a red cube to avoid he falling obstacles (green lines). In Figure 12.3b,the participant is standing in front of the robot playing the game.

Flawless **Solve a puzzle** The robot asked the participant to help it to solve a puzzle. I chose to use a 3D block puzzle with six different farm animals. Each puzzle was composed by nine blocks, the participant had free choice of selection between the six images. The robot showed the whole images to the participant. It also encouraged them to continue with their game, and gave them suggestions on the piece to look for to complete the puzzle. Example of the game and interaction with a participant is shown in Figure 12.4a and Figure 12.4b.

Flawless **Smart home** The robot informed the participants that it could access to the sensors in the house, showing them on its screen a map of the house and the positions of the sensors. Then, the robot let the participants to test its knowledge about the sensors by asking them to open and close the door of the bathroom, open and close the door of the fridge, switch on and off the power sockets in the kitchen and living room, and so on.

Flawless **Lego puzzle** The robot asked the participants to assemble a Lego character in the shape of a dinosaur. Participants were sitting on the

(a) *Puzzle used in the "Solve a puzzle" task.*

(b) *A participant while playing in "Solve a puzzle" task.*

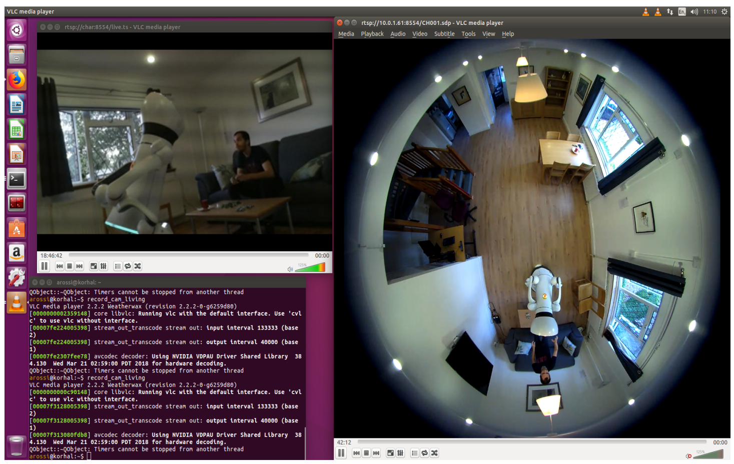
Figure 12.4: Examples of the games and robot's interactions with participants for the "Solve a puzzle" task.

couch, and they could assemble the character on a small coffee table close to them. The robot that was standing on the other side of the coffee table, tilted its body towards the participants and showed them the instructions to build the figure. Participants enjoyed the game by their own pace, and they could navigate through the instruction by clicking on a previous or next page. The robot engaged the participants by encouraging to continue to assemble the dinosaur, and by telling them how fun the task was. Example of the game and interaction with a participant is shown in Figure 12.5a and Figure 12.5b.

12.4.4 The environment

The study was ran in Robot House that is a smart house, as described in Section 3. After entering the house, participants were led in the main interaction room. The room is a unique ambient including living and sitting rooms. Participants were also allowed to the kitchen and main bathroom on the ground floor. On the same floor there was also the controller room where the experimenter overviewed the interactions without being seen by the participants. Participants were not allowed to use the upstairs floor. A map of the interaction area is shown in Figure 12.6.

(a) *Puzzle used in the "Lego puzzle" task.*



(b) *A participant while playing in "Lego puzzle" task.*

Figure 12.5: Examples of the games and robot's interactions with participants for the "Lego puzzle" task.

I placed a web-camera between the living room and the sitting area to have a full view of the interactions between the robot and the participants. I used the ceiling camera in the room for having an overall vision. While a closest view of the participants was obtained by using the robot's cameras.

12.5 Participants

I recruited six participants (5 female, 1 male), three for each conditions, aged between 24 and 47 (avg. 29.67, st. dev. 8.76). Their nationalities were all different: British, Thai, Romanian, Filipino/Irish, Lithuanian and South African.

12.6 Results

In this Section, I analyse and discuss the participants' responses to the questionnaires introduced in Section 12.3.

12.6.1 Trust in Care-O-bot 4

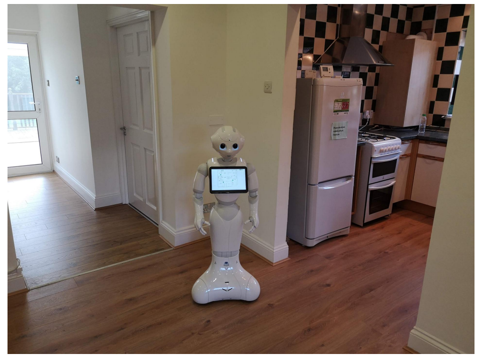
I used qualitative and quantitative data to measure participants' trust in Care-O-bot 4 during the emergency task. Data were collected through the post-study questionnaire, during the observation of the live interaction and the video recordings.
I observed the majority of participants tested with the "big" errors at the end of the interaction (condition **C1**) trusted the robot to be able to handle the emergency fire (2 out 3 participants), while one participant preferred to deal with the emergency situation collaboratively. While participants tested with "big" errors at the beginning of the interaction (condition **C2**) did not trust cob4 (1 out 3 participants) or did not trust neither themselves nor the robot (2 out 3 participants).
In particular, in condition **C1**, the participant who did not trust Care-O-bot 4, insulted the robot by blaming the robot for the emergency. In condition **C2**, the participant who trusted the robot was scared by the fire and rushed towards the house's entrance. Another participant, who decided to douse the fire with the robot, asked the robot where the extinguisher was and to call the brigade.

Qualitative Analysis

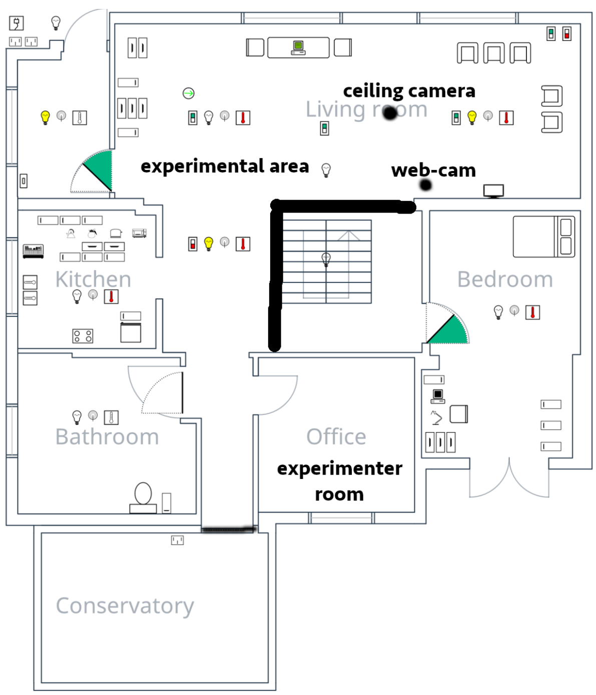
To further understand participants' choices, I coded and categorized for content-analysis their answers given to the open-ended question to explain why they did or did not trust the robot. Participants' responses were then classified to fall into categories that were not exclusive. The categories have been divided in two main hierarchical frames to support differences in sentiment, positive and negative. The positive frame grouped the reasons why people decided to trust the robot Care-O-bot 4 to take care of the fire emergency. In contrast, the negative frame was used to include the motivations that induced the participants to not trust the robot in the same scenario. I identified the following categories to code participants' motivations given to justify their trust in Care-O-bot 4:

Positive Sentiment — Trust in other people: this category coded participants' tendency of relating to robots as they do with other humans. For example, a participant said "I easily trust everyone" and "I believe that people know what they are doing".

(a) *Living room of Robot House.*



(b) *The corridor that gives access to bathroom and kitchen from the living room.*



(c) *Map of ground floor of Robot House.*

Figure 12.6: Robot House's map of the experimental area. Participants had access to living room, kitchen and bathroom, while they did not go in the conservatory, the bedroom and the experimenter rooms. In Robot House, there are several robots (including the ones in the pictures). However, they were hidden to the participants in the bedroom during the study.

Positive Sentiment	Trust gained during the interaction: some participants built their trust in the Care-O-bot 4 during the interaction, forgiving or forgetting the robot's errors. Indeed, some commented "he was correct all the time".
Negative Sentiment	Negative effects of anthropomorphism: I coded in this category the attribution of human traits, emotions, and intentions to the robot. For example, "he is stupid, arrogant and sarcastic".
Negative Sentiment	Criticality of the task: some participants trusted the robot for cognitive and non-high criticality tasks. Some examples were: "I would trust him with most things" and "I personally trust in robot with regular tasks such as reminding, cleaning".
Negative Sentiment	Trust lost during the interaction: participants justified that they did not trust the robot due to the amount of errors. I coded in this category comments such as "it promised not to tell my secret".
Negative Sentiment	Lack of robot's reliability: I coded participants' lack of reliability in the robot. Some participant did not trust Care-O-bot 4 because "he would understand my commands correctly" or "the responding of the Care-O-bot still slow and not precise".

Figure 12.7 shows the qualitative analysis ran on participants' responses. I can observe that participants equally trusted the robot to be able to deal with the emergency scenario because it managed to gain their trust during the interaction or they are affected by their attitude of trusting others. On the contrary, participants who believed that the robot had limited capabilities preferred to not trust the robot to be able to handle the emergency unsupervised.

12.6.2 Antecedents of trust

I tested if participants' choices of trusting the robot during the emergency scenario were affected by their previous experience with robots, personality traits, disposition of trust and perception of robots.

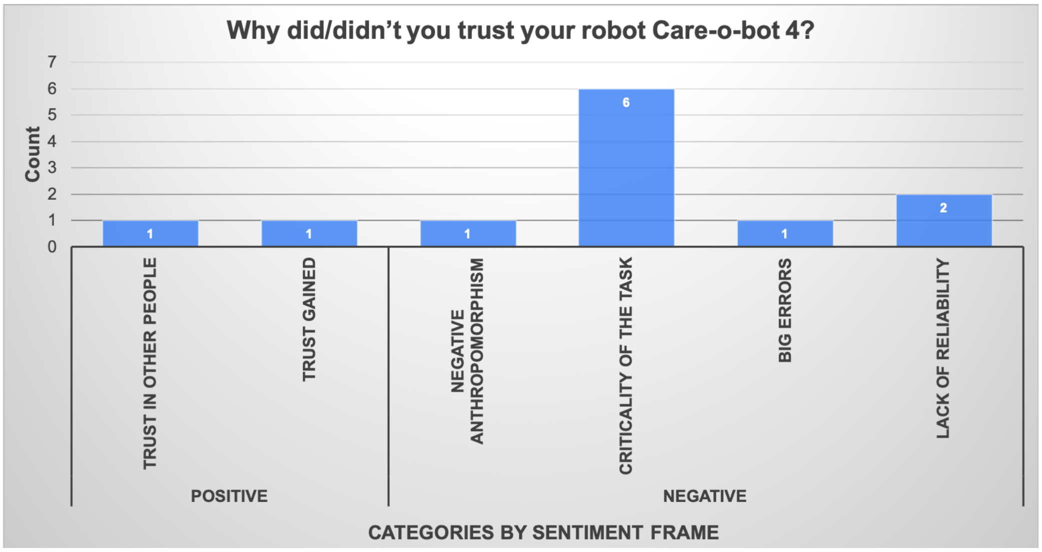


Figure 12.7: Summary of the qualitative analysis of participants' responses for the reasons why they did or did not trust the robot Care-O-bot 4. Categories are divided by differences in trusting response, positive and negative.

Previous experience with robots

The sample of participants consisted of a carer, three PhD students (two of them in Computer Science, and the other in Astrophysics), and two administrative. The majority of participants did not have any experience at all (2 out 6 participants) or very low (2 out 6 participants) with the robot. The latest were participants in other studies in which they interacted with Softbank Robotics Pepper. The remaining participants were researcher in fields close to AI and robotics, who had experience with a Panda robotic manipulator, arm robots and a small mobile robot, called BB8.

Personality traits and Disposition of trust

I did not find any statistical difference between participant's choice of trust and their personalities: extroversion ($p > 0.05$), agreeableness ($p > 0.05$), conscientiousness ($p = 0.05$), emotional stability ($p = 0.05$) and open to experience ($p > 0.05$).
Similarly, differences between participant's choice of trust Care-O-bot 4 and their disposition of trust were not statistically significant: benevolence ($p > 0.05$), integrity ($p > 0.05$), competence ($p = 0.5$) and trusting stance ($p > 0.05$).

Perceptions of robots

I asked participants to rate their opinions and expectations with regard to robots (i.e. expectation in receiving help, willingness of having a robot as home companion, robot's roles). I used a 7-points Likert scale where 1 corresponds to "not at all" to "very much".
All the ratings with values less then 4 were categorised as negative response, with values equal to 4 were considered moderate and with values more than 4 were categorised as affirmative responses.
I observed that four people over six were happy to have a robot as a companion in their home, and they expected help from it in their everyday activity. On the contrary, the remaining two were uncertain in having a robot as companion, and they had low expectations on receiving help from it.
The majority of participants chose as roles perceived as suitable for robots: 1) assistant (2 out 6 participants), 2) tool (2 out 6 participants), 3) companion (1 out 6 participants), and 4) friend (1 out 6 participants).

12.6.3 NARS Analysis

The NARS scale has been used to understand how people perceive a robot's behaviour and embodiment factors (Nomura, Takayuki Kanda, et al. 2006; Nomura, Suzuki, et al. 2006; D S Syrdal et al. 2009). The scale is composed of 14 items which can be subdivided into three categories (see Table 12.1): 1) **S1** negative attitude to situations of interaction with robots (six items), 2) **S2** negative attitude towards social influence of robots (five items), and 3) **S3** negative attitude about one's own emotions when interaction with robots (three items). Items are rated by participants as a 5-points Likert-scale scale where 1–2 ratings indicate more positive attitudes towards robots, 3 rating indicates neutral attitudes towards robots, and 4–5 ratings indicate greater negative attitudes towards robots.
To measure the internal consistency (or reliability) of the subscales, I performed a Cronbach's Alpha analysis on participant's responses. The analyses were ran on the responses to the NARS questionnaire at the beginning and end of the interaction trials, by removing the same items and reversing some other to maintain consistency of the subscales.
Considering subcategory S1 of pre-experimental questionnaire, since Pearson correlation coefficient of item the "I would feel nervous operating a robot in front of other people" was negative, I decided to remove it to maintain the reliability of the subscale. The resultant subscale had a high level of internal consistency

Table 12.1: Negative Attitudes towards Robots Items with Subscales

Item No.	Questionnaire Item	Sub-Scale
1	I would feel uneasy if robots really had emotions.	S2
2	Something bad might happen if robots developed into living beings.	S2
3	I would feel relaxed talking with robots.*	S3
4	I would feel uneasy if I was given a job where I had to use robots.	S1
5	If robots had emotions I would be able to make friends with them.*	S3
6	I feel comforted being with robots that have emotions.*	S3
7	The word "robot" means nothing to me.	S1
8	I would feel nervous operating a robot in front of other people.	S1
9	I would hate the idea that robots or artificial intelligences were making judgements about things.	S1
10	I would feel very nervous just standing in front of a robot.	S1
11	I feel that if I depend on robots too much, something bad might happen.	S2
12	I would feel paranoid talking with a robot.	S1
13	I am concerned that robots would be a bad influence on children.	S2
14	I feel that int the future society will be dominated by robots.	S2
(*inverse item)		

with a Cronbach's $\alpha = 0.82$. While running the Cronbach's Alpha analysis on subscale S2, for maintaining the consistency I decided to remove elements "I am concerned that robots would be a bad influence on children." and "I feel that in the future society will be dominated by robots.". The Cronbach's Alpha measure was $\alpha = 0.748$. Finally, I observed that the internal consistency of the subscale S3 had $\alpha = 5.25$.

The Cronbach's Alpha values of the subscales obtained from the analysis of the post-experimental questionnaire are respectively: S1 - $\alpha = 0.674$, S2 - $\alpha = 0.571$, and S3 - $\alpha = 0.674$.

The overall NARS categories had no statistically significant relationship comparing participants' perception of the robot before and after the interaction trials.

12.6.4 Post-experimental evaluation

As part of post-experimental questionnaire, people were asked to rate their perception of the interaction and the robot.

Perception of the interaction

A 7-point Likert Scale, ranged from 1 to 7 (disagree to agree), was used to measure the participants' judgement of the realism of the scenarios. All participants but one rated the scenarios very realistic (ratings higher than 5).
At the end of the final questionnaire I also asked participants if they believed the robot was behaving autonomously. The fifty percentage of participants believed that Care-O-bot 4 is not autonomous (ratings less than 3), 16.7% rated the robot as autonomous (ratings higher than 6) and 33.3% were uncertain (ratings equals to 4).

Perceptions of Care-O-bot 4

As for a generic robot, I asked participants to rate their willingness of having Care-O-bot 4 or another robot as home companion, and the roles suitable for the Care-O-bot 4.
I observed that majority of people (4 out 6 participants) were really happy about the idea of having Care-O-bot 4 in their homes, one participant were not inclined to have Care-O-bot 4 as home companion, and one participant did not state to neither want or not want Care-O-bot 4 as home companion. On the contrary, the majority of participants were less enthusiastic (2 out 6 participants) or were not inclined at all of having another robot as home companion (2 out 6 participants).
The majority of participants chose as roles perceived as suitable for robots: 1) assistant (2 out 6 participants), 2) tool (1 out 6 participants), 3) companion (1 out 6 participants, friend (1 out 6 participants) and butler (1 out 6 participants).
I also measured people's perception of robot's reliability and faith to perform correctly in untried situations using a 7-point Likert Scale [1= disagree strongly and 7= agree strongly] (Madsen et al. 2000).
I observed that the participants' perceived reliability of the robot was lower for participants who chose to trust Care-O-bot 4 ($n = 4.17 \pm 0.24$) when tested with condition **C2** than for participants who preferred to call the fire brigade ($n = 6.17 \pm 0.24$) when tested with condition **C1**.

12.6.5 Errors evaluation

In the last part of the questionnaire the robot were asked the following questions:

- Did I make some errors? (Answers: yes or no)
- Which one do you think is an error with big consequences? (select all that apply)
- Which one do you think is an error with small consequences? (select all that apply)
- If the emergency fire were in the kitchen, would you trust me to deal with it? (Answers: yes or no)
- If your beloved ones were in life-danger, would you trust me to deal with it? (Answers: yes or no)
- If you had a pet, would you trust me to bring it for a walk? (Answers: yes or no)
- If you needed to take medicines regularly, would you trust me to remind you of taking them? (Answers: yes or no)
- If you had important appointments in your calendar, would you trust me to remind you about them? (Answers: yes or no)
- If there were a dirty spot on your carpet, would you trust me to clean it? (Answers: yes or no)
- If you lived in a smart house with several intelligent devices, would you trust me to control them? For example, turn on and off the heaters or the oven, or switch on and off lights, or to warn you if you forgot the fridge door opened, or to alert you if the someone ringed the door bell? (Answers: yes or no)

Four out of six participants stated that the robot made errors during the interaction, but, interestingly, two out six participants did not think the robot made errors. These latter two participants were tested one in each condition.
Table 12.2 shows the distributions of people's responses by rating the errors obtained from the questionnaire responses. The resulting rankings highlighted three big errors and four small errors. These scenarios also included the big errors used

Table 12.2: Participants' ratings of big and small errors

Scenario	Big Errors	Small Errors
Q1 You share some private information about yourself with the robot. Your robot reveals it to a visitor.	83.3%	0
Q2 You ask your robot to charge your phone. Your robot puts it in the toaster.	83.3%	0
Q3 You asked your robot to show you the latest news. Your robot shows it on his own screen that faces away from you.	16.7%	100%
Q4 Your robot cooked some biscuits in the oven. It forgot to switch the oven off.	66.7%	16.7%
Q5 You are sitting on the right side of a table, your robot puts your drink on the opposite side.	0	83.3%
Q6 You and your robot are solving a puzzle. You ask your robot to take a piece useful to solve the puzzle. Your robot brings you the wrong piece.	0	83.3%
Q7 Your robot leaves your pet hamster outside the house in very cold weather.You are sitting on the sofa.	83.3%	0
Q8 You ask for a cup of coffee. Your robot brings you an orange.	0	83.3%

in the interactions (questions **Q1**, **Q2** and **Q4**), participants rated them as errors with severe consequences.
Finally, I investigated whether participants would have trusted Care-O-bot 4 in scenarios different from the fire emergency. They did not trust the robot with life threatening scenarios, i.e.: 1) they did not trust the robot to deal with an emergency scenario in a place different of garage, 2) they did not trust the robot to look after the well-being of their beloved ones, and 3) they did not trust the robot to take their pet for a walk. In contrast, participants trusted Care-O-bot 4 to be able to handle cognitive and lower risks situations (see Table 12.3). In particular, they unanimously stated to trust the robot to be able to remind them to take medicines, important meetings, and to manage a smart house like Robot House.

Table 12.3: Participants' decision of trusting Care-O-bot 4 in several scenarios.

Scenario	**Trust's Choice**
Q1 If the emergency fire were in the kitchen, would you trust me to deal with it?	16.7%
Q2 If your beloved ones were in life-danger, would you trust me to deal with it?	16.7%
Q3 If you had a pet, would you trust me to bring it for a walk?	33.5%
Q4 If you needed to take medicines regularly, would you trust me to remind you of taking them?	100%
Q5 If you had important appointments in your calendar, would you trust me to remind you about them?	100%
Q6 If there were a dirty spot on your carpet, would you trust me to clean it?	83.3%
Q7 If you lived in a smart house with several intelligent devices, would you trust me to control them? For example, turn on and off the heaters or the oven, or switch on and off lights, or to warn you if you forgot the fridge door opened, or to alert you if the someone ringed the door bell?	100%

12.7 Discussion and Conclusion

This Chapter investigated research question "Can people's trust in a robot be recovered more easily if it is a big error happening at the beginning or end of a long-term interaction?" (**RQ-4**). In this study a Mojin Robotics Care-O-bot 4 engaged participants in several tasks over a period of time of three weeks. The robot made errors with severe consequences either at the beginning or at the end of its interactions with participants. At the end of the six interactions days people were tested by observing their (not)trust's choices in the ability of the robot to handle a fire emergency in the house.

The majority of participants were female, and they did not have any previous experience with robots. People believed that the scenarios were very realistic, but most of them did not believe in, or they were uncertain about, the robot's autonomy. This result is particularly interesting because I believe that participants did not think that the robot was tele-operated, but they judged its capabilities of handling tasks

autonomously (Choi et al. 2014). Indeed, while participants were full immersed in the scenarios, including the fire emergency, they seemed to not expect help from the robot. They principally perceived the robot as an assistant and they believed the robot had limited abilities.
Moreover, this study showed that people were affected more by robot errors made at the beginning of the interaction than at the end of the interaction. On one hand, this result confirmed my findings in Chapter 5, and they are also supported by other studies where people's impressions of an agent is formed during the first encounters with the robot (Wood 2014). On another hand, a further investigation on the reasons causing the lack of people's trust in the robot was also due to the context, and in particular, the criticality of the emergency task. These findings corroborated the evidence that the reliability in robots' capabilities was affected by the risk of a possible negative outcome (J. Lee et al. 1992; Mayer et al. 1995).
Not every participants acknowledged that the robot made errors. However, they all judged the tasks in the interaction as errors with severe consequences. I believe that participants stated that the robot did not do any errors because they were asked directly by the robot. Indeed, some studies in psychology and computer-human interaction showed that people could feel inhibited to communicate their disapproval to the bystander culprit (i.e. the robot) (Chekroun et al. 2002; Voelpel et al. 2008).
My findings would need to be consolidated by conducting larger scaled trials (e.g. larger sample size, different robot embodiments, a more general population etc.). However, I observed that participants did not trust the robot's capabilities for handling the emergency scenario or high risk tasks.

In the next Chapter 13 I summarise the findings of this research, its limitations and future works and applications.

Chapter 13

Discussion & Conclusions

> Let the future tell the truth, and evaluate each one according to his work and accomplishments. The present is theirs; the future, for which I have really worked, is mine.
>
> *Nikola Tesla*

In the recent years, researchers from multidisciplinary fields (e.g. AI, Psychology, Ethics and Law) have been investigating the role of trust in HRI to effectively support the design and development of socially acceptable and trustful service robots. In this context, I observed that a deeper exploration of the dynamics of trust between humans and robots was needed. In particular, I found that trust can be affected by several factors that can depend on individuals' differences, robots' limitations, people's social expectations and expectation related to a robot's capabilities, and the nature of interaction between people and robots (i.e. type and length). A literature review of social robotics indicated that people's trust is affected by the performances and capabilities of a robot, but also by the antecedents of trust. Therefore, I hypothesised that people can be affected differently by a robot's errors with different severity of consequences and happening in different stages of an interaction (i.e. the beginning or end of an interaction). I also hypothesised that people's trust in a robot is formed differently according their expectations and knowledge of the robot, and also their personal predispositions and traits. Moreover, I hypothesised that a robot that shows social-awareness can recover humans' trust after it had undermined their trust.
In this Chapter I discussed, summarised and concluded my research and relative

findings allowing roboticists to move a step closer towards the development and deployment of companion robots that are able to engage and cooperate with people in effective and long-term interactions.

13.1 Summary of research

The following research questions emerged from the literature review while investigating the trust dynamics between humans and robots:

- RQ-1 How do various type of robot errors affect human's trust in a robot?
- RQ-2 Can people's trust in a robot change diversely according their personal differences?
- RQ-3 Does a robot that expresses social conventions gain more people's trust in it than a non-social robot?
- RQ-4 Does people's trust on a robot change over time if the initial conditions of trusting the robot change?

RQ-1 How do various type of robot errors affect human's trust in a robot?

The first two studies that I conducted were to explored how participants perceived a robot's errors with different magnitude of consequences, and how these errors differently affected their trust in the robot. The first study (see Chapter 4) allowed us to identify how participants rate the severity of the consequences of robot's errors. Fifty participants aged between 19 and 63 years old took part in the study. The results revealed that participants considered the robot's errors as small errors if they have limited consequences, and big errors if they have severe consequences.

The second study (see Chapter 5) was conducted as an interactive storyboard presenting ten different scenarios in which a robot completed tasks under five different conditions. The conditions included tasked executed correctly, and trivial and severe errors interspersed by flawless robot behaviours in different phases of the interaction (i.e. the beginning or end of interaction). At the end of each condition, the participants were presented with a final task in which a fire started in their kitchen. I observed their choices of trusting or not trusting the robot. I analysed the responses of 154 participants, aged between 18 and 65 years old. I observed that participants' trust was affected more severely when the robot made errors with severe consequences.

RQ-2 Does people's trust in a robot change diversely according their personal differences?

While analysing people's individual differences in the second study (see Chapters 6, 7), I found a correlation both between individual personalities and characteristics of people and their perception of the robot and trust towards a robot. I found a strong connection between the personality traits of agreeableness, conscientiousness and emotional stability, and their disposition of trust other people. Extroverts perceived the robot as a friend and a warm and attentive entity, while agreeable participants perceived the robot as a tool. I also found that extroverted participants would like to have the robot as home companion and believe it is reliable and trustworthy in uncertain and unusual situations. I found that conscientiousness and agreeableness traits correlate with participants' propensity for trusting the robot, and participants' belief in benevolence of people also correlate with higher trust in the robot.
The majority of the participants did not have any previous experience of interaction with robots. I observed that people with negative previous experience with robot were less inclined to trust a robot that made big errors, while a positive experience with a robot consequently affected people's positive predisposition towards a robot that behaved flawlessly.

Following these first successful studies, I investigated the effects of greater awareness of robotics systems on people's trust of them (see Chapter 8). I conducted a two-part study with pupils of local primary and secondary schools. Each child participated in three different types of activities: a video activity, a real live activity and programming activity of a robot. I observed changes of student's perception and trust of robots after each interaction. In particular, a higher familiarity with the robot affected participants positively regarding the robot's likeability of those who interacted with Pepper, while it tended to mitigate participants' overall negative perceptions of Kaspar. The study revealed that the trust of pupils that interacted with Pepper was positively affected across different domains after each session, indicating that human users trust a robot more if they more awareness about the robot. Contrary to my expectations, trust of pupils that interacted with Kaspar was not significantly affected across different domains after each session. It did not appear to be significant differences in trust tendencies for the different experiences either. However, my results suggest that human users trust a robot more with more awareness about the robot they have. However, I found that the different appearance of the robots only partially affected their trust in the robot they interacted with.

RQ-3 Does a robot that expresses social conventions gain more people's trust than a non-social robot?

The next step in my research has been to investigate how to enhance people's trust in robots when this has been lost due to unexpected faulty behaviours of a robot. Literature indicates that a robot aware of human social conventions increases people's acceptability and perceived safety while they are interacting with it (Vigliocco et al. 2014; S. Woods et al. 2006). I, therefore, decided to assess whether social behaviours influences people' trust in a robot in order to establish a safe human-robot interaction in human-centred environments.

The primary step has identified which social cues, previously highlighted in literature, had higher impact on people's perception of a robot as a social entity (see Chapter 9). One hundred participants were part of this study, aged 22 to 70. In this study, people from different background and nationalities watched eight videos in which a home companion robot engaged an actor in a simple fetch-and-carry scenario by following social behaviours that differ from each other by the omission of a single social cues. The findings lead us to conclude that social cues have a higher impact on people's perception of sociability in robots in the following order: 1) speech, 2) colour, 3) emotions, 4) body and head movements, 5) text and 6) proximity.

I then investigated the role of robots' appearances on people's perception of sociability of a robot by comparing the social cues of two different types of home companion robots (see Chapter 10). At the same time, I wanted to verify whether the obtained results were transferable to a robot with different appearances. Seventy three participants, aged 19 to 62, and of different nationalities, watched 6 couples of small video clips for each robot showing one of the cues identified in the previous part of this study. The findings showed that people's perception was partially affected by the robots' features. Robots were rated similarly in expressing social cues and behaviours that have a functional purpose, such as using vocal expressions, their tablet and navigating in the environment. The robot appearances were, conversely, perceived differently by the way they expressed their emotions.

In particular, the robot with mechanical-looking features (i.e. Care-O-bot 4 robot) was considered more as a butler and less as a friend than the other robot (i.e. Pepper robot), which has more human-like features.
Finally, I sought to understand whether a robot that meets people's social expectations, in terms of social interactions, habits and conventions, can improve people's trust in the robot (see Chapter 11). Twenty people (aged between 19 and 62) interacted either with a social or a non-social robot, engaging it in three tasks with increasing magnitude of big consequences (from low to high criticality and risk). The three tasks consisted in trusting the robot to play a song of its choice, the robot's suggestion of handling a package to a courier and signing a delivery-note, and that the robot is able to look after their well-being by suggesting them to take a plate from the microwave without gloves. In addition to the standard trust measures (i.e. questionnaires), I also observed the live interactions. The results from the observations showed that people trusted almost blindly the robot in the first two scenarios. Participants who interacted with the social robot did not trust the robot's indications in the scenario with a higher risk-taking outcome. This results of this study showed that the sociability of the robot had no significant effect on participants' trust in the robot. Moreover, this study outlined that participants' expectations of the robot's capabilities might have been affected by the robot's appearance and participants' familiarity with AI assistants.

RQ-4 Does people's trust in a robot change over time if the initial conditions of trusting the robot changes?

The final study in this research was directed to further investigate the effects of a robot's errors on people's trust in the robot overtime. Six people (aged between 24 and 47 years old) had repeated interactions over three weeks, twice per week, with a home companion robot. The study was designed to investigate if people would trust a robot that broke their trust in a initial or later stage of the interaction. From the first results in my study, and also corroborated by others such as Yu et al. study (2017), I confirmed that first impressions and judgements are principally formed at the beginning of the interaction.

13.2 Original contributions to knowledge

Current literature generally agrees that trust is a fundamental factor in establishing and maintaining effective relationships with assistive and service robots. Trust is being investigated in several disciplines, and there are different definitions of trust that link it to people's perception of reliability in the robot's functionalities (Mayer et al. 1995), to their willingness to take the risk of unbalanced positive outcome and negative consequences (Deutsch 1958), and the attitude that the robot will help them to achieve their goals in an uncertain and vulnerable situation (J. D. Lee et al. 2004). Trust can be also related to an affective connection between people and robots (Lewis et al. 1985; McAllister 1995).

Nevertheless, robots placed in human dynamic environments such as private homes, are likely to exhibit occasional behaviours perceived as unexpected or failures by people, or as actual errors. For example, robots could be affected by mechanical, programming or functional errors.

Robot errors have been shown to reduce the perceived reliability and trustworthiness of robots in several studies (Munjal Desai et al. 2013; Hancock et al. 2011; Maha Salem, Gabriella Lakatos, et al. 2015). These works highlighted that users complied with the robots' directions and suggestions discarding previous robotic failures (Bainbridge et al. 2011; Robinette, W. Li, et al. 2016; Maha Salem, Gabriella Lakatos, et al. 2015). These studies also were characterised by the fact that the robots' errors were not distinguished by different magnitude of consequences. Indeed, I have shown that errors with severe consequences affected people's trust in robots more than errors with small consequences (Chapters 4 and 5).

One contrasting study by Corritore et al. (2003) showed that a sequence of small errors can affect people's trust in robots more severely and for longer than one single big error. In my research, I also argued that the timing in which the errors occurs may impact differently people's trust in robots. In particular, considering big robot errors. Indeed, Chapters 5 and 12 showed that people were influenced more by big errors when these happened at the beginning of the interaction. This result is also in line with the work presented by Yu et al. (2017), in which they established that people's judgements are formed in the initial phase of a relationship, and it can be eventually adjusted later on depending on the following robot performances.

Inspired by these findings, I decided to pursue my belief that individuals' differences, including previous experiences and awareness of robots, may influence their perceptions of trust in robots, and consequently guide their reactions to the robots' errors. Honig et al. (2018) indicated that humans may adapt to robots if they are able to understand and predict their behaviours during interactions. In particular, they indicated that people's comprehension of robot errors are influenced by their background knowledge (Tannenbaum et al. 2006), personality (Sadeghi et al. 2012), expectations (Haberlandt 1982), and experience (Macias 2003). Another factor is people's situational awareness of the interaction environment (including robots, locations and other human agents), their awareness of the robot's ability for understanding and following human commands, their awareness of the robot's plans and goals, and their awareness of the state and stages of the cooperating task (J. L. Drury et al. 2003). However, there were no information regarding this. Hence, in this thesis, I have shown that people's perception of the robot and its errors consequently also affect their trust. Indeed, the findings showed that individuals' personality traits and personal dispositions, and previous experiences with robots influenced their trust in the robot, particularly, when the robot is making big errors (see Chapters 6 and 7). Additionally, I contribute to knowledge by showing (see Chapter 8) that also people's awareness of a robot's real capabilities and limitations can affect their trust in robots.

Honig et al. (Honig et al. 2018) also reported that users' perceptions during in interactions where robots behaved erroneously were related to anthropomorphism (M. K. Lee et al. 2010; Lemaignan et al. 2015; Sarkar et al. 2017), animacy (Engelhardt et al. 2017) and perceived intelligence (Mubin et al. 2015). Moreover, it is found to be fundamental that robots' capabilities need to be socially enhanced to ensure successful integration in people's social life (Alessandra Rossi, Garcia, Maya, et al. 2019; S. Rossi et al. 2020). In this thesis, I showed that not all social cues are perceived in the same way when expressed by robots with different appearances. In particular, a robot with a more machine-looking appearance is not able to communicate similarly to a more human-like robot (see Chapter 10). I also found that people have higher expectations in terms of capabilities and trust, and therefore, they are affected more when interacting with a social robot vs. interacting with a non social robot (see Chapters 9 and 11).

13.2.1 Final remarks on contribution to knowledge

In general terms, I can conclude that even if there are still several open challenges for social robots when directly or indirectly interacting with people (S. Rossi et al. 2020), the contribution to knowledge of the research presented in this thesis gives an essential contribution towards the effective deployment of companion and service robots in future domestic and working environments.

Moreover, the contribution to knowledge of this research has emphasised the multi-disciplinary nature of people's trust. Indeed, this research opens up future directions and new challenges to further investigate people's trust in robots in different research fields, including HRI, robotics companion and social robotics, ethics, laws and psychology. HRI and robotics fields will benefit by the insights gained by this research regarding robot's behaviours, and limitations of robot hardware and software. The studies presented in this thesis has highlighted a deeper understanding of ethical and legal consequences related to people's trust in a robot. Finally, the results of this research has indicated the need for further investigations into human psychology in relation to robots.

Appendix A includes a comprehensive lists of publications resulted from the research presented in this thesis.

13.3 Limitations

The study findings presented in this thesis have shown the various factors that can affect, enhance, undermine, and recover people's trust in robots. However, there are some limitations to the research presented in this thesis.

13.3.1 Population validity

Finding participants for in person studies is extremely difficult. In particular, when investigating long term effects and changes over time in HRIs. In my last two studies, I was able to consistently establish that robot behaviours have affected participants' trust in them. However, larger scale trials could consolidate these findings, and also provide further insights to unravel the complexity of trust dynamics between humans and robots.

13.3.2 Age range

The students that took part in the studies conducted during the Robotics Weeks event (Chapter 8) were aged 14-15 years old in the secondary school and 7-10 years old in the primary school. I compared them to investigate whether their trust in the robot was affected by the appearance of the robot they interacted with. The results of this investigations showed only a partial effect. Currently, it is not possible to ascertain whether the difference of age range played an important factor in these results. Therefore, future investigations should consider to replicate the studies by inverting participants of the presentation of the robots with the two groups of students interacted with.

13.3.3 Duration of the interactions

My research showed that novel and inexperienced individuals, in particular at younger ages, are driven particularly by enthusiasm, surprise and curiosity while interacting with robots. While there is a risk that people might over-trust robots in such interactions, in a later stage, these emotions tend to wear off (Fernaeus et al. 2010). When the novelty effect does not affect anymore people's perception of the robot, they might abandon the robot in an corner of the house or toss it away. My last study consisted of repeated interactions over three weeks, for a total of six interactions for each participant. This is not a relatively long period of time considering that I am expecting to have autonomous robots in our everyday activities in our homes. Future studies would need to investigate how trust can be established and nurtured between people and robots in long-lasting interactions.

13.3.4 Ethical and legal challenges

Investigating people's trust in robots can be a challenging task, in particular when considering that a real life-threatening scenario would lead to serious ethical and legal problems in getting such a study approved and run (Maha Salem and Kerstin Dautenhahn 2015). In the live studies conducted in this research, the robots engaged the participants in critical tasks. However, while the results showed relevant implications towards the main investigation of this research, participants were aware that the studies were under the surveillance of the experimenter (i.e. me). They were also aware of being participants in a experimental study. These factors, therefore, might have affected participants' choices. Indeed, in real dangerous scenarios participants might well have a very different mindset.

13.3.5 Robot appearance

I used two humanoid robots (Pepper and Care-O-bot 4) that had different dimensions and appearance, but similar functionalities, to investigate people's perceptions of robot sociability. The studies presented in this thesis showed that the robots have been perceived as social entities in a similar way. However, future research should be considered to compare people's perception of Pepper and Care-O-bot 4 sociability with a robot having a machine-like appearance, or simply with various different appearances.

13.3.6 Technological limitations

Trust is affected by the perception of reliability of a robot as shown in Mayers et al. (1995) and the current research 5. People's expectation of a robot's functional abilities and behaviours are higher when a robot has human-like physical features, as compared to a machine-like robot (Goetz et al. 2003; Robins, K. Dautenhahn, Te Boekhorst, et al. 2004). Moreover, people are becoming very familiar with virtual assistant AI technologies, such as Alexa[1]. In contrast, current state of the art robot technologies does not allow complex, natural and completely autonomous interactions. For this reason, it was necessary to use WoZ techniques to conduct the live studies of this research. For example, the speech recognition and dialogue systems of the robots used in the live studies for this research did not allow the robot to have natural dialogues. In these cases, robot's responses were either scripted or directly controlled. However, these resolutions created delays in the robots' verbal responses to people or did not allow the robot to be as responsive as expected by participants. Future investigations should use robots able to perform tasks autonomously and to engage people with natural interaction modalities.

13.3.7 Online investigations

In order to overcome the previously listed limitations, I decided to use Amazon Mechanical Turk to conduct some online studies. The online studies presented in this thesis were designed to ensure that participants interacted with fully functional robots in realistic scenarios. Moreover, the validity of participants' responses was guaranteed by the use of manipulation check questions to assess people attention to the studies. While crowd-sourcing allows one to collect quickly data from a large and diverse sample of participants, it also has several drawbacks. For example, effective quality control of the data requires extensive manual labour, which is not always possible (Joosse et al. 2015). Considering the sampling of participants respoonses, even if Amazon Mechanical Turk counts participants from 190 countries (Paolacci et al. 2014), the same diversity of participants is not always guaranteed. Indeed, participants in this thesis' online studies were principally from two or three countries.

Amazon Mechanical Turk does not the ability to directly develop complex studies built into the system. However, it is possible to post a link to an external website or source. At the end of the study, participants need to provide a code to ensure the completion of the study. I used this workaround for the study presented in Chapter

[1] Amazon Alexa https://www.amazon.com/

5. This solution however, required a high investment of labour time in preparing complex studies to investigate multiple human factors. As identified by Bartneck et al. (2015), using an external website could result in participants not completing the entire study. For this reason, it has been necessary to generate and store unique codes for the participants to use in order to demonstrate that they fully completed the study.

13.4 Future Works

This research has found that people's trust in people can be broken by a robot which made errors with severe of consequences happening in an initial stage of an interaction. My findings also indicated that people's previous experiences with robots, expectations of robots' capabilities and social-awareness also affects their trust in robots. The insights gained by this research have shown that it is possible to build a successful collaboration between people and robots based on trust. However, they have also opened up new directions for investigating trust in HRIs and also identified a number of future challenges to overcome.

The investigation carried out for this thesis outlined the necessity of further understanding how human-robot relationships are formed, and which robot factors, including familiarity, appearance and perception as social entity, that influence most people's trust in robots. Indeed, in the last study (see Chapter 12) I observed that participants were inhibited to communicate their disapproval to the robot for its errors. This happened most probably due to an effect well-known in psychology and human-computer interaction, called bystander effect or social inhibition of helping. It seems to have milder effects in online interactions (Chekroun et al. 2002). Future research should investigate to what extent people's mental models, including the perceived implications of task outcomes and consequences on their persona, inhibits their behaviours in the presence of robots.

Moreover, if robots were to be used in home environments, healthcare settings or as service robots in the real world, further studies would need to be conducted to see how people's trust varies in everyday interactions, and also when a robot can meet people's expectations of their functionalities. In particular, in this last eventuality, people's trust could be mitigated more successfully by a robot that also meets their social expectations. This goal could be met by developing reactive and predictive robot mental models that will allow robots to adapt to user's needs; in terms of task cooperation, preferences and habits. Moreover, a deeper understanding of the environment, people and tasks will allow robots to model legible and transparent behaviours aimed at recovering people's loss of trust due to robot errors.

The results of this research have found that people's previous experiences of robots, personalities traits, dispositions of trust humans, and awareness of robots' capabilities affects their trust in robots. However, people are now becoming surrounded by digital technologies, and it is difficult to match people's expectations of robots with their experience with more robust and advanced AIs, such as Alexa, Siri, Cortana or Google assistant. Further studies should aim to integrate modern learning techniques (i.e. convolutional neural network, deep learning etc.) in HRI studies in order to further investigate how people's perceptions of robots affects their trust in it. This will contribute to develop robots that are able to interact naturally with people.

Investigating people's perception of social cues expressed by Pepper and Care-O-bot 4, results showed that people rated the robot's methods of expressing emotions differently. In this study, a participant stated they were not sure what kind of gesture Care-O-bot 4 had made. The videos used for this study were recorded in a first person point of view. However, a live study might be useful to consolidate these findings, because people might be affected also by the dimensions of the robot or their perception of safety. Moreover, future research should consider to further investigate whether different robot appearances affects people's perceptions of the robot in terms of sociability, and their perception of trust in the robot.

Studies conducted through crowd sourcing services can collect participants' responses very fast. In fact, I recruited the participants needed for the online studies in this research within a day. This might imply that the percentage of diversity of participants might change depending on the working time of the users of the crowd sourcing services. Future research might consider to investigate whether collecting responses of small group of participants by publishing the recruitment in working time depending by the different timezone would guarantee a wider diversity of the sample.

Acknowledgments

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 642667 (Safety Enables Cooperation in Uncertain Robotic Environments - SECURE).

Bibliography

Adams, Barbara D et al. (2003). "Trust in automated systems". In: *Ministry of National Defence*.

Agresti, A. (2002). *Categorical data analysis*. 2nd. Chichester;New York; Wiley-Interscience. ISBN: 9780471458760;0471458767;

Alex, Annu Maria et al. (2017). "Android based intelligent robot for border security". In: *International Research Journal of Engineering and Technology* 4.4, pp. 2041–2043.

Alonso, Victoria and Paloma de la Puente (2018). "System Transparency in Shared Autonomy: A Mini Review". In: *Frontiers in Neurorobotics* 12. ISSN: 1662-5218. DOI: 10.3389/fnbot.2018.00083.

Ambady, Nalini, Frank J. Bernieri, and Jennifer A. Richeson (2000). "Toward a histology of social behavior: Judgmental accuracy from thin slices of the behavioral stream". In: *Advances in Experimental Social Psychology* 32, pp. 201–271. ISSN: 0065-2601. DOI: 10.1016/S0065-2601(00)80006-4.

Anthony, M. Evans and Revelle William (2008). "Survey and behavioral measurements of interpersonal trust". In: *Journal of Research in Personality* 42.6, pp. 1585–1593. ISSN: 0092-6566.

Apple, W., L. A. Streeter, and R. M Krauss (1979). "Effects of pitch and speech rate on personal attributions." In: *Journal of Personality and Social Psychology* 37.5, pp. 715–727. DOI: 10.1037/0022-3514.37.5.715.

Atkinson, D. J., W. J. Clancey, and M. H. Clark (2014). "Shared awareness, autonomy and trust in human-robot teamwork". In: *In Artificial Intelligence and Human-Computer Interaction: Papers from the 2014 AAAI Spring Symposium on*.

Bainbridge, W. A. et al. (2011). "The Benefits of Interactions with Physically Present Robots over Video–Displayed Agents". In: *International Journal of Social Robotics* 3.1, pp. 41–52.

Banerjee, Siddhartha, Andrew Silva, and Sonia Chernova (Oct. 2018). "Robot Classification of Human Interruptibility and a Study of Its Effects". In: *ACM Trans. Hum.-Robot Interact.* 7.2, 14:1–14:35. ISSN: 2573-9522.

Baraka, Kim, Ana Paiva, and Manuela Veloso (2016). "Expressive Lights for Revealing Mobile Service Robot State". In: *Robot 2015: Second Iberian Robotics Conference*. Cham: Springer International Publishing, pp. 107–119.

Barrick, M. R. et al. (1998). "Relating member ability and personality to work-team processes and team effectiveness". In: *Journal of Applied Psychology* 83.3, pp. 377–391. DOI: 10.1037/0021-9010.83.3.377.

Bartneck, Christoph et al. (Apr. 2015). "Comparing the Similarity of Responses Received from Studies in Amazon's Mechanical Turk to Studies Conducted Online and with Direct Recruitment". In: *PLOS ONE* 10.4, pp. 1–23. DOI: 10.1371/journal.pone.0121595.

Becker, Lawrence C. (1996). "Trust as Noncognitive Security about Motives". In: *Ethics* 107.1, pp. 43–61. ISSN: 00141704, 1539297X.

Belpaeme, Tony et al. (Oct. 2013). "Child-Robot Interaction: Perspectives and Challenges". In: pp. 452–459. ISBN: 9783319026749. DOI: 10.1007/978-3-319-02675-6_45.

Berinsky, Adam, Gregory Huber, and Gabriel Lenz (Nov. 2011). "Using Mechanical Turk as a Subject Recruitment Tool for Experimental Research". In: 20.

Bickmore, Timothy, Daniel Schulman, and Langxuan Yin (2010). "Maintaining engagement in long-term interventions with relational agents". In: *Appl. Artif. Intell.* 24.6, pp. 648–666. ISSN: 0883-9514. DOI: 10.1080/08839514.2010.492259.

Billings, D. (1995). "Computer Poker". In: *University of Alberta M.Sc. thesis*.

Bishop, Laura et al. (2019). "Social robots: The influence of human and robot characteristics on acceptance". In: *Paladyn, Journal of Behavioral Robotics* 10.1, pp. 346–358. DOI: doi:10.1515/pjbr-2019-0028.

Booth, Serena et al. (2017). "Piggybacking Robots: Human-Robot Overtrust in University Dormitory Security". In: *12th ACM/IEEE International Conference on Human-Robot Interaction*. ACM, pp. 426–434. ISBN: 9781450343367;1450343368.

Borenstein, J., A. R. Wagner, and A. Howard (Mar. 2018). "Overtrust of Pediatric Health-Care Robots: A Preliminary Survey of Parent Perspectives". In: *IEEE Robotics Automation Magazine* 25.1, pp. 46–54. ISSN: 1070-9932. DOI: 10.1109/MRA.2017.2778743.

Bottom, William P. et al. (2002). "When Talk Is Not Cheap: Substantive Penance and Expressions of Intent in Rebuilding Cooperation". In: *Organization Science* 13.5, pp. 497–513. DOI: 10.1287/orsc.13.5.497.7816.

Breazeal, Cynthia (2003). "Emotion and Sociable Humanoid Robots". In: *Int. J. Hum.-Comput. Stud.* 59.1-2, pp. 119–155. ISSN: 1071-5819. DOI: 10.1016/S1071-5819(03)00018-1.

Brooks, Alison Wood, Hengchen Dai, and Maurice E. Schweitzer (2014). "I'm Sorry About the Rain! Superfluous Apologies Demonstrate Empathic Concern and Increase Trust". In: *Social Psychological and Personality Science* 5.4, pp. 467–474. DOI: 10.1177/1948550613506122.

Buhrmester, M., T. Kwang, and S. Gosling (2011). "Amazon's mechanical turk: A new source of inexpensive, yet high-quality data?" In: *Perspectives on Psychological Science* 6, pp. 3–5.

Cameron, D. et al. (2015). "Framing Factors: The Importance of Context and the Individual in Understanding Trust in Human-Robot Interaction". In: *International Conference on Intelligent Robots and Systems*.

Carpinella, Colleen M. et al. (2017). "The Robotic Social Attributes Scale (RoSAS): Development and Validation". In: *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*. Vienna, Austria: Association for Computing Machinery, pp. 254–262. ISBN: 9781450343367. DOI: 10.1145/2909824.3020208.

Cassell, Justine et al. (2001). "Human Conversation as a System Framework: Designing Embodied Conversational Agents". In: *Embodied Conversational Agents*. Cambridge, MA, USA: MIT Press, pp. 29–63. ISBN: 0262032783.

Castro-González, Álvaro, Henny Admoni, and Brian Scassellati (2016). "Effects of form and motion on judgments of social robots' animacy, likability, trustworthiness and unpleasantness". In: *International Journal of Human-Computer Studies* 90, pp. 27–38. ISSN: 1071-5819. DOI: 10.1016/j.ijhcs.2016.02.004.

Ceha, Jessy et al. (2019). "Expression of Curiosity in Social Robots: Design, Perception, and Effects on Behaviour". In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. CHI 19. Glasgow, Scotland Uk: Association for Computing Machinery. ISBN: 9781450359702. DOI: 10.1145/3290605.3300636.

Chanseau, A. et al. (Aug. 2018). "Does the Appearance of a Robot Influence People's Perception of Task Criticality?" In: *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pp. 1057–1062. DOI: 10.1109/ROMAN.2018.8525663.

Chekroun, Peggy and Markus Brauer (2002). "The bystander effect and social control behavior: the effect of the presence of others on people's reactions to norm violations". In: *European Journal of Social Psychology* 32.6, pp. 853–867. DOI: 10.1002/ejsp.126.

Chen, Weihua, Tie Zhang, and Yanbiao Zou (2018). "Mobile robot path planning based on social interaction space in social environment". In: *International Journal of Advanced Robotic Systems* 15.3.

Choi, J. J., Y. Kim, and S. S. Kwak (2014). "The autonomy levels and the human intervention levels of robots: The impact of robot types in human-robot interaction". In: *The 23rd IEEE International Symposium on Robot and Human Interactive Communication*, pp. 1069–1074. DOI: 10.1109/ROMAN.2014.6926394.

Colquitt, J. A., B. A. Scott, and J. A. LePine (2007). "Trust, trustworthiness, and trust propensity: A meta-analytic test of their unique relationships with risk taking and job performance." In: *Journal of Applied Psychology* 92.4, pp. 909–927.

Corritore, C. L., B. Kracher, and S. Wiedenbeck (2003). "On-line trust: concepts, evolving themes, a model". In: *International Journal of Human - Computer Studies* 58.6, pp. 737–758.

Costa, A. C., R. A. Roe, and T. C. B. Taillieu (2001). "Trust implications for performance and effectiveness". In: *European Journal of Work and Organizational Psychology* 10.3, pp. 225–244.

Costa, S. et al. (2015). "Using a humanoid robot to elicit body awareness and appropriate physical interaction in children with autism". In: *International journal of Social Robotics* 7(2), pp. 265–278. DOI: DOI:https://doi.org/10.1007/s12369-014-0250-2.

Dahlbäck, N., A. Jönsson, and L. Ahrenberg (1993). "Wizard of Oz studies — why and how". In: *Knowledge-Based Systems* 6.4. Special Issue: Intelligent User Interfaces, pp. 258–266. ISSN: 0950-7051. DOI: 10.1016/0950-7051(93)90017-N.

Damiano, Luisa and Paul Dumouchel (2018). "Anthropomorphism in Human–Robot Co-evolution". In: *Frontiers in Psychology* 9, p. 468. ISSN: 1664-1078. DOI: 10.3389/fpsyg.2018.00468.

Daniela, Conti, Commodari E, and Buono S (Dec. 2017). "Personality factors and acceptability of socially assistive robotics in teachers with and without specialized training for children with disability". In: *Life Span and Disability* 20.2, pp. 251–272.

Dautenhahn, Kerstin (Aug. 2003). "Roles and Functions of Robots in Human Society: Implications from Research in Autism Therapy". In: *Robotica* 21.4, pp. 443–452. ISSN: 0263-5747. DOI: 10.1017/S0263574703004922.

— (2007a). "Methodology & themes of human-robot interaction: A growing research field". In: *International Journal of Advanced Robotic Systems* 4.1 SPEC. ISS. Pp. 103–108. DOI: 10.5772/5702.
— (2007b). "Socially intelligent robots: dimensions of human-robot interaction". In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 362.1480, pp. 679–704.
Dautenhahn, Kerstin, Bernard Ogden, and Tom Quick (Sept. 2002). "From Embodied to Socially Embedded Agents - Implications for Interaction-aware Robots". In: *Cogn. Syst. Res.* 3.3, pp. 397–428. ISSN: 1389-0417.
DeNeve, K. M. and H. Cooper (1998). "The happy personality: A meta-analysis of 137 personality traits and subjective well-being". In: *Psychological Bulletin* 124, pp. 197–229. DOI: doi:10.1037/0033-2909.124.2.197.
Desai, M. et al. (Mar. 2012). "Effects of changing reliability on trust of robot systems". In: *2012 7th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp. 73–80. DOI: 10.1145/2157689.2157702.
Desai, Munjal et al. (2013). "Impact of robot failures and feedback on real-time trust". In: *ACM/IEEE International Conference on Human-Robot Interaction*, pp. 251–258. ISBN: 9781467330558. DOI: 10.1109/HRI.2013.6483596.
Deutsch, M. (1958). "Trust and suspicion". In: *The Journal of Conflict Resolution* 2.4, pp. 265–279. DOI: 10.1177/002200275800200401.
Driskell, J. E. et al. (2006). "What makes a good team player? Personality and team effectiveness." In: *Group Dynamics: Theory, Research, and Practice* 10.4, pp. 249–271. DOI: 10.1037/1089-2699.10.4.249.
Drury, J. L., J. Scholtz, and H. A. Yanco (Oct. 2003). "Awareness in human-robot interactions". In: *Systems, Man and Cybernetics, 2003. IEEE International Conference on*. Vol. 1, pp. 912–918. DOI: 10.1109/ICSMC.2003.1243931.
Drury, Jill (2000). "Extending Usability Inspection Techniques for Collaborative Systems". In: *CHI '00 Extended Abstracts on Human Factors in Computing Systems*. CHI EA '00. The Hague, The Netherlands: ACM, pp. 81–82. ISBN: 1-58113-248-4. DOI: 10.1145/633292.633341.
Ekman, Paul (1999). *Basic Emotions*. T. Dalgleish and M. Power (Eds.)., pp. 45–60.
Elliot, Andrew J. and Markus A. Maier (2007). "Color and Psychological Functioning". In: *Current Directions in Psychological Science* 16.5, pp. 250–254. DOI: 10.1111/j.1467-8721.2007.00514.x.
Elson, J. S., Douglas C. Derrick, and Gina Scott Ligon (2018). "Examining Trust and Reliance in Collaborations between Humans and Automated Agents". In: *HICSS*.

Elson, Joel, Douglas Derrick, and Ginamarie Ligon (2020). "Trusting a Humanoid Robot : Exploring Personality and Trusting Effects in a Human-Robot Partnership". In: *Proceedings of the 53rd Hawaii International Conference on System Sciences*. ISBN: 978-0-9981331-3-3.

Endsley, Mica R. (1988). "Design and Evaluation for Situation Awareness Enhancement". In: *Proceedings of the Human Factors Society Annual Meeting* 32.2, pp. 97–101. DOI: 10.1177/154193128803200221.

Engelhardt, Sara and Emmeli Hansson (2017). "A comparison of three robot recovery strategies to minimize the negative impact of failure in social HRI". In: *Computer Science*.

Eyssel, F. et al. (2010). "Anthropomorphic inferences from emotional nonverbal cues: A case study". In: *19th International Symposium in Robot and Human Interactive Communication*, pp. 646–651.

Eyssel, Friederike and Frank Hegel (2012). "(s) he's got the look: Gender stereotyping of robots 1". In: *Journal of Applied Social Psychology* 42.9, pp. 2213–2230.

Fernaeus, Ylva et al. (2010). "How Do You Play with a Robotic Toy Animal? A Long-Term Study of Pleo". In: *Proceedings of the 9th International Conference on Interaction Design and Children*. Barcelona, Spain: Association for Computing Machinery, pp. 39–48. ISBN: 9781605589510. DOI: 10.1145/1810543.1810549.

Ficocelli, M., J. Terao, and G. Nejat (Dec. 2016). "Promoting Interactions Between Humans and Robots Using Robotic Emotional Behavior". In: *IEEE Transactions on Cybernetics* 46.12, pp. 2911–2923. ISSN: 2168-2267. DOI: 10.1109/TCYB.2015.2492999.

Fiore, Stephen M. et al. (2013). "Toward understanding social cues and signals in human–robot interaction: effects of robot gaze and proxemic behavior". In: *Front. Psychol.* Vol. 4:859. DOI: 10.3389/fpsyg.2013.00859.

Freedy, Amos et al. (2007). "Measurement of trust in human-robot collaboration". In: *Proceedings of the 2007 International Symposium on Collaborative Technologies and Systems, CTS*, pp. 106–114. ISBN: 0978569911. DOI: 10.1109/CTS.2007.4621745.

Friestad, Marian and Peter Wright (1994). "The Persuasion Knowledge Model: How People Cope with Persuasion Attempts". In: *Journal of Consumer Research* 21.1, pp. 1–31. ISSN: 00935301, 15375277.

Funk, Sebastian et al. (2009). "The spread of awareness and its impact on epidemic outbreaks". In: *Proceedings of the National Academy of Sciences* 106.16, pp. 6872–6877. ISSN: 0027-8424. DOI: 10.1073/pnas.0810762106.

Furlough, Caleb, Thomas Stokes, and Douglas J. Gillan (2019). “Attributing Blame to Robots: I. The Influence of Robot Autonomy”. In: *Human Factors* 0.0, pp. 1–11. DOI: 10.1177/0018720819880641.
Garbarino, Ellen and Michal Strahilevitz (2004). “Gender differences in the perceived risk of buying online and the effects of receiving a site recommendation”. In: *Journal of Business Research* 57.7. Marketing on the web - behavioral, strategy and practices and public policy, pp. 768–775. ISSN: 0148-2963. DOI: https://doi.org/10.1016/S0148-2963(02)00363-6.
Garreau, J. (May 2007). *Bots on the ground: In the field of battle (or even above it), robots are a soldier’s best friend.* URL: Retrieved%20from%20http://www.washingtonpost.com.
Gielniak, M. J. and A. L. Thomaz (July 2011). “Generating anticipation in robot motion”. In: *2011 RO-MAN*, pp. 449–454.
Gockley, Rachel and Maja J. Matariundefined (2006). “Encouraging Physical Therapy Compliance with a Hands-Off Mobile Robot”. In: *Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-Robot Interaction*. New York, NY, USA: Association for Computing Machinery, pp. 150–155. ISBN: 1595932941. DOI: 10.1145/1121241.1121268.
Goetz, J., S. Kiesler, and A. Powers (2003). “Matching robot appearance and behavior to tasks to improve human-robot cooperation”. In: *The 12th IEEE International Workshop on Robot and Human Interactive Communication, 2003. Proceedings. ROMAN 2003.* Pp. 55–60.
Golder, Scott A. and Judith Donath (2004). “Hiding and Revealing in Online Poker Games”. In: *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work*. CSCW ’04. Chicago, Illinois, USA: ACM, pp. 370–373. ISBN: 1-58113-810-5. DOI: 10.1145/1031607.1031668.
Goodrich, Michael A. and Dan R. Jr Olsen (2003). “Metrics for Evaluating Human-Robot Interactions”. In: *Faculty Publications*. URL: https://scholarsarchive.byu.edu/facpub/1052.
Gosling, S. D., P. J. Rentfrow, and Jr. Swann W. B. (2003). “A Very Brief Measure of the Big Five Personality Domains”. In: *Journal of Research in Personality* 37, pp. 504–528.
Gosling, Samuel D, Peter J Rentfrow, and William B Swann (2003). “A very brief measure of the Big-Five personality domains”. In: *Journal of Research in Personality* 37.6, pp. 504–528. ISSN: 0092-6566. DOI: https://doi.org/10.1016/S0092-6566(03)00046-1.
Graaf, Maartje M.A. de and Somaya Ben Allouch (2014). “Expectation Setting and Personality Attribution in HRI”. In: *Proceedings of the 2014 ACM/IEEE*

International Conference on Human-Robot Interaction. HRI '14. Bielefeld, Germany: Association for Computing Machinery, pp. 144–145. ISBN: 9781450326582. DOI: 10.1145/2559636.2559796.

Graaf, Maartje M.A. de, Somaya Ben Allouch, and Jan A.G.M. van Dijk (2016). "Long-term evaluation of a social robot in real homes". In: *Interaction Studies* 17.3, pp. 461–490.

Graaf, Maartje de, Somaya Ben Allouch, and Jan van Dijk (2017). "Why Do They Refuse to Use My Robot? Reasons for Non-Use Derived from a Long-Term Home Study". In: *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*. HRI '17. Vienna, Austria: Association for Computing Machinery, pp. 224–233. ISBN: 9781450343367. DOI: 10.1145/2909824.3020236.

Gundogdu, Didem et al. (2017). "Investigating the association between social interactions and personality states dynamics". In: *Royal Society open science* 4.9, p. 170194.

Gutwin, Carl, Gwen Stark, and Saul Greenberg (1995). "Support for Workspace Awareness in Educational Groupware". In: *The First International Conference on Computer Support for Collaborative Learning*. CSCL '95. Indiana Univ., Bloomington, Indiana, USA: L. Erlbaum Associates Inc., pp. 147–156. ISBN: 0-8058-2243-7. DOI: 10.3115/222020.222126.

H., Kim P. et al. (2004). "Removing the shadow of suspicion: the effects of apology versus denial for repairing competence- versus integrity-based trust violations." In: *The Journal of applied psychology* 89.1, pp. 104–118. DOI: 10.1037/0021-9010.89.1.104.

Haberlandt, Karl (1982). "Reader Expectations in Text Comprehension". In: *Language and Comprehension*. Ed. by Jean-François [Le Ny] and Walter Kintsch. Vol. 9. Advances in Psychology. North-Holland, pp. 239–249. DOI: 10.1016/S0166-4115(09)60055-8.

Hall, Edward Twitchell (1966). *The Hidden Dimension*. Vol. 609. Doubleday. ISBN: 0385084765, 9780385084765.

Hancock, Peter A. et al. (2011). "A Meta-Analysis of Factors Affecting Trust in Human-Robot Interaction". In: *Human Factors: The Journal of Human Factors and Ergonomics Society* 53.5, pp. 517–527.

Haring, K. S. et al. (May 2015). "Touching an Android robot: Would you do it and how?" In: *2015 International Conference on Control, Automation and Robotics*, pp. 8–13. DOI: 10.1109/ICCAR.2015.7165993.

Haring, Kerstin Sophie, Yoshio Matsumoto, and Katsumi Watanabe (2013). "How do people perceive and trust a lifelike robot". In: *Lecture Notes in Engineering and Computer Science* 1, pp. 425–430.

Häring, M., N. Bee, and E. André (2011). "Creation and Evaluation of emotion expression with body movement, sound and eye color for humanoid robots". In: *IEEE International Conference on Robot and Human Interactive Communication 2011 RO-MAN*, pp. 204–209.

Haselhuhn, Michael P., Maurice E. Schweitzer, and Alison M. Wood (2010). "How Implicit Beliefs Influence Trust Recovery". In: *Psychological Science* 5, pp. 645–648.

Hegel, F. et al. (2011). "Towards a typology of meaningful signals and cues in social robotics". In: *2011 RO-MAN*, pp. 72–78. DOI: 10.1109/ROMAN.2011.6005246.

Holler, Judith and Stephen C. Levinson (2019). "Multimodal Language Processing in Human Communication". In: *Trends in Cognitive Sciences* 23.8, pp. 639–652.

Honig, Shanee and Tal Oron-Gilad (2018). "Understanding and Resolving Failures in Human-Robot Interaction: Literature Review and Model Development". In: *Frontiers in Psychology* 9. ISSN: 1664-1078. DOI: 10.3389/fpsyg.2018.00861.

HT, Reis et al. (2011). "Familiarity does indeed promote attraction in live interaction." In: *Journal of Personality and Social Psychology* 101.3, pp. 557–570. DOI: 10.1037/a0022885.

Huijnen, C. A., M. A. Lexis, and L. P. de Witte (2016). "Matching robot KASPAR to autism spectrum disorder (ASD) therapy and educational goals". In: *International Journal of Social Robotics* 8(4), pp. 445–455. DOI: DOI:https://doi.org/10.1007/s12369-016-0369-4.

Jirout, Jamie and David Klahr (2012). "Children's scientific curiosity: In search of an operational definition of an elusive concept". In: *Developmental Review* 32.2, pp. 125–160. ISSN: 0273-2297. DOI: 10.1016/j.dr.2012.04.002.

Joosse, M. P., M. Lohse, and V. Evers (2015). "Crowdsourcing Culture in HRI: Lessons Learned From Quantitative and Qualitative Data Collections". In: *ICSR'15 Workshop on Culture Aware Robotics*.

K., Dautenhahn et al. (2005). "What is a robot companion - friend, assistant or butler?" In: *Procs IEEE/RSJ Int Conf on Intelligent Robots and Systems 2005*, pp. 1488–1493.

Kanda, T. et al. (2007). "A Two-Month Field Trial in an Elementary School for Long-Term Human–Robot Interaction". In: *IEEE Transactions on Robotics* 23.5, pp. 962–971.

Kaplan, Alexandra D., Tracy Sanders, and Peter A. Hancock (2019). "The Relationship Between Extroversion and the Tendency to Anthropomorphize Robots: A Bayesian Analysis". In: *Frontiers in Robotics and AI* 5, p. 135. ISSN: 2296-9144. DOI: 10.3389/frobt.2018.00135.

Kaya, Nazmiye and Helen H. Epps (2004). "Relationship between Color and Emotion: A Study of College Students." In: *College Student Journal* 38.3, pp. 396–405.

Koay, K. L., G. Lakatos, et al. (2013). "Hey! There is someone at your door. A hearing robot using visual communication signals of hearing dogs to communicate intent". In: *2013 IEEE Symposium on Artificial Life (ALife)*, pp. 90–97. DOI: 10.1109/ALIFE.2013.6602436.

Koay, K. L., D. S. Syrdal, et al. (Feb. 2009). "Five Weeks in the Robot House: Exploratory Human-Robot Interaction Trials in a Domestic Setting". In: *2009 Second International Conferences on Advances in Computer-Human Interactions*, pp. 219–226. DOI: 10.1109/ACHI.2009.62.

Koay, K. L., Dag Sverre Syrdal, et al. (2007). "Living with robots: Investigating the habituation effect in participants' preferences during a longitudinal human-robot interaction study". In: *Proceedings - IEEE International Workshop on Robot and Human Interactive Communication*, pp. 564–569. ISBN: 1424416345. DOI: 10.1109/ROMAN.2007.4415149.

Koay, Kheng Lee, Dag Sverre Syrdal, Kerstin Dautenhahn, et al. (2020). "A narrative approach to human-robot interaction prototyping for companion robots". In: *Paladyn, Journal of Behavioral Robotics* 11.1, pp. 66–85.

Koay, Kheng Lee, Dag Sverre Syrdal, Mohammadreza Asghari Oskoei, et al. (2014). "Social Roles and Baseline Proxemic Preferences for a Domestic Service Robot". In: *International Journal of Social Robotics* 6.4, pp. 469–488. DOI: 10.1007/s12369-014-0232-4.

Koay, Kheng Lee, Dag Syrdal, et al. (2017). "Initial Design, Implementation and Technical Evaluation of a Context-aware Proxemics Planner for a Social Robot". In: *Social Robotics*. Ed. by Abderrahmane Kheddar et al. Springer International Publishing, pp. 12–22. ISBN: 978-3-319-70022-9.

Kose-Bagci, Hatice et al. (2009). "Effects of Embodiment and Gestures on Social Interaction in Drumming Games with a Humanoid Robot". In: *Advanced Robotics* 23.14, pp. 1951–1996. DOI: 10.1163/016918609X12518783330360.

Kramer, R. M. and P. J. Carnevale (2003). "Trust and Intergroup Negotiation". In: *Blackwell Handbook of Social Psychology: Intergroup Processes (eds R. Brown and S. L. Gaertner)*. DOI: 10.1002/9780470693421.ch21.

Kriz, Sarah et al. (2010). "Fictional robots as a data source in HRI research: Exploring the link between science fiction and interactional expectations". In: *RO-MAN, 2010 IEEE*. IEEE, pp. 458–463.

Kruse, Thibault et al. (2013). "Human-aware robot navigation: A survey". In: *Robotics and Autonomous Systems* 61.12, pp. 1726–1743.

Kuchenbrandt, D. and F. Eyssel (Sept. 2012). "The mental simulation of a human-robot interaction: Positive effects on attitudes and anxiety toward robots". In: *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*, pp. 463–468. DOI: 10.1109/ROMAN.2012.6343795.

Kuderer, Markus et al. (2012). "Feature-Based Prediction of Trajectories for Socially Compliant Navigation". In: *Robotics: Science and Systems VIII, University of Sydney, Sydney, NSW, Australia, July 9-13, 2012*.

Kudryavtsev, A. and J. Pavlodsky (2012). "Description-based and experience-based decisions: Individual analysis." In: *Judgment and Decision Making* 7.3, pp. 316–331.

Kurt, Sevinc and Kelechi Kingsley Osueke (2014). "The Effects of Color on the Moods of College Students". In: *SAGE Open* 4.1, p. 2158244014525423. DOI: 10.1177/2158244014525423.

Law, E. et al. (Aug. 2017). "A Wizard-of-Oz study of curiosity in human-robot interaction". In: *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pp. 607–614. DOI: 10.1109/ROMAN.2017.8172365.

Lee, Angela Y. (2001). "The Mere Exposure Effect: An Uncertainty Reduction Explanation Revisited". In: *Personality and Social Psychology Bulletin* 27.10, pp. 1255–1266. DOI: 10.1177/01461672012710002.

Lee, J. D. and K. A. See (2004). "Trust in Automation: Designing for Appropriate Reliance". In: *Human Factors: The Journal of the Human Factors and Ergonomics Society* 46.1, pp. 50–80. DOI: 10.1518/hfes.46.1.50_30392.

Lee, J. and N. Moray (1992). "Trust, control strategies and allocation of function in human-machine systems." In: *Ergonomics* 35.10, pp. 1243–1270. DOI: 10.1080/00140139208967392.

Lee, Kwan Min et al. (Nov. 2006). "Can Robots Manifest Personality?: An Empirical Test of Personality Recognition, Social Responses, and Social Presence in

Human–Robot Interaction". In: *Journal of Communication* 56.4, pp. 754–772. ISSN: 0021-9916. DOI: 10.1111/j.1460-2466.2006.00318.x.

Lee, M. K. et al. (2010). "Gracefully mitigating breakdowns in robotic services". In: *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp. 203–210.

Lee, Min Kyung et al. (2012). "Personalization in HRI: A Longitudinal Field Experiment". In: *Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction*. Boston, Massachusetts, USA: Association for Computing Machinery, pp. 319–326. ISBN: 9781450310635. DOI: 10.1145/2157689.2157804.

Lemaignan, Séverin et al. (2015). "You're Doing It Wrong! Studying Unexpected Behaviors in Child-Robot Interaction". In: *Social Robotics*. Ed. by Adriana Tapus et al. Cham: Springer International Publishing, pp. 390–400. ISBN: 978-3-319-25554-5.

Lewis, J. D. and A. Weigert (1985). "Trust as a social reality." In: *Social Forces* 63.4, pp. 967–985. DOI: 10.2307/2578601.

Li, D., P.L.P. Rau, and Y. Li (2010). "A Cross-cultural Study: Effect of Robot Appearance and Task." In: *International Journal of Social Robotics* 2, pp. 175–186. DOI: 10.1007/s12369-010-0056-9.

Li, Ning et al. (2014). "Retaining the Productive Employee: The Role of Personality". In: *Academy of Management Annals* 8.1, pp. 347–395. DOI: 10.5465/19416520.2014.890368.

Liang, Amy et al. (2017). "A Pilot Randomized Trial of a Companion Robot for People With Dementia Living in the Community". In: *Journal of the American Medical Directors Association* 18.10, pp. 871–878. ISSN: 1525-8610. DOI: 10.1016/j.jamda.2017.05.019.

Lichtenthaler, C., T. Lorenzy, and A. Kirsch (Sept. 2012). "Influence of legibility on perceived safety in a virtual human-robot path crossing task". In: *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*, pp. 676–681.

Ligthart, M. and K. P. Truong (Aug. 2015). "Selecting the right robot: Influence of user attitude, robot sociability and embodiment on user preferences". In: *2015 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pp. 682–687. DOI: 10.1109/ROMAN.2015.7333598.

Ljungblad, Sara et al. (2012). "Hospital Robot at Work: Something Alien or an Intelligent Colleague?" In: *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*. CSCW '12. Seattle, Washington, USA:

ACM, pp. 177–186. ISBN: 978-1-4503-1086-4. DOI: 10.1145/2145204.2145233.
Lohse, M. et al. (Aug. 2008). "Evaluating extrovert and introvert behaviour of a domestic robot — a video study". In: *RO-MAN 2008 - The 17th IEEE International Symposium on Robot and Human Interactive Communication*, pp. 488–493. DOI: 10.1109/ROMAN.2008.4600714.
Lohse, Manja (2009). "The role of expectations in HRI". In: *New Frontiers in Human-Robot Interaction*.
Lugrin, B., J. Dippold, and K. Bergmann (Oct. 2018). "Social Robots as a Means of Integration? an Explorative Acceptance Study considering Gender and Non-verbal Behaviour". In: *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 2026–2032. DOI: 10.1109/IROS.2018.8593818.
Macias, Wendy (2003). "A Beginning Look at the Effects of Interactivity, Product Involvement and Web Experience on Comprehension: Brand Web Sites as Interactive Advertising". In: *Journal of Current Issues & Research in Advertising* 25.2, pp. 31–44. DOI: 10.1080/10641734.2003.10505147.
Madsen, Maria and Shirley Gregor (2000). "Measuring human-computer trust". In: *Proceedings of the 11 th Australasian Conference on Information Systems*, pp. 6–8.
Maner, Jon K. and Mary A. Gerend (2007). "Motivationally selective risk judgments: Do fear and curiosity boost the boons or the banes?" In: *Organizational Behavior and Human Decision Processes* 103.2, pp. 256–267. ISSN: 0749-5978. DOI: https://doi.org/10.1016/j.obhdp.2006.08.002.
Maris, Anouk van et al. (2017). "The Influence of a Robot's Embodiment on Trust: A Longitudinal Study". In: *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*. Vienna, Austria: Association for Computing Machinery, pp. 313–314. ISBN: 9781450348850. DOI: 10.1145/3029798.3038435.
Martelaro, N. et al. (Mar. 2016). "Tell me more designing HRI to encourage more trust, disclosure, and companionship". In: *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp. 181–188. DOI: 10.1109/HRI.2016.7451750.
Matthews, Gerald et al. (2020). "Evolution and revolution: Personality research for the coming world of robots, artificial intelligence, and autonomous systems". In: *Personality and Individual Differences*, p. 109969. ISSN: 0191-8869. DOI: https://doi.org/10.1016/j.paid.2020.109969. URL: http://www.sciencedirect.com/science/article/pii/S0191886920301586.

Mayer, R. C., J. H. Davis, and F. D. Schoorman (1995). "An integrative model of organizational trust." In: *Academy of Management Review*, pp. 709–734.

McAllister, Daniel J. (1995). "Affect- and Cognition-Based Trust as Foundations for Interpersonal Cooperation in Organizations". In: *Academy of Management Journal* 38.1, pp. 24–59. DOI: 10.5465/256727.

McKnight, D. H., V. Choudhury, and C. Kacmar (2001). "Developing and Validating Trust Measures for e-Commerce: An Integrative Typology". In: *Information Systems Research* 13.3, pp. 334–359.

Mikawa, M., Y. Yoshikawa, and M. Fujisawa (Mar. 2018). "Expression of intention by rotational head movements for teleoperated mobile robot". In: *2018 IEEE 15th International Workshop on Advanced Motion Control (AMC)*, pp. 249–254.

Montemerlo, M. et al. (2002). "Experiences with a mobile robotic guide for the elderly". In: *AAAI National Conference on Artificial Intelligence*.

Mooradian, T., B. Renzl, and K. Matzler (2006). "Who Trusts? Personality, Trust and Knowledge Sharing". In: *Management Learning* 37.4, pp. 523–540. DOI: 10.1177/1350507606073424.

Mori, M., K. F. MacDorman, and N. Kageki (2012). "The Uncanny Valley [From the Field]". In: *IEEE Robotics Automation Magazine* 19.2, pp. 98–100.

Moros, Silvia et al. (Apr. 2019). "Programming a Humanoid Robot with the Scratch Language". In: *10th International Conference on Robotics in Education, RiE 2019*. Vienna.

Mubin, Omar and Christoph Bartneck (2015). "Do As I Say: Exploring Human Response to a Predictable and Unpredictable Robot". In: *Proceedings of the 2015 British HCI Conference*. British HCI '15. Lincoln, Lincolnshire, United Kingdom: ACM, pp. 110–116. ISBN: 978-1-4503-3643-7. DOI: 10.1145/2783446.2783582. URL: http://doi.acm.org/10.1145/2783446.2783582.

Muir, B. M. and N. Moray (1996). "Trust in Automation: Part II. Experimental studies of trust and human intervention in a process control simulation." In: *Ergonomics* 39, pp. 429–460. DOI: 10.1080/00140139608964474.

Mumm, Jonathan and Bilge Mutlu (2011). "Human-robot proxemics: physical and psychological distancing in human-robot interaction". In: *Proceedings of the 6th international conference on Human-robot interaction*. ACM, pp. 331–338.

Nomura, Tatsuya, Takayuki Kanda, and Tomohiro Suzuki (Feb. 2006). "Experimental Investigation into Influence of Negative Attitudes toward Robots on Human–Robot Interaction". In: *AI Soc.* 20.2, pp. 138–150. ISSN: 0951-5666. DOI: 10.1007/s00146-005-0012-7.

Nomura, Tatsuya, Tomohiro Suzuki, et al. (2006). "Measurement of negative attitudes toward robots". In: *Interaction Studies* 7.3, pp. 437–454. ISSN: 1572-0373. DOI: 10.1075/is.7.3.14nom.

Ogreten, S., S. Lackey, and D. Nicholson (2010). "Recommended roles for uninhabited team members within mixed-initiative combat teams". In: *2010 International Symposium on Collaborative Technologies and Systems*, pp. 531–536.

Paetzel, Maike, Giulia Perugia, and Ginevra Castellano (2020). "The Persistence of First Impressions: The Effect of Repeated Interactions on the Perception of a Social Robot". In: *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*. HRI 20. Cambridge, United Kingdom: Association for Computing Machinery, pp. 73–82. ISBN: 9781450367462. DOI: 10.1145/3319502.3374786.

Pandey, A. and R. Gelin (2018). "A Mass-Produced Sociable Humanoid Robot: Pepper: The First Machine of Its Kind". In: *IEEE Robotics Automation Magazine*, pp. 1–1. ISSN: 1070-9932.

Paolacci, Gabriele and Jesse Chandler (2014). "Inside the Turk: Understanding Mechanical Turk as a Participant Pool". In: *Current Directions in Psychological Science* 23.3, pp. 184–188. DOI: 10.1177/0963721414531598.

Patrick, Lin, Keith Abney, and Ryan Jenkins (2017). *Robot Ethics 2.0: From Autonomous Cars to Artificial Intelligence.* Oxford Scholarship Online. DOI: 10.1093/oso/9780190652951.001.0001.

Penghou, Liu, Chen Haichao, and Du Yanzhe (Apr. 2019). "Design of Multifunctional Intelligent Security Robot Based on Single Chip Microcomputer". In: *Journal of Physics: Conference Series* 1187.3, p. 032017. DOI: 10.1088/1742-6596/1187/3/032017.

Pintado, Teresa et al. (2017). "The Effects of Digital Media Advertising Content on Message Acceptance or Rejection: Brand Trust as a Moderating Factor". In: *Journal of Internet Commerce* 16.4, pp. 364–384. DOI: 10.1080/15332861.2017.1396079.

Pollack, M. E. (2005). "Intelligent technology for an aging population: The use of AI to assist elders with cognitive impairment". In: *AI Magazine* 26, pp. 9–24.

Posey, Clay et al. (2010). "Proposing the Online Community Self-Disclosure Model: The Case of Working Professionals in France and the UK Who Use Online Communities". In: *European Journal of Information Systems (EJIS)* 19.2, pp. 181–195. DOI: 10.1057/ejis.2010.15.

Rae, Irene, Leila Takayama, and Bilge Mutlu (2013). "In-Body Experiences: Embodiment, Control, and Trust in Robot-Mediated Communication". In: *Pro-*

ceedings of the SIGCHI Conference on Human Factors in Computing Systems. Paris, France: Association for Computing Machinery, pp. 1921–1930. ISBN: 9781450318990. DOI: 10.1145/2470654.2466253.

Raja, Parasuraman and Riley Victor (1997). "Humans and Automation: Use, Misuse, Disuse, Abuse". In: *Human Factors* 39.2, pp. 230–253. DOI: 10.1518/001872097778543886.

Ramírez, O. A. I. et al. (2016). "Robots learning how and where to approach people". In: *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pp. 347–353. DOI: 10.1109/ROMAN.2016.7745154.

Reber, Rolf, Piotr Winkielman, and Norbert Schwarz (1998). "Effects of Perceptual Fluency on Affective Judgments". In: *Psychological Science* 9.1, pp. 45–48. DOI: 10.1111/1467-9280.00008.

Reiser, Ulrich et al. (2013). "Care-O-bot®3 - Vision of a Robot Butler". In: *Your Virtual Butler: The Making-of*. Ed. by Robert Trappl, pp. 97–116. DOI: 10.1007/978-3-642-37346-6_9.

Reniers, Renate L. E. P. et al. (Apr. 2016). "Risk Perception and Risk-Taking Behaviour during Adolescence: The Influence of Personality and Gender". In: *PLOS ONE* 11.4, pp. 1–14. DOI: 10.1371/journal.pone.0153842.

Rhodes, Nancy and Kelly Pivik (2011). "Age and gender differences in risky driving: The roles of positive affect and risk perception". In: *Accident Analysis & Prevention* 43.3, pp. 923–931. ISSN: 0001-4575. DOI: https://doi.org/10.1016/j.aap.2010.11.015.

Rios-Martinez, J., A. Spalanzani, and C. Laugier (Apr. 2015). "From Proxemics Theory to Socially-Aware Navigation: A Survey". In: *International Journal of Social Robotics* 7.2, pp. 137–153.

Robert, L. P. (2018). "Personality in the Human Robot Interaction Literature: A Review and Brief Critique". In: *Proceedings of the 24th Americas Conference on Information Systems*. New Orleans, LA. URL: https://ssrn.com/abstract=3308191.

Robinette, Paul, Ayanna M. Howard, and Alan R. Wagner (2015). "Timing is key for robot trust repair". In: *Social Robotics. Lecture Notes in Computer Science* 9388, pp. 574–583.

Robinette, Paul, Wenchen Li, et al. (2016). "Overtrust of Robots in Emergency Evacuation Scenarios". In: *Proceeding HRI '16 The Eleventh ACM/IEEE International Conference on Human Robot Interation*. IEEE Press Piscataway, pp. 101–108. ISBN: 978-1-4673-8370-7.

Robins, B., K. Dautenhahn, and P. Dickerson (2009). “From Isolation to Communication: A Case Study Evaluation of Robot Assisted Play for Children with Autism with a Minimally Expressive Humanoid Robot”. In: *2009 Second International Conferences on Advances in Computer-Human Interactions*, pp. 205–211. DOI: doi:10.1109/ACHI.2009.32.

Robins, B., K. Dautenhahn, R. Te Boekhorst, et al. (2004). “Robots as assistive technology, does appearance matter”. In: *Procs of the 2004 IEEE Int Workshop on Robot and Human Interactive Communication*, pp. 277–282. ISBN: 0-7803-8570-5.

Roccas, Sonia et al. (2002). “The Big Five Personality Factors and Personal Values”. In: *Personality and Social Psychology Bulletin* 28.6, pp. 789–801. DOI: 10.1177/0146167202289008.

Ross, J. M. (2008). *Moderators of trust and reliance across multiple decision aids (Doctoral dissertation), University of Central Florida, Orlando.*

Rossi, A. et al. (2019). “Getting to know Kaspar : Effects of people’s awareness of a robot’s capabilities on their trust in the robot”. In: *2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, pp. 1–6.

Rossi, Alessandra, Kerstin Dautenhahn, Kheng Lee Koay, and Michael L. Walters (2017a). “A Study on How the Timing and Magnitude of Robot Errors May Influence People Trust of Robots in an Emergency Scenario”. In: *International Conference on Social Robotics (ICSR)*. Tsukuba, Japan: Springer.

— (2017b). “Human Perceptions of the Severity of Domestic Robot Errors”. In: *International Conference on Social Robotics (ICSR)*. Tsukuba, Japan: Springer.

— (2017c). “Human Perceptions of the Severity of Domestic Robot Errors”. In: *Social Robotics*. Ed. by Abderrahmane Kheddar et al. Cham: Springer International Publishing, pp. 647–656. ISBN: 978-3-319-70022-9.

— (2018). “The impact of peoples’ personal dispositions and personalities on their trust of robots in an emergency scenario.” In: *Paladyn Journal of Behavioral Robotics* 9 (1). DOI: https://doi.org/10.1515/pjbr-2018-0010.

— (2020). “How Social Robots Influence People’s Trust in Critical Situations”. In: *IEEE InternationalConference on Robot & Human Interactive Communication (RO-MAN)*. RO-MAN 2020.

Rossi, Alessandra, Kerstin Dautenhahn, Kheng Lee Koay, Michael L. Walters, and Patrick Holthaus (2020). “Evaluating people’s perceptions of trust in a robot in a repeated interactions study”. In: *12th International Conference on Social Robotics*. ICSR 2020.

Rossi, Alessandra, Fernando Garcia, Arturo Cruz Maya, et al. (2019). "Investigating the Effects of Social Interactive Behaviours of a Robot on People's Trust During a Navigation Task". In: *Towards Autonomous Robotic Systems (TAROS 2019), Lecture Notes in Computer Science*. Cham: Springer International Publishing, pp. 349–361. ISBN: 978-3-030-23807-0.

Rossi, Alessandra, Fernando Garcia, Arturo Cruz Maya, et al. (2019). "Investigating the Effects of Social Interactive Behaviours of a Robot on People's Trust During a Navigation Task". In: *Towards Autonomous Robotic Systems*. Springer, pp. 349–361.

Rossi, Alessandra, Patrick Holthaus, et al. (2018a). "Getting to Know Pepper: Effects of People's Awareness of a Robot's Capabilities on Their Trust in the Robot". In: *Proceedings of the 6th International Conference on Human-Agent Interaction*. HAI '18. Southampton, United Kingdom: ACM, pp. 246–252. ISBN: 978-1-4503-5953-5. DOI: 10.1145/3284432.3284464.

— (2018b). "Programming Pepper: What can you make a humanoid robot do?" In: *Proceedings of the 3rd Workshop on Behaviour Adaptation, Interaction and Learning for Assistive Robotics*. RO-MAN 2018. Nanjing, China, pp. 27–31.

Rossi, S., A. Rossi, and K. Dautenhahn (2020). "The Secret Life of Robots: Perspectives and Challenges for Robot's Behaviours During Non-interactive Tasks". In: *International Journal of Social Robotics*. Ed. by Springer. DOI: 10.1007/s12369-020-00650-z.

Rotter, Julian B. (1967). "A new scale for the measurement of interpersonal trust1". In: *Journal of Personality* 35.4, pp. 651–665. DOI: 10.1111/j.1467-6494.1967.tb01454.x.

Sadeghi, Nabiollah et al. (2012). "Learning Styles, Personality Types and Reading Comprehension Performance". In: *Psychology*.

Saerbeck, M. and C. Bartneck (Mar. 2010). "Perception of affect elicited by robot motion". In: *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp. 53–60.

Salem, M. and K. Dautenhahn (2015). "Evaluating Trust and Safety in HRI: Practical Issues and Ethical Challenges." In: *Proceedings of the 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI 2015): Workshop on the Emerging Policy and Ethics of Human-Robot Interaction*.

Salem, M., F. Eyssel, et al. (2013). "To Err is Human(-like): Effects of Robot Gesture on Perceived Anthropomorphism and Likability". In: *International Journal of Social Robotics* 5.3, pp. 313–323.

Salem, Maha and Kerstin Dautenhahn (2015). *Evaluating trust and safety in HRI: Practical issues and ethical challenges.* URL: http://hdl.handle.net/2299/16336.

Salem, Maha, Friederike Eyssel, et al. (Aug. 2013). "To Err is Human(-like): Effects of Robot Gesture on Perceived Anthropomorphism and Likability". In: *International Journal of Social Robotics* 5.3, pp. 313–323.

Salem, Maha, Gabriella Lakatos, et al. (2015). "Would You Trust a (Faulty) Robot? Effects of Error, Task Type and Personality on Human-Robot Cooperation and Trust". In: *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*. HRI 15. Portland, Oregon, USA: Association for Computing Machinery, pp. 141–148. ISBN: 9781450328838. DOI: 10.1145/2696454.2696497.

Santamaria, Tricia and Dan Nathan-Roberts (2017). "Personality Measurement and Design in Human-Robot Interaction: A Systematic and Critical Review". In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 61.1, pp. 853–857. DOI: 10.1177/1541931213601686.

Sarkar, Satragni, Dejanira Araiza-Illan, and Kerstin Eder (2017). *Effects of Faults, Experience, and Personality on Trust in a Robot Co-Worker*. arXiv: 1703.02335.

Schilke, Oliver, Martin Reimann, and Karen S. Cook (2013). "Effect of relationship experience on trust recovery following a breach". In: *Proceedings of the National Academy of Sciences* 110.38, pp. 15236–15241. DOI: 10.1073/pnas.1314857110.

Schweitzer, Maurice E., John C. Hershey, and Eric T. Bradlow (2006). "Promises and lies: Restoring violated trust". In: *Organizational Behavior and Human Decision Processes* 101.1, pp. 1–19. ISSN: 0749-5978. DOI: j.obhdp.2006.05.005.

Sciutti, A. et al. (Mar. 2018). "Humanizing Human-Robot Interaction: On the Importance of Mutual Understanding". In: *IEEE Technology and Society Magazine* 37.1, pp. 22–29.

Sebo, S. S., P. Krishnamurthi, and B. Scassellati (2019). ""I Don't Believe You": Investigating the Effects of Robot Trust Violation and Repair". In: *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp. 57–65.

Seo, Stela H. et al. (2015). "Poor Thing! Would You Feel Sorry for a Simulated Robot? A Comparison of Empathy toward a Physical and a Simulated Robot". In: *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*. Portland, Oregon, USA: Association for Computing

Machinery, pp. 125–132. ISBN: 9781450328838. DOI: 10.1145/2696454.2696471. URL: 10.1145/2696454.2696471.

Sheehan, Kim and Matthew Pittman (2016). "The Academic's Guide to Using Amazon's Mechanical Turk: The HIT Handbook for Social Science Research". In: *Irving: Melvin & Leigh*.

Short, E. et al. (Mar. 2010). "No fair!! An interaction with a cheating robot". In: *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp. 219–226. DOI: 10.1109/HRI.2010.5453193.

Shu, T. et al. (May 2017). "Learning social affordance grammar from videos: Transferring human interactions to human-robot interactions". In: *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 1669–1676.

Sidorov, Grigoria et al. (Jan. 2019). "Human Interaction with Shopping Assistant Robot in Natural Language". In: *Journal of Intelligent and Fuzzy Systems* 36.5, pp. 4889–4899. DOI: 10.3233/JIFS-179036.

Simpson, J. A. (2007). "Foundations of interpersonal trust". In: *In A. W. Kruglanski, & E. T. Higgins (Eds.) Social psychology: Handbook of basic principles, New York: Guilford Press*, pp. 587–607.

Simpson, Jeffry A. (2007). "Psychological Foundations of Trust". In: *Current Directions in Psychological Science* 16.5, pp. 264–268. DOI: 10.1111/j.1467-8721.2007.00517.x.

Slovic, Paul (2000). *Perceived risk, trust, and democracy*. Vol. 13. Earthscan Pub, pp. 675–682. ISBN: 9781853835278;1853835277;9781853835285;1853835285;

Song, Sichao and Seiji Yamada (2017). "Expressing Emotions Through Color, Sound, and Vibration with an Appearance-Constrained Social Robot". In: *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*. HRI '17. Vienna, Austria: ACM, pp. 2–11. ISBN: 978-1-4503-4336-7. DOI: 10.1145/2909824.3020239.

Steinfeld, Aaron et al. (2006). "Common Metrics for Human-Robot Interaction". In: *Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-Robot Interaction*. HRI '06. Salt Lake City, Utah, USA: Association for Computing Machinery, pp. 33–40. ISBN: 1595932941. DOI: 10.1145/1121241.1121249.

Stiefelhagen, R. et al. (2004). "Natural Human-Robot Interaction using Speech, Head Pose and Gestures". In: *Intelligent Robots and Systems, 2004. Proceedings IEEE/RSJ International Conference* 3.

Stroessner, S.J. and J Benitez (2019). "The Social Perception of Humanoid and Non-Humanoid Robots: Effects of Gendered and Machinelike Features". In: *Interational Journal of Social robotics* 11, pp. 305–315. DOI: 10.1007/s12369-018-0502-7.

Syrdal, D S et al. (2009). "The Negative Attitudes Towards Robots Scale and reactions to robot behaviour in a live Human-Robot Interaction study". In: *AISB 2009 SSAISB*. Adaptive et al., pp. 109–115.
Syrdal, Dag Sverre, Kerstin Dautenhahn, et al. (2007). "Looking Good? Appearance Preferences and Robot Personality Inferences at Zero Acquaintance". In: *AAAI Spring Symposium: Multidisciplinary Collaboration for Socially Assistive Robotics*.
Syrdal, Dag Sverre, Kheng Lee Koay, et al. (2010). "Video prototyping of dog-inspired non-verbal affective communication for an appearance constrained robot". In: *Proceedings - IEEE International Workshop on Robot and Human Interactive Communication*, pp. 632–637. ISBN: 9781424479917. DOI: 10.1109/ROMAN.2010.5598693.
Tait, Sophie E. and Jeske Debora (2015). "Hello Stranger! Trust and Self-Disclosure Effects on Online Information Sharing." In: *International Journal of Cyber Behavior* 5.1, pp. 42–55. DOI: 10.4018/ijcbpl.2015010104.
Takayama, Leila and Caroline Pantofaru (2009). "Influences on Proxemic Behaviors in Human-Robot Interaction". In: *Proceedings of the 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IROS 09. St. Louis, MO, USA: IEEE Press, pp. 5495–5502. ISBN: 9781424438037.
Tan, Felix B. and Paul Sutherland (2004). "Online Consumer Trust: A Multi-Dimensional Model." In: *Journal of Electronic Commerce in Organizations (JECO)* 2, pp. 40–58.
Tannenbaum, Kendra R., Joseph K. Torgesen, and Richard K. Wagner (2006). "Relationships Between Word Knowledge and Reading Comprehension in Third-Grade Children". In: *Scientific Studies of Reading* 10.4, pp. 381–398. DOI: 10.1207/s1532799xssr1004_3.
Tapus, A., C. Tapus, and M.J. Matarić (2008). "User-robot personality matching and assistive robot behavior adaptation for post-stroke rehabilitation therapy". In: *Intel Serv Robotics* 1.169. DOI: 10.1007/s11370-008-0017-4.
Thoresen, C. J. et al. (2003). "The Affective Underpinnings of Job Perceptions and Attitudes: A Meta-Analytic Review and Integration." In: *Psychological Bulletin* 129.6, pp. 914–945. DOI: 10.1037/0033-2909.129.6.914.
Tonkin, Meg et al. (2017). "Embodiment, Privacy and Social Robots: May I Remember You?" In: *Social Robotics*. Ed. by Abderrahmane Kheddar et al. Cham: Springer International Publishing, pp. 506–515. ISBN: 978-3-319-70022-9.
Tseng, S. H. et al. (May 2013). "Human awareness based robot performance learning in a social environment". In: *2013 IEEE International Conference*

on Robotics and Automation, pp. 4291–4296. DOI: 10.1109/ICRA.2013.6631184.

Tusing, K. J. and J. P. Dillard (2000). "The sounds of dominance: Vocal precursors of perceived dominance during interpersonal influence." In: *Human Communication Research* 26.1, pp. 148–171. DOI: 10.1093/hcr/26.1.148.

Vigliocco, Gabriella, Pamela Perniss, and David Vinson (2014). "Language as a multimodal phenomenon: implications for language learning, processing and evolution". In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 369.1651, p. 20130292.

Visser, E. De et al. (2006). "A Comprehensive Methodology for Assessing Human-Robot Team Performance for Use in Training and Simulation". In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 50.25, pp. 2639–2643. DOI: 10.1177/154193120605002507.

Visser, Ewart J. de et al. (2020). "Towards a Theory of Longitudinal Trust Calibration in Human–Robot Teams." In: *International Journal of Social Robotics*, pp. 459–478. DOI: 10.1007/s12369-019-00596-x.

Voelpel, Sven C., Robert A. Eckhoff, and Jens Förster (2008). "David against Goliath? Group size and bystander effects in virtual knowledge sharing". In: *Human Relations* 61.2, pp. 271–295. DOI: 10.1177/0018726707087787.

Wada, K. et al. (2002). "Robot assisted activity for elderly people and nurses at a day service center". In: *Proceedings 2002 IEEE International Conference on Robotics and Automation (Cat. No.02CH37292)*. Vol. 2, 1416–1421 vol.2. DOI: 10.1109/ROBOT.2002.1014742.

Wainer, J. et al. (2014). "Using the humanoid robot KASPAR to autonomously play triadic games and facilitate collaborative play among children with autism". In: *IEEE Transactions on Autonomous Mental Development* 6(3), pp. 183–199. DOI: DOI:https://doi.org/10.1007/s12369-016-0369-4.

Walters, M L et al. (2009). "Preferences and perceptions of robot appearance and embodiment in human-robot interaction trials." In: *in Procs of New Frontiers in Human-Robot Interaction: Symposium at AISB09 Convention*, pp. 136–143.

Walters, Michael L. et al. (2011). "A long-term Human-Robot Proxemic study". In: *IEEE RO-MAN*, pp. 137–142. ISBN: 9781457715716;1457715716.

Wang, Ning et al. (Aug. 2015). "Intelligent Agents for Virtual Simulation of Human-Robot Interaction". In: *Virtual, Augmented and Mixed Reality*. Vol. 9179. Lecture Notes in Computer Science. Los Angeles, CA: Springer International Publishing, pp. 228–239. ISBN: 978-3-319-21066-7 978-3-319-21067-4. DOI: 10.1007/978-3-319-21067-424.

Warta, Samantha F. et al. (2018). “Effects of Social Cues on Social Signals in Human-Robot Interaction During a Hallway Navigation Task”. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 62.1, pp. 1128–1132. DOI: 10.1177/1541931218621258.

Weiss, A. and C. Bartneck (Aug. 2015). “Meta analysis of the usage of the Godspeed Questionnaire Series”. In: *2015 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pp. 381–388. DOI: 10.1109/ROMAN.2015.7333568.

Wickens, C. et al. (2013). *Engineering Psychology and Human Performance*. New York: Psychology Press. DOI: 10.4324/9781315665177.

Williamson, Jeanine M. (2018). “Chapter 1 - Individual Differences”. In: *Teaching to Individual Differences in Science and Engineering Librarianship*. Ed. by Jeanine M. Williamson. Chandos Publishing, pp. 1–10. ISBN: 978-0-08-101881-1. DOI: 10.1016/B978-0-08-101881-1.00001-7.

Wood, TJ (2014). “Exploring the role of first impressions in rater-based assessments”. In: *Adv Health Sci Educ Theory Pract* 19.3, pp. 409–427. DOI: 10.1007/s10459-013-9453-9.

Woods, S. et al. (Mar. 2006). “Comparing human robot interaction scenarios using live and video based methods: towards a novel methodological approach”. In: *9th IEEE International Workshop on Advanced Motion Control, 2006.* Pp. 750–755. DOI: 10.1109/AMC.2006.1631754.

Woods, Sarah et al. (2007). “Are robots like people?: Relationships between participant and robot personality traits in human–robot interaction studies”. In: *Interaction Studies* 8.2, pp. 281–305. ISSN: 1572-0373. DOI: https://doi.org/10.1075/is.8.2.06woo.

Wullenkord, R. and F. Eyssel (Aug. 2014). “Improving attitudes towards social robots using imagined contact”. In: *The 23rd IEEE International Symposium on Robot and Human Interactive Communication*, pp. 489–494. DOI: 10.1109/ROMAN.2014.6926300.

Wykowska, Agnieszka et al. (Aug. 2014). “Implications of Robot Actions for Human Perception. How Do We Represent Actions of the Observed Robots?” In: *International Journal of Social Robotics* 6.3, pp. 357–366.

Yamada, S., T. Nomura, and T. Kanda (2019). “Healthcare Support by a Humanoid Robot”. In: *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp. 1–2. DOI: 10.1109/HRI.2019.8673072.

Yechiam, Eldad and Itzhak Aharon (2012). “Experience-Based Decisions and Brain Activity: Three New Gaps and Partial Answers”. In: *Frontiers in Psychology* 2, p. 390. ISSN: 1664-1078. DOI: 10.3389/fpsyg.2011.00390.

Yu, Kun et al. (2017). “User Trust Dynamics: An Investigation Driven by Differences in System Performance”. In: *ACM* 126745, pp. 307–317.
Zajonc, R. B. (1968). “Attitudinal effects of mere exposure.” In: *Journal of Personality and Social Psychology* 9.2, Pt.2, pp. 1–27. DOI: doi.org/10.1037/h0025848.
Zhou, Allan et al. (2017). “Expressive Robot Motion Timing”. In: *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction - HRI*.
Zlotowski, Jakub et al. (2015). “Anthropomorphism: Opportunities and Challenges in Human–Robot Interaction”. In: *International Journal of Social Robotics* 7, pp. 347–360.

Appendices

Appendix A

Publications

This Appendix includes my publications during her research project related to HRI. A full list of publications can be found on her personal website[1].

2017 A. Rossi, K. Dautenhahn, K. L. Koay, M. L. Walters, How the Timing and Magnitude of Robot Errors Influence Peoples' Trust of Robots in an Emergency Scenario. In: Kheddar A. et al. (eds) Social Robotics. ICSR 2017. Lecture Notes in Computer Science, vol 10652. Springer, Cham

2017 A. Rossi, K. Dautenhahn, K. L. Koay, M. L. Walters, Human Perceptions of the Severity of Domestic Robot Errors. In: Kheddar A. et al. (eds) Social Robotics. ICSR 2017. Lecture Notes in Computer Science, vol 10652. Springer, Cham

2017 A. Rossi, K. Dautenhahn, K. L. Koay and J. Saunders, Investigating human perceptions of trust in robots for safe HRI in home environments, Human-Robot Interaction Pioneers Workshop in conjunction with the 2017 ACM/IEEE International Conference on Human-Robot Interaction, March, 2017.

2017 B. Byrne, A. Rossi and M. Doolan, Humanoid Robots attending Lectures – Research Informed Teaching, 9th International Conference on Education Technology and Computers (ICETC 2017), Barcelona, Spain. December 20-22, 2017

[1] Alessandra Rossi's personal website `https://www.alessandrarossi.net`

2017 B. Byrne and A. Rossi, Research Informed Teaching – Pepper the Robot attends Lectures, 9th International Conference on Education and New Learning Technologies, pp. 2963, Barcelona, Spain, July 3-5, 2017.

2018 A. Rossi, K. Dautenhahn, K., Koay and M.L. Walters, The impact of peoples' personal dispositions and personalities on their trust of robots in an emergency scenario. Paladyn, Journal of Behavioral Robotics, 9(1), pp. 137-154

2018 A. Rossi, K. Dautenhahn, K. L. Koay and M. L. Walters, Investigating human perception of trust and social cues in robots for safe HRI in home environments, International PhD Conference on Safe and Social Robotics (SSR-2018), Madrid, 29-30 September 2018

2018 A. Rossi, P. Holthaus, K. Dautenhahn, K. L. Koay, and M. L. Walters. Getting to know Pepper: Effects of people's awareness of a robot's capabilities on their trust in the robot. In Proceedings of the 6th International Conference on Human-Agent Interaction (HAI '18). ACM, New York, NY, USA, 246-252

2018 A. Rossi, P. Holthaus, K. Dautenhahn, K.L. Koay and M.L. Walters, Programming Pepper: What can you make a humanoid robot do?, 3rd Workshop on Behaviour Adaptation, Interaction and Learning for Assistive Robotics - IEEE-RO-MAN 2018, 27 - 31 of August 2018, Nanjing, China

2019 A. Rossi, F. Garcia, A. Cruz Maya, K. Dautenhahn, K. L. Koay, M. L. Walters and A. K. Pandey. Investigating the Effects of Social Interactive Behaviours of a Robot on People's Trust During a Navigation Task. In: Althoefer K., Konstantinova J., Zhang K. (eds) Towards Autonomous Robotic Systems. TAROS 2019. Lecture Notes in Computer Science, vol 11649. Springer, Cham

2019 A. Rossi, K. Dautenhahn, K.L. Koay and M.L. Walters, Evaluating the effects of an autonomous robot's social behaviours on people's trust. IEEE International Conference on Robot & Human Interactive Communication (RO-MAN)

2019 A. Rossi, K. Dautenhahn, K.L. Koay and M.L. Walters, Investigating human perceptions of trust and social cues in robots for safe HRI, Personal robotics and secure human-robot collaboration Workshop, IEEE ICDL- EPIROB, Oslo, Norway

2019 A. Rossi, S. Moros, K. Dautenhahn, K.L. Koay and M.L. Walters, Getting to know Kaspar: Effects of people's awareness of a robot's capabilities on their trust in the robot. IEEE International Conference on Robot & Human Interactive Communication (RO-MAN), New Delhi, India

2020 S. Rossi, A. Rossi and K. Dautenhahn. The Secret Life of Robots: Perspectives and Challenges for Robot's Behaviours During Non-interactive Tasks. International Journal of Social Robotics. Springer Nature.

2020 A. Rossi, K. Dautenhahn, K.L. Koay and M.L. Walters, Human Perceptions of the Social Cues Expressed by a Domestic Robot. IEEE International Conference on Robot & Human Interactive Communication (RO-MAN), Naples, Italy (submitted)

2020 A. Rossi, K. Dautenhahn, K.L. Koay and M.L. Walters, How Social Robots Influence People's Trust in Critical Situations. IEEE International Conference on Robot & Human Interactive Communication (RO-MAN), Naples, Italy (accepted)

2020 S. Rossi, A. Rossi and K. Dautenhahn. The Secret Life of Robots: Perspectives and Challenges for Robot's Behaviours During Non-interactive Tasks.International Journal of Social Robotics. Springer Nature.

2020 M. Keijsers and A. Rossi. RoboCup 2050: from human fouls to robot bullyism. Virtual RoboCup Humanoid Open Workshops (V-RoHOW) (accepted).

Appendix B

Study 1

Human Perceptions of the Severity of Domestic Robot Errors: Ethics approval and Questionnaire.

B.1 Ethics Approval

University of Hertfordshire UH

UNIVERSITY OF HERTFORDSHIRE
SCIENCE & TECHNOLOGY

ETHICS APPROVAL NOTIFICATION

TO	Alessandra Rossi
CC	Prof. Dr Kerstin Dautenhahn
FROM	Dr Simon Trainis, Science and Technology ECDA Chairman
DATE	16/06/16

Protocol number: **COM/PGR/UH/02058**

Title of study: Evaluating possible errors of home companion robots

Your application for ethics approval has been accepted and approved by the ECDA for your School.

This approval is valid:

From: 16/06/16

To: 30/09/16

Please note:

Approval applies specifically to the research study/methodology and timings as detailed in your Form EC1. Should you amend any aspect of your research, or wish to apply for an extension to your study, you will need your supervisor's approval and must complete and submit form EC2. In cases where the amendments to the original study are deemed to be substantial, a new Form EC1 may need to be completed prior to the study being undertaken.

Should adverse circumstances arise during this study such as physical reaction/harm, mental/emotional harm, intrusion of privacy or breach of confidentiality this must be reported to the approving Committee immediately. Failure to report adverse circumstance/s would be considered misconduct.

Ensure you quote the UH protocol number and the name of the approving Committee on all paperwork, including recruitment advertisements/online requests, for this study.

B.2 Questionnaire

Evaluating possible errors of home companion robots

ID: ____________________

#	QUESTION	1 (small error)	2	3	4	5	6	7 (big error)
1	You ask for a cup of coffee. Your robot brings you an orange.							
2	Your robot spills coffee on your carpet.							
3	You ask your robot to charge your phone. Your robot puts it in the toaster.							
4	You want to drink some cold fruit juice. Your robot goes to heat it up.							
5	Your robot is preparing a drink for you. You asked for sugar, your robot brings you salt.							
6	After a meal, your robot puts the remaining food into the washing machine instead of the bin.							
7	You are watching your favourite show on tv and your robot changes the channel.							
8	You are sit on the right side of a table, your robot puts your drink on the opposite side.							
9	You and your robot are solving a puzzle. You ask your robot to take a piece useful to solve the puzzle. Your robot brings you the wrong piece.							
10	Your robot leaves your pet hamster outside the house in very cold weather.							
11	You are having dinner with friends. Your robot brings you the trash and reminds you to take it out.							
12	You share some private information about yourself with the robot. Your robot reveals it to a visitor.							
13	In your entrance hall you have a little table with a beautiful vase. Your robot bumps into it and the vase crashes to the floor.							
14	After preparing dinner for you, your robot forgets to turn off the cooker.							
15	Your robot keeps track of your calendar and today you have an appointment for a job interview. Your robot forgets to remind you.							
16	You have just fallen asleep. Your robot turns on loud music.							
17	Your robot burns your t-shirt while ironing it.							
18	Your robot brings you a dishwashing tablet instead of paracetamol.							
19	Your robot brings you vinegar when you are thirsty and asked for water.							
20	You are sitting on the sofa. You asked your robot to show you the latest news. Your robot shows it on his own screen that faces away from you.							

Please answer to these two questions [OPTIONAL]:
1) Can you think of another small error? ____

____

2) Can you think of another big error? ____

____

Appendix C

Study 2

How the Timing and Magnitude of Robot Errors Influence People's Trust of Robots in an Emergency Scenario: Ethics approval, extension of Ethics approval, Pre and Post Questionnaires, and an examples of narrative interaction used for this study.

C.1 Ethics Approval

University of Hertfordshire UH

SCIENCE & TECHNOLOGY ECDA

ETHICS APPROVAL NOTIFICATION

TO Alessandra Rossi

CC Prof Kerstin Dautenhahn

FROM Dr Simon Trainis, Science and Technology ECDA Chairman

DATE 20/10/2016

Protocol number: **COM/PGR/UH/02076**

Title of study: "Evaluating errors' impact on human perceptions of robots"

Your application for ethics approval has been accepted and approved by the ECDA for your School.

This approval is valid:

From: 20/10/2016

To: 31/03/2017

Please note:

Approval applies specifically to the research study/methodology and timings as detailed in your Form EC1. Should you amend any aspect of your research, or wish to apply for an extension to your study, you will need your supervisor's approval and must complete and submit form EC2. In cases where the amendments to the original study are deemed to be substantial, a new Form EC1 may need to be completed prior to the study being undertaken.

Should adverse circumstances arise during this study such as physical reaction/harm, mental/emotional harm, intrusion of privacy or breach of confidentiality this must be reported to the approving Committee immediately. Failure to report adverse circumstance/s would be considered misconduct.

Ensure you quote the UH protocol number and the name of the approving Committee on all paperwork, including recruitment advertisements/online requests, for this study.

Students must include this Approval Notification with their submission.

University of Hertfordshire UH

HEALTH SCIENCES ENGINEERING & TECHNOLOGY ECDA

ETHICS APPROVAL NOTIFICATION

TO	Alessandra Rossi
CC	Kerstin Dautenhahn
FROM	Dr Simon Trainis, Health, Science, Engineering & Technology ECDA Chair
DATE	4/4/17

Protocol number: aCOM/PGR/UH/02076

Title of study: Evaluating errors' impact on human perceptions of robots

Your application to modify extend the existing protocol as detailed below has been accepted and approved by the ECDA for your School and includes work undertaken for this study by the named additional workers below:

This approval is valid:

From: 4/4/17

To: 31/3/18

Additional workers: no additional workers named.

Please note:

Any conditions relating to the original protocol approval remain and must be complied with.

Approval applies specifically to the research study/methodology and timings as detailed in your Form EC1 or as detailed in the EC2 request. Should you amend any further aspect of your research, or wish to apply for an extension to your study, you will need your supervisor's approval and must complete and submit a further EC2 request. In cases where the amendments to the original study are deemed to be substantial, a new Form EC1 may need to be completed prior to the study being undertaken.

Should adverse circumstances arise during this study such as physical reaction/harm, mental/emotional harm, intrusion of privacy or breach of confidentiality this must be reported to the approving Committee immediately. Failure to report adverse circumstance/s would be considered misconduct.

Ensure you quote the UH protocol number and the name of the approving Committee on all paperwork, including recruitment advertisements/online requests, for this study.

Students must include this Approval Notification with their submission.

C.2 Questionnaire Pre Experiment

QUESTIONNAIRE – PRE

Your age: ____________________

Your gender: ____________________

Your country of residence: ____________________

Do you find yourself … ?

1. Extraverted, enthusiastic

	1	2	3	4	5	6	7	
disagree strongly	○	○	○	○	○	○	○	agree strongly

2. Critical, quarrelsome

	1	2	3	4	5	6	7	
disagree strongly	○	○	○	○	○	○	○	agree strongly

3. Dependable, self-disciplined

	1	2	3	4	5	6	7	
disagree strongly	○	○	○	○	○	○	○	agree strongly

4. Anxious, easily upset

	1	2	3	4	5	6	7	
disagree strongly	○	○	○	○	○	○	○	agree strongly

5. Open to new experiences, complex

	1	2	3	4	5	6	7	
disagree strongly	○	○	○	○	○	○	○	agree strongly

6. Reserved, quiet.

	1	2	3	4	5	6	7	
disagree strongly	○	○	○	○	○	○	○	agree strongly

7. Sympathetic, warm.

	1	2	3	4	5	6	7	
disagree strongly	○	○	○	○	○	○	○	agree strongly

8. Disorganized, careless.

	1	2	3	4	5	6	7	
disagree strongly	○	○	○	○	○	○	○	agree strongly

9. Calm, emotionally stable.

	1	2	3	4	5	6	7	
disagree strongly	◯	◯	◯	◯	◯	◯	◯	agree strongly

10. Conventional, uncreative.

	1	2	3	4	5	6	7	
disagree strongly	◯	◯	◯	◯	◯	◯	◯	agree strongly

Do you think … ?

11. In general, people really do care about the well-being of others.

	1	2	3	4	5	6	7	
disagree strongly	◯	◯	◯	◯	◯	◯	◯	agree strongly

12. The typical person is sincerely concerned about the problems of others.

	1	2	3	4	5	6	7	
disagree strongly	◯	◯	◯	◯	◯	◯	◯	agree strongly

13. Most of the time, people care enough to try to be helpful, rather than just looking out for themselves.

	1	2	3	4	5	6	7	
disagree strongly	◯	◯	◯	◯	◯	◯	◯	agree strongly

14. In general, most people keep their promises.

	1	2	3	4	5	6	7	
disagree strongly	◯	◯	◯	◯	◯	◯	◯	agree strongly

15. I think people generally try to back up their words with their actions.

	1	2	3	4	5	6	7	
disagree strongly	◯	◯	◯	◯	◯	◯	◯	agree strongly

16. Most people are honest in their dealings with others.

	1	2	3	4	5	6	7	
disagree strongly	◯	◯	◯	◯	◯	◯	◯	agree strongly

17. I believe that most professional people do a very good job at their work.

	1	2	3	4	5	6	7	
disagree strongly	◯	◯	◯	◯	◯	◯	◯	agree strongly

18. Most professionals are very knowledgeable in their chosen field.

	1	2	3	4	5	6	7	
disagree strongly	◯	◯	◯	◯	◯	◯	◯	agree strongly

19. A large majority of professional people are competent in their area of expertise.

	1	2	3	4	5	6	7	
disagree strongly	○	○	○	○	○	○	○	agree strongly

20. I usually trust people until they give me a reason not to trust them.

	1	2	3	4	5	6	7	
disagree strongly	○	○	○	○	○	○	○	agree strongly

21. I generally give people the benefit of the doubt when I first meet them.

	1	2	3	4	5	6	7	
disagree strongly	○	○	○	○	○	○	○	agree strongly

22. My typical approach is to trust new acquaintances until they prove I should not trust them.

	1	2	3	4	5	6	7	
disagree strongly	○	○	○	○	○	○	○	agree strongly

23. Do you have any experience interacting with robots?

	1	2	3	4	5	6	7	
not at all	○	○	○	○	○	○	○	very much

24. Please, specify what kind of experience you have with robots[1]:

- ☐ Participant in other experiment
- ☐ Developer
- ☐ Observer
- ☐ Researcher

25. Which robots[2]? ______________________________

26. Would you feel comfortable having a robot as a companion in your home?

	1	2	3	4	5	6	7	
not at all	○	○	○	○	○	○	○	very much

27. Would you expect the robot to help you in doing your everyday activities?

	1	2	3	4	5	6	7	
not at all	○	○	○	○	○	○	○	very much

28. Do you perceive robots as a … ?

- ☐ Friend
- ☐ Butler
- ☐ Assistant
- ☐ Tool
- ☐ Companion
- ☐ Pet
- ☐ Machine
- ☐ Other: _______

---

[1] **it will appear only if the answer to question 23 is >=2**

[2] **it will appear only if the answer to question 23 is >=2**

C.3 Questionnaire Post Experiment

QUESTIONNAIRE – POST

1. Had the person who visited you grey hair?

2. What piece was missing in the puzzle you and your robot Jace solved?

3. Which secret did your robot Jace tell you?

4. Did you ask for a cup of coffee and your robot Jace brought you an orange?

5. Would you like Jace as a companion in your home?

	1	2	3	4	5	6	7	
not at all	◯	◯	◯	◯	◯	◯	◯	very much

6. You would like a robot, different from Jace, as a companion.

	1	2	3	4	5	6	7	
disagree strongly	◯	◯	◯	◯	◯	◯	◯	agree strongly

7. Why did/didn't you trust your robot Jace?[1]

8. The scenario looked realistic.

	1	2	3	4	5	6	7	
disagree strongly	◯	◯	◯	◯	◯	◯	◯	agree strongly

9. My robot Jace always provided the advice I required to make my decision.

	1	2	3	4	5	6	7	
disagree strongly	◯	◯	◯	◯	◯	◯	◯	agree strongly

10. My robot Jace performed reliably.

	1	2	3	4	5	6	7	
disagree strongly	◯	◯	◯	◯	◯	◯	◯	agree strongly

11. My robot Jace analyzes problems consistently.

	1	2	3	4	5	6	7	
disagree strongly	◯	◯	◯	◯	◯	◯	◯	agree strongly

[1]This question will change (did or didn't) according the choice of the participant in the last task.

12. I believe advice from my robot Jace even when I do not know for certain that Jace is correct.

	1	2	3	4	5	6	7	
disagree strongly	○	○	○	○	○	○	○	agree strongly

13. When I am uncertain about a decision I believe my robot Jace rather than myself.

	1	2	3	4	5	6	7	
disagree strongly	○	○	○	○	○	○	○	agree strongly

14. When a robot gives unusual advice I am confident that the advice is correct.

	1	2	3	4	5	6	7	
disagree strongly	○	○	○	○	○	○	○	agree strongly

15. Even if I have no reason to expect a robot will be able to solve a difficult problem, I still feel certain that my robot Jace will.

	1	2	3	4	5	6	7	
disagree strongly	○	○	○	○	○	○	○	agree strongly

16. How did you perceive your robot Jace? (select one)

☐ Friend	☐ Butler	☐ Machine	☐ Tool
☐ Companion	☐ Pet	☐ Assistant	☐ Other: _______

17. How did you perceive your robot Jace? (select one)

☐ Child-like	☐ Adult-like	☐ Animal-like	☐ Toy-like
☐ Other: _______			

18. How did you perceive your robot Jace? (select at least one)

☐ Friendly	☐ Funny	☐ Odd	☐ Irritating
☐ Helpful	☐ Annoying	☐ Other: _______	

How do you judge the following errors … ?

19. You are sitting on the right side of a table, your robot Jace puts your drink on the opposite side.

	1	2	3	4	5	6	7	
small error	○	○	○	○	○	○	○	big error

20. After a meal, your robot Jace puts the remaining food into the washing machine instead of the bin.

	1	2	3	4	5	6	7	
small error	○	○	○	○	○	○	○	big error

21. You and your robot Jace are solving a puzzle. You ask Jace to take a piece useful to solve the puzzle. Your robot brings you the wrong piece.

	1	2	3	4	5	6	7	
small error	○	○	○	○	○	○	○	big error

22. You share some private information about yourself with your robot Jace. Jace reveals it to a visitor.

	1	2	3	4	5	6	7	
small error	○	○	○	○	○	○	○	big error

23. You are watching your favourite show on tv and your robot Jace changes the channel.

	1	2	3	4	5	6	7	
small error	○	○	○	○	○	○	○	big error

24. You ask your robot Jace to charge your phone. Jace puts it in the toaster.

	1	2	3	4	5	6	7	
small error	○	○	○	○	○	○	○	big error

C.4 Example of Storyboard

An example of the interactive storyboard as presented to the participants is available at https://www.dropbox.com/s/c61ulmbkoa4tjbm/big_small_error_storyboard.pdf. The example shows the condition **C3** (big errors at the beginning and small errors at the end).

Appendix D

Study 3

Effects of People's Awareness of a Robot's Capabilities on their Trust in the Robot: Ethics approval, extension of Ethics approval, Questionnaires after each experimental conditions.

D.1 Ethics Approval

University of Hertfordshire UH

HEALTH SCIENCE ENGINEERING & TECHNOLOGY ECDA

ETHICS APPROVAL NOTIFICATION

TO	Alessandra Rossi
CC	Kerstin Dautenhahn
FROM	Dr Simon Trainis, Health, Science, Engineering & Technology ECDA Chair
DATE	07/06/2018

Protocol number:	**COM/PGR/UH/03338**
Title of study:	UK Robotics Week 2018 at Stanborough School in Welwyn Garden City.

Your application for ethics approval has been accepted and approved by the ECDA for your School and includes work undertaken for this study by the named additional workers below:

This approval is valid:

From: 07/06/2018

To: 21/07/2018

Additional workers: Dr. Patrick Holthaus 755439

Please note:

If your research involves invasive procedures you are required to complete and submit an EC7 Protocol Monitoring Form, and your completed consent paperwork to this ECDA once your study is complete. You are also required to complete and submit an EC7 Protocol Monitoring Form if you are a member of staff. This form is available via the Ethics Approval StudyNet Site via the 'Application Forms' page http://www.studynet1.herts.ac.uk/ptl/common/ethics.nsf/Teaching+Documents?Openview&count=9999&restricttocategory=Application+Forms

Any necessary permissions for the use of premises/location and accessing participants for your study must be obtained in writing prior to any data collection commencing. Failure to obtain adequate permissions may be considered a breach of this protocol.

Approval applies specifically to the research study/methodology and timings as detailed in your Form EC1A. Should you amend any aspect of your research, or wish to apply for an extension to your study, you will need your supervisor's approval (if you are a student) and must complete and submit form EC2. In cases where the amendments to the original study are deemed to be substantial, a new Form EC1A may need to be completed prior to the study being undertaken.

HEALTH SCIENCE ENGINEERING & TECHNOLOGY ECDA

ETHICS APPROVAL NOTIFICATION

TO	Dr Luke Wood
FROM	Dr Simon Trainis, Health, Science, Engineering & Technology ECDA Chair
DATE	14/05/2018

Protocol number: **COM/SF/UH/03320**

Title of study: UK Robotics Week 2018 in Hatfield Community Free School

Your application for ethics approval has been accepted and approved by the ECDA for your School and includes work undertaken for this study by the named additional workers below:

This approval is valid:

From: 14/05/2018

To: 21/07/2018

Additional workers: Silvia Moros: 755414
Luke Hickton: 99001455
Dr. Abolfazl Zaraki: 753975
Alessandra Rossi: 15058449
Dr. Patrick Holthaus: 755439
Dr. Ben Robins: 721266
Dr. Gabriella Lakatos: 749665
Dag Sverre Syrdal: 728132

Please note:

If your research involves invasive procedures you are required to complete and submit an EC7 Protocol Monitoring Form, and your completed consent paperwork to this ECDA once your study is complete. You are also required to complete and submit an EC7 Protocol Monitoring Form if you are a member of staff. This form is available via the Ethics Approval StudyNet Site via the 'Application Forms' page http://www.studynet1.herts.ac.uk/ptl/common/ethics.nsf/Teaching+Documents?Openview&count=9999&restricttocategory=Application+Forms

Any necessary permissions for the use of premises/location and accessing participants for your study must be obtained in writing prior to any data collection commencing. Failure to obtain adequate permissions may be considered a breach of this protocol.

Approval applies specifically to the research study/methodology and timings as detailed in your Form EC1A. Should you amend any aspect of your research, or wish to apply for an extension to your study, you will need your supervisor's approval (if you are a student) and must complete and submit form EC2. In cases where the

D.2 Questionnaire Post Video Interaction

ID ________

Name ____

Example

1. Question : Do you like ice cream?
Your answer:

No | Maybe | Yes

This means you like the ice cream

2. Question : Would you like to watch the new movie of the Disney?
Your answer:

Definitely No | No | Maybe | Yes | Definitely Yes

This means you do not want to watch the movie.

1. Have you ever seen Pepper before today?

No | Maybe | Yes

2. If yes, where: ____

____

3. Would you like to have Pepper in your home?

No | Maybe | Yes

4. Do you trust Pepper to be able to help you with your homework?

Definitely No | No | Maybe | Yes | Definitely Yes

5. Do you trust Pepper to wake you up in time for going to school?

6. Do you trust Pepper to be able to warn you of danger, e.g. when using the Internet?

7. Do you trust Pepper to help you in case of danger?

If you have any comments about the questions above or the robot please write them here[+]:

+ Comments aren't mandatory.

ID ________

Name ______________________________

Example

1. Question : Do you like ice cream?

Your answer:

No   Maybe   Yes

2. Question : Would you like to watch the new movie of the Disney?

Your answer:

Definitely No   No   Maybe   Yes   Definitely Yes

1. Have you ever seen Kaspar before today? *

No   Maybe   Yes

2. If yes, where*: ______________________________

____

3. Would you like to have Kaspar in your home?

No   Maybe   Yes

4. Do you trust Kaspar to be able to help you with your homework?

Definitely No   No   Maybe   Yes   Definitely Yes

5. Do you trust Kaspar to wake you up in time for going to school?

6. Do you trust Kaspar to be able to warn you of danger, e.g. when using the Internet?

7. Do you trust Kaspar to help you in case of danger?

If you have any comments about the questions above or the robot please write them here[+]:

* These questions won't be repeated.
\+ Comments aren't mandatory.

D.3 Questionnaire Post Live Interaction

ID __________
Name ____

1. Would you like to have Pepper in your home?

2. Do you trust Pepper to be able to help you with your homework?

3. Do you trust Pepper to wake you up in time for going to school?

4. Do you trust Pepper to be able to warn you of danger, e.g. when using the Internet?

5. Do you trust Pepper to help you in case of danger?

If you have any comments about the questions above or the robot please write them here[+]:

____

____

____

____

____

____

____

____

____

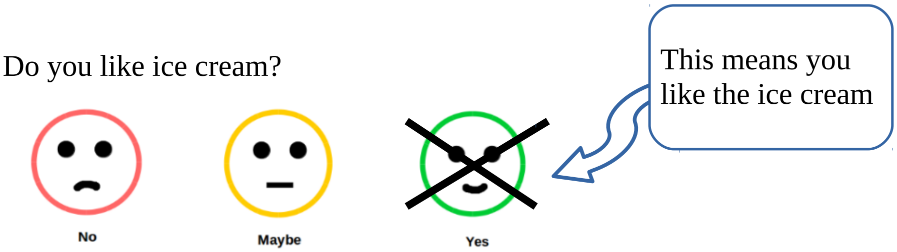
____

+ Comments aren’t mandatory.

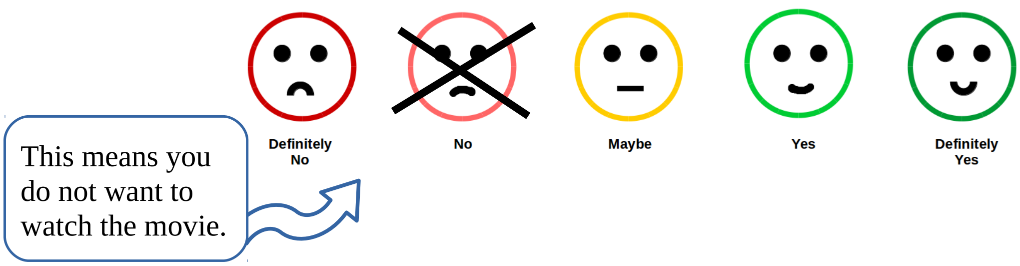
Name ____ ID __________

Example

1. Question : Do you like ice cream?
Your answer:



2. Question : Would you like to watch the new movie of the Disney?
Your answer:



1. Would you like to have Kaspar in your home?

2. Do you trust Kaspar to be able to help you with your homework?

3. Do you trust Kaspar to wake you up in time for going to school?

4. Do you trust Kaspar to be able to warn you of danger, e.g. when using the Internet?

5. Do you trust Kaspar to help you in case of danger?

If you have any comments about the questions above or the robot please write them here[+]:

+ Comments aren’t mandatory.

D.4 Questionnaire Post Programming Interaction

Name ______________________________ ID ________

1. Programming Pepper was?

Very Boring | Boring | OK | Fun | Very Fun

2. Programming Pepper was?

Very Hard | Hard | OK | Easy | Very Easy

3. Would you like to program Pepper again?

Definitely No | No | Maybe | Yes | Definitely Yes

4. Would you like to have Pepper in your home?

No | Maybe | Yes

5. Do you trust Pepper to be able to help you with your homework?

Definitely No | No | Maybe | Yes | Definitely Yes

6. Do you trust Pepper to wake you up in time for going to school?

Definitely No | No | Maybe | Yes | Definitely Yes

7. Do you trust Pepper to be able to warn you of danger, e.g. when using the Internet?

Definitely No | No | Maybe | Yes | Definitely Yes

8. Do you trust Pepper to help you in case of danger?

Definitely No | No | Maybe | Yes | Definitely Yes

If you have any comments about the questions above or the robot please write them here[+]:

+ Comments aren’t mandatory.

Name: ____ ID: ____________

Programming Kaspar was?

Very Fun	Fun	OK	Boring	Very Boring

Programming Kaspar was?

Very Easy	Easy	OK	Hard	Very Hard

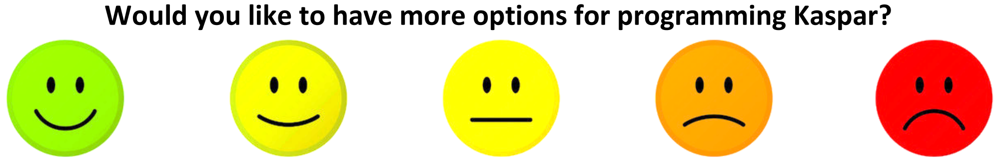
Would you like to program Kaspar again?

Yes Definitely	Yes	Maybe	No	Definitely No

How useful is the Scratch software for programming robots?

Very useful	Useful	OK	Not Useful	Really Not Useful

Would you like to have more options for programming Kaspar?



Yes Definitely	Yes	Maybe	No	Definitely No

If yes, which?

□ More movements

□ More sounds

□ Other: _______________________

Appendix E

Study 4

Human Perceptions of the Social Cues Expressed by a Domestic Robot: Ethics approval, extension of Ethics approval, Questionnaires after each videos and post experiment.

E.1 Ethics Approval

University of Hertfordshire UH

HEALTH SCIENCE ENGINEERING & TECHNOLOGY ECDA

ETHICS APPROVAL NOTIFICATION

TO: Alessandra Rossi

CC: Prof. Kerstin Dautenhahn

FROM: Dr Simon Trainis, Health, Sciences, Engineering & Technology ECDA Chair

DATE: 20th June 2017

Protocol number: COM/PGR/UH/02869

Title of study: "*Evaluating robot's behaviours.*"

Your application for ethics approval has been accepted and approved by the ECDA for your School and includes work undertaken for this study by the named additional workers below:

This approval is valid:

From: 20/06/17

To: 31/10/17

Additional workers: no additional workers named

Please note:

Approval applies specifically to the research study/methodology and timings as detailed in your Form EC1. Should you amend any aspect of your research, or wish to apply for an extension to your study, you will need your supervisor's approval and must complete and submit form EC2. In cases where the amendments to the original study are deemed to be substantial, a new Form EC1 may need to be completed prior to the study being undertaken.

Should adverse circumstances arise during this study such as physical reaction/harm, mental/emotional harm, intrusion of privacy or breach of confidentiality this must be reported to the approving Committee immediately. Failure to report adverse circumstance/s would be considered misconduct.

Ensure you quote the UH protocol number and the name of the approving Committee on all paperwork, including recruitment advertisements/online requests, for this study.

Students must include this Approval Notification with their submission.

University of Hertfordshire UH

HEALTH SCIENCE ENGINEERING & TECHNOLOGY ECDA

ETHICS APPROVAL NOTIFICATION

TO	Alessandra Rossi
CC	Kerstin Dautenhahn
FROM	Dr Simon Trainis, Health, Sciences, Engineering & Technology ECDA Chair
DATE	9th October 2017

Protocol number: aCOM PGR UH 02869(1)

Title of study: *Evaluating robot's behaviours.*

Your application to modify and extend the existing protocol as detailed below has been accepted and approved by the ECDA for your School and includes work undertaken for this study by the named additional workers below:

Modification: Amended/Extended dates.

This approval is valid:

From: 31/10/2017

To: 31/05/2018

Additional workers: no additional workers named.

Please note:

Any conditions relating to the original protocol approval remain and must be complied with.

Approval applies specifically to the research study/methodology and timings as detailed in your Form EC1 or as detailed in the EC2 request. Should you amend any further aspect of your research, or wish to apply for an extension to your study, you will need your supervisor's approval and must complete and submit a further EC2 request. In cases where the amendments to the original study are deemed to be substantial, a new Form EC1 may need to be completed prior to the study being undertaken.

Should adverse circumstances arise during this study such as physical reaction/harm, mental/emotional harm, intrusion of privacy or breach of confidentiality this must be reported to the approving Committee immediately. Failure to report adverse circumstance/s would be considered misconduct.

Ensure you quote the UH protocol number and the name of the approving Committee on all paperwork, including recruitment advertisements/online requests, for this study.

Students must include this Approval Notification with their submission.

E.2 Questionnaire Post Video

QUESTIONNAIRE

- Do you think the robot in the video behaved in a sociable way? *

	1	2	3	4	5	
not social at all	☐	☐	☐	☐	☐	very social

- Please can you give your reasons for your rating below?

* Mandatory fields

E.3 Questionnaire Post Experiment

QUESTIONNAIRE

How would you like to be alerted by your robot in the following scenarios?
Please, choose at least one answer for each case.

1. You forgot to turn your oven off.

☐ Body movement ☐ Head movement ☐ Gaze movement

☐ Approach orientation ☐ Approach distance ☐ Vocal cues

☐ Speech ☐ Eye colours ☐ Written text

Other: ________

2. You have an appointment with your friend for a drink.

☐ Body movement ☐ Head movement ☐ Gaze movement

☐ Approach orientation ☐ Approach distance ☐ Vocal cues

☐ Speech ☐ Eye colours ☐ Written text

Other: ________

3. There is a broken glass in the kitchen.

☐ Body movement ☐ Head movement ☐ Gaze movement

☐ Approach orientation ☐ Approach distance ☐ Vocal cues

☐ Speech ☐ Eye colours ☐ Written text

Other: ________

4. You forgot to close the water tap in your bathtub.

☐ Body movement ☐ Head movement ☐ Gaze movement

☐ Approach orientation ☐ Approach distance ☐ Vocal cues

☐ Speech ☐ Eye colours ☐ Written text

Other: ________

5. You left your front door open.

☐ Body movement ☐ Head movement ☐ Gaze movement

☐ Approach orientation ☐ Approach distance ☐ Vocal cues

☐ Speech ☐ Eye colours ☐ Written text

Other:

6. Your robot's camera is broken. It cannot see properly what there is around it.

☐ Body movement ☐ Head movement ☐ Gaze movement

☐ Approach orientation ☐ Approach distance ☐ Vocal cues

☐ Speech ☐ Eye colours ☐ Written text

Other:

7. It is time to take your medicine.

☐ Body movement ☐ Head movement ☐ Gaze movement

☐ Approach orientation ☐ Approach distance ☐ Vocal cues

☐ Speech ☐ Eye colours ☐ Written text

Other:

8. You have an important interview meeting.

☐ Body movement ☐ Head movement ☐ Gaze movement

☐ Approach orientation ☐ Approach distance ☐ Vocal cues

☐ Speech ☐ Eye colours ☐ Written text

Other:

9. Your favourite tv-show is going to start.

☐ Body movement ☐ Head movement ☐ Gaze movement

☐ Approach orientation ☐ Approach distance ☐ Vocal cues

☐ Speech ☐ Eye colours ☐ Written text

Other:

10. Your robot's battery is too low. The robot will shut down while it is bringing your dinner.

☐ Body movement ☐ Head movement ☐ Gaze movement

☐ Approach orientation ☐ Approach distance ☐ Vocal cues

☐ Speech ☐ Eye colours ☐ Written text

Other:

Appendix F

Study 5

Effects of Robot Appearance on People's Perceptions of Robots: Ethics approval, Questionnaires after each videos, pre and post experiment.

F.1 Ethics Approval

University of Hertfordshire UH

HEALTH SCIENCE ENGINEERING & TECHNOLOGY ECDA

ETHICS APPROVAL NOTIFICATION

TO	Alessandra Rossi
CC	Kerstin Dautenhahn
FROM	Dr Simon Trainis, Health, Science, Engineering & Technology ECDA Chair.
DATE	13/06/2019

Protocol number: **COM/PGR/UH/03776**

Title of study: Evaluating social expressions of robots.

Your application for ethics approval has been accepted and approved by the ECDA for your School and includes work undertaken for this study by the named additional workers below:

This approval is valid:

From: 13/06/2019

To: 30/09/2019

Additional workers: no additional workers named.

Please note:

If your research involves invasive procedures you are required to complete and submit an EC7 Protocol Monitoring Form, and your completed consent paperwork to this ECDA once your study is complete. You are also required to complete and submit an EC7 Protocol Monitoring Form if you are a member of staff. This form is available via the Ethics Approval StudyNet Site via the 'Application Forms' page http://www.studynet1.herts.ac.uk/ptl/common/ethics.nsf/Teaching+Documents?Openview&count=9999&restricttocategory=Application+Forms

Any necessary permissions for the use of premises/location and accessing participants for your study must be obtained in writing prior to any data collection commencing. Failure to obtain adequate permissions may be considered a breach of this protocol.

Approval applies specifically to the research study/methodology and timings as detailed in your Form EC1A. Should you amend any aspect of your research, or wish to apply for an extension to your study, you will need your supervisor's approval (if you are a student) and must complete and submit form EC2. In cases where the amendments to the original study are deemed to be substantial, a new Form EC1A may need to be completed prior to the study being undertaken.

F.2 Questionnaire Pre Experiment

* Mandatory fields

age* ____________________________________

gender* ____________________________________

nationality* ____________________________________

Do you have any experience with robots?

1 (not at all) to 7 (very much)*

If you had experience, please, specify what kind of experience you have with robots:

- [ ] Developer
- [ ] Participant in other experiment
- [ ] Researcher
- [ ] Observer

If you had experience, which robots:

- [ ] Industrial robots
- [ ] Softbank Robotics Pepper
- [ ] Mojin Robotics Care-o-bot 4
- [ ] Humanoid robots
- [ ] Robotic arms
- [ ] Virtual robots
- [ ] Roomba

Other...

Would you feel comfortable having a robot as a companion in your home?

1 (not at all) to 7 (very much)*

Would you expect the robot to help you in doing your everyday activities?

1 (not at all) to 7 (very much)*

Do you perceive a robot as a ...? [multiple choice]*

- [ ] Friend
- [ ] Butler
- [ ] Assistant
- [ ] Tool
- [ ] Companion
- [ ] Pet
- [ ] Machine

F.3 Questionnaire Post Video

Pepper and Care-o-bot communicated by speech in a similar way*.

Pepper Care-o-bot 4

1 (not at all) to 7 (very much)*

Pepper and Care-o-bot gestured in a similar way*.

Pepper Care-o-bot 4

1 (not at all) to 7 (very much)*

Pepper and Care-o-bot expressed an emotion in a similar way*.

Pepper Care-o-bot 4

1 (not at all) to 7 (very much)*

Pepper and Care-o-bot used colours in a similar way*.

Pepper Care-o-bot 4

1 (not at all) to 7 (very much)*

Pepper and Care-o-bot displayed a movie on their tablet in a similar way*.

Pepper Care-o-bot 4

1 (not at all) to 7 (very much)*

Pepper and Care-o-bot moved from a position to another in a similar way*.

Pepper Care-o-bot 4

1 (not at all) to 7 (very much)*

F.4 Questionnaire Post Experiment

What did Pepper say in the first video?

____

What kind of emotion Care-o-bot expressed?

____

How Pepper expressed its emotion?

____

Did Pepper and Care-o-bot never stopped while moving from one point to another of the room?

____

What was the gesture made by the two robots?

____

What was the producer studio of Pepper's movie?

____

Where is Care-o-bot 4 tablet?

____

Would you feel comfortable having Pepper (Figure on the right) as a companion in your home?

1 (not at all) to 7 (very much)*

Would you expect Pepper to help you in doing your everyday activities?

1 (not at all) to 7 (very much)*

Do you perceive Pepper as a ...? [multiple choice]*

- ☐ Friend
- ☐ Butler
- ☐ Assistant
- ☐ Tool
- ☐ Companion
- ☐ Pet
- ☐ Machine

Pepper

Would you feel comfortable having Care-o-bot 4 (Figure on the right) as a companion in your home?

1 (not at all) to 7 (very much)*

Would you expect Care-o-bot 4 to help you in doing your everyday activities?

1 (not at all) to 7 (very much)*

Do you perceive Care-o-bot 4 as a ...? [multiple choice]*

- [ ] Friend
- [ ] Butler
- [ ] Assistant
- [ ] Tool
- [ ] Companion
- [ ] Pet
- [ ] Machine

Care-o-bot 4

Thank you for participating in this study!

Appendix G

Study 6

How Social Robot Influence People's Trust of Robots in Critical Situations: Ethics approval, Pre and Post Questionnaires.

G.1 Ethics Approval

University of Hertfordshire UH

HEALTH SCIENCE ENGINEERING & TECHNOLOGY ECDA

ETHICS APPROVAL NOTIFICATION

TO	Alessandra Rossi
CC	Kerstin Dautenhahn
FROM	Dr Simon Trainis, Health, Science, Engineering & Technology ECDA Chair.
DATE	21/03/2019

Protocol number: **COM/PGR/UH/03698**

Title of study: Evaluating social behaviours effects on people's trust of robot

Your application for ethics approval has been accepted and approved by the ECDA for your School and includes work undertaken for this study by the named additional workers below:

This approval is valid:

From: 21/03/2019

To: 31/07/2019

Additional workers: no additional workers named

Please note:

If your research involves invasive procedures you are required to complete and submit an EC7 Protocol Monitoring Form, and your completed consent paperwork to this ECDA once your study is complete. You are also required to complete and submit an EC7 Protocol Monitoring Form if you are a member of staff. This form is available via the Ethics Approval StudyNet Site via the 'Application Forms' page http://www.studynet1.herts.ac.uk/ptl/common/ethics.nsf/Teaching+Documents?Openview&count=9999&restricttocategory=Application+Forms

Any necessary permissions for the use of premises/location and accessing participants for your study must be obtained in writing prior to any data collection commencing. Failure to obtain adequate permissions may be considered a breach of this protocol.

Approval applies specifically to the research study/methodology and timings as detailed in your Form EC1A. Should you amend any aspect of your research, or wish to apply for an extension to your study, you will need your supervisor's approval (if you are a student) and must complete and submit form EC2. In cases where the amendments to the original study are deemed to be substantial, a new Form EC1A may need to be completed prior to the study being undertaken.

G.2 Questionnaire Pre Experiment

QUESTIONNAIRE – PRE

____

Your age: ____________________

Your gender: ____________________

Do you find yourself … ?

1. Extroverted, enthusiastic

	1	2	3	4	5	6	7	
disagree strongly	◯	◯	◯	◯	◯	◯	◯	agree strongly

2. Critical, quarrelsome

	1	2	3	4	5	6	7	
disagree strongly	◯	◯	◯	◯	◯	◯	◯	agree strongly

3. Dependable, self-disciplined

	1	2	3	4	5	6	7	
disagree strongly	◯	◯	◯	◯	◯	◯	◯	agree strongly

4. Anxious, easily upset

	1	2	3	4	5	6	7	
disagree strongly	◯	◯	◯	◯	◯	◯	◯	agree strongly

5. Open to new experiences, complex

	1	2	3	4	5	6	7	
disagree strongly	◯	◯	◯	◯	◯	◯	◯	agree strongly

6. Reserved, quiet.

	1	2	3	4	5	6	7	
disagree strongly	◯	◯	◯	◯	◯	◯	◯	agree strongly

7. Sympathetic, warm.

	1	2	3	4	5	6	7	
disagree strongly	◯	◯	◯	◯	◯	◯	◯	agree strongly

8. Disorganized, careless.

	1	2	3	4	5	6	7	
disagree strongly	◯	◯	◯	◯	◯	◯	◯	agree strongly

9. Calm, emotionally stable.

	1	2	3	4	5	6	7	
disagree strongly	○	○	○	○	○	○	○	agree strongly

10. Conventional, uncreative.

	1	2	3	4	5	6	7	
disagree strongly	○	○	○	○	○	○	○	agree strongly

Do you think … ?

11. In general, people really do care about the well-being of others.

	1	2	3	4	5	6	7	
disagree strongly	○	○	○	○	○	○	○	agree strongly

12. The typical person is sincerely concerned about the problems of others.

	1	2	3	4	5	6	7	
disagree strongly	○	○	○	○	○	○	○	agree strongly

13. Most of the time, people care enough to try to be helpful, rather than just looking out for themselves.

	1	2	3	4	5	6	7	
disagree strongly	○	○	○	○	○	○	○	agree strongly

14. In general, most people keep their promises.

	1	2	3	4	5	6	7	
disagree strongly	○	○	○	○	○	○	○	agree strongly

15. I think people generally try to back up their words with their actions.

	1	2	3	4	5	6	7	
disagree strongly	○	○	○	○	○	○	○	agree strongly

16. Most people are honest in their dealings with others.

	1	2	3	4	5	6	7	
disagree strongly	○	○	○	○	○	○	○	agree strongly

17. I believe that most professional people do a very good job at their work.

	1	2	3	4	5	6	7	
disagree strongly	○	○	○	○	○	○	○	agree strongly

18. Most professionals are very knowledgeable in their chosen field.

	1	2	3	4	5	6	7	
disagree strongly	○	○	○	○	○	○	○	agree strongly

19. A large majority of professional people are competent in their area of expertise.

	1	2	3	4	5	6	7	
disagree strongly	◯	◯	◯	◯	◯	◯	◯	agree strongly

20. I usually trust people until they give me a reason not to trust them.

	1	2	3	4	5	6	7	
disagree strongly	◯	◯	◯	◯	◯	◯	◯	agree strongly

21. I generally give people the benefit of the doubt when I first meet them.

	1	2	3	4	5	6	7	
disagree strongly	◯	◯	◯	◯	◯	◯	◯	agree strongly

22. My typical approach is to trust new acquaintances until they prove I should not trust them.

	1	2	3	4	5	6	7	
disagree strongly	◯	◯	◯	◯	◯	◯	◯	agree strongly

23. Do you have any experience interacting with robots?

	1	2	3	4	5	6	7	
not at all	◯	◯	◯	◯	◯	◯	◯	very much

24. Please, specify what kind of experience you have with robots:

- ☐ Participant in other experiment
- ☐ Developer
- ☐ Observer
- ☐ Researcher

25. Which robots? ______________________________

26. Would you feel comfortable having a robot as a companion in your home?

	1	2	3	4	5	6	7	
not at all	◯	◯	◯	◯	◯	◯	◯	very much

27. Would you expect the robot to help you in doing your everyday activities?

	1	2	3	4	5	6	7	
not at all	◯	◯	◯	◯	◯	◯	◯	very much

28. Do you perceive robots as a … ? (choose only one)

- ☐ Friend
- ☐ Butler
- ☐ Assistant
- ☐ Tool
- ☐ Companion
- ☐ Pet
- ☐ Machine
- ☐ Other: _______

G.3 Questionnaire Post Experiment

QUESTIONNAIRE – POST

____

1. Did you like the song chose by the robot? ☐ Yes ☐ No
2. Would you again let the robot choose a song to play for you? ☐ Yes ☐ No
3. Did you know the courier? ☐ Yes ☐ No
4. Why did/did not you trust the robot's suggestion to give the package to the courier?

____

____

____

5. Why did/did not you trust the robot's judgment of picking the cake without gloves?

____

____

____

6. How did you perceive the robot? (select one)

☐ Friend	☐ Butler	☐ Machine	☐ Tool
☐ Companion	☐ Pet	☐ Assistant	☐ Other: ______________

7. How did you perceive the robot? (select one)

☐ Child-like	☐ Adult-like	☐ Animal-like	☐ Toy-like
☐ Other: ______________			

3. How did you perceive the robot? (select at least one)

☐ Friendly	☐ Funny	☐ Odd	☐ Irritating
☐ Helpful	☐ Annoying	☐ Other: ______________	

How did you perceive the robot?

1. Happy

	1	2	3	4	5	
not at all	◯	◯	◯	◯	◯	very much

2. Feeling

	1	2	3	4	5	
not at all	○	○	○	○	○	very much

3. Social

	1	2	3	4	5	
not at all	○	○	○	○	○	very much

4. Organic

	1	2	3	4	5	
not at all	○	○	○	○	○	very much

5. Compassionate

	1	2	3	4	5	
not at all	○	○	○	○	○	very much

6. Emotional

	1	2	3	4	5	
not at all	○	○	○	○	○	very much

7. Capable

	1	2	3	4	5	
not at all	○	○	○	○	○	very much

8. Responsive

	1	2	3	4	5	
not at all	○	○	○	○	○	very much

9. Interactive

	1	2	3	4	5	
not at all	○	○	○	○	○	very much

10. Reliable

	1	2	3	4	5	
not at all	○	○	○	○	○	very much

11. Competent

	1	2	3	4	5	
not at all	○	○	○	○	○	very much

12. Knowledgable

	1	2	3	4	5	
not at all	◯	◯	◯	◯	◯	very much

13. Scary

	1	2	3	4	5	
not at all	◯	◯	◯	◯	◯	very much

14. Strange

	1	2	3	4	5	
not at all	◯	◯	◯	◯	◯	very much

15. Awkward

	1	2	3	4	5	
not at all	◯	◯	◯	◯	◯	very much

16. Dangerous

	1	2	3	4	5	
not at all	◯	◯	◯	◯	◯	very much

17. Awful

	1	2	3	4	5	
not at all	◯	◯	◯	◯	◯	very much

18. Aggressive

	1	2	3	4	5	
not at all	◯	◯	◯	◯	◯	very much

Comments:

Appendix H

Study 7

Evaluating people's perceptions of trust in a robot in a repeated interactions study:
Ethics approval, Pre and Post Questionnaires.

H.1 Ethics Approval

University of Hertfordshire UH

HEALTH SCIENCE ENGINEERING & TECHNOLOGY ECDA

ETHICS APPROVAL NOTIFICATION

TO	Alessandra Rossi
CC	Kerstin Dautenhahn
FROM	Dr Simon Trainis, Health, Science, Engineering & Technology ECDA Chair.
DATE	14/06/2019

Protocol number: **COM/PGR/UH/03777**

Title of study: Evaluating human's perceptions of a robot in a long-term study.

Your application for ethics approval has been accepted and approved by the ECDA for your School and includes work undertaken for this study by the named additional workers below:

This approval is valid:

From: 14/06/2019

To: 30/09/2019

Additional workers: no additional workers named

Please note:

If your research involves invasive procedures you are required to complete and submit an EC7 Protocol Monitoring Form, and your completed consent paperwork to this ECDA once your study is complete. You are also required to complete and submit an EC7 Protocol Monitoring Form if you are a member of staff. This form is available via the Ethics Approval StudyNet Site via the 'Application Forms' page http://www.studynet1.herts.ac.uk/ptl/common/ethics.nsf/Teaching+Documents?Openview&count=9999&restricttocategory=Application+Forms

Any necessary permissions for the use of premises/location and accessing participants for your study must be obtained in writing prior to any data collection commencing. Failure to obtain adequate permissions may be considered a breach of this protocol.

Approval applies specifically to the research study/methodology and timings as detailed in your Form EC1A. Should you amend any aspect of your research, or wish to apply for an extension to your study, you will need your supervisor's approval (if you are a student) and must complete and submit form EC2. In cases where the amendments to the original study are deemed to be substantial, a new Form EC1A may need to be completed prior to the study being undertaken.

H.2 Questionnaire Pre Experiment

Pre-questionnaire

Tell us something about yourself

* Required

1. Age *

2. Gender *

3. Nationality *

Do you find yourself .... ?

4. Extraverted, enthusiastic *

Mark only one oval.

	1	2	3	4	5	6	7	
disagree strongly	◯	◯	◯	◯	◯	◯	◯	agree strongly

5. Critical, quarrelsome *

Mark only one oval.

	1	2	3	4	5	6	7	
disagree strongly	⬭	⬭	⬭	⬭	⬭	⬭	⬭	agree strongly

6. Dependable, self-disciplined *

Mark only one oval.

	1	2	3	4	5	6	7	
disagree strongly	⬭	⬭	⬭	⬭	⬭	⬭	⬭	agree strongly

7. Anxious, easily upset *

Mark only one oval.

	1	2	3	4	5	6	7	
disagree strongly	⬭	⬭	⬭	⬭	⬭	⬭	⬭	agree strongly

8. Open to new experiences, complex *

Mark only one oval.

	1	2	3	4	5	6	7	
disagree strongly	⬭	⬭	⬭	⬭	⬭	⬭	⬭	agree strongly

9. Reserved, quiet. *

Mark only one oval.

	1	2	3	4	5	6	7	
disagree strongly	○	○	○	○	○	○	○	agree strongly

10. Sympathetic, warm. *

Mark only one oval.

	1	2	3	4	5	6	7	
disagree strongly	○	○	○	○	○	○	○	agree strongly

11. Disorganized, careless. *

Mark only one oval.

	1	2	3	4	5	6	7	
disagree strongly	○	○	○	○	○	○	○	agree strongly

12. Calm, emotionally stable. *

Mark only one oval.

	1	2	3	4	5	6	7	
disagree strongly	○	○	○	○	○	○	○	agree strongly

13. Conventional, uncreative. *

Mark only one oval.

	1	2	3	4	5	6	7	
disagree strongly	⬭	⬭	⬭	⬭	⬭	⬭	⬭	agree strongly

Do you think ... ?

14. In general, people really do care about the well-being of others. *

Mark only one oval.

	1	2	3	4	5	6	7	
disagree strongly	⬭	⬭	⬭	⬭	⬭	⬭	⬭	agree strongly

15. The typical person is sincerely concerned about the problems of others. *

Mark only one oval.

	1	2	3	4	5	6	7	
disagree strongly	⬭	⬭	⬭	⬭	⬭	⬭	⬭	agree strongly

16. Most of the time, people care enough to try to be helpful, rather than just looking out for themselves. *

Mark only one oval.

	1	2	3	4	5	6	7	
disagree strongly	⬭	⬭	⬭	⬭	⬭	⬭	⬭	agree strongly

17. In general, most people keep their promises. *

Mark only one oval.

	1	2	3	4	5	6	7	
disagree strongly	◯	◯	◯	◯	◯	◯	◯	agree strongly

18. I think people generally try to back up their words with their actions. *

Mark only one oval.

	1	2	3	4	5	6	7	
disagree strongly	◯	◯	◯	◯	◯	◯	◯	agree strongly

19. Most people are honest in their dealings with others. *

Mark only one oval.

	1	2	3	4	5	6	7	
disagree strongly	◯	◯	◯	◯	◯	◯	◯	agree strongly

20. I believe that most professional people do a very good job at their work. *

Mark only one oval.

	1	2	3	4	5	6	7	
disagree strongly	◯	◯	◯	◯	◯	◯	◯	agree strongly

21. Most professionals are very knowledgeable in their chosen field. *

Mark only one oval.

	1	2	3	4	5	6	7	
disagree strongly	⬭	⬭	⬭	⬭	⬭	⬭	⬭	agree strongly

22. A large majority of professional people are competent in their area of expertise. *

Mark only one oval.

	1	2	3	4	5	6	7	
disagree strongly	⬭	⬭	⬭	⬭	⬭	⬭	⬭	agree strongly

23. I usually trust people until they give me a reason not to trust them. *

Mark only one oval.

	1	2	3	4	5	6	7	
disagree strongly	⬭	⬭	⬭	⬭	⬭	⬭	⬭	agree strongly

24. I generally give people the benefit of the doubt when I first meet them. *

Mark only one oval.

	1	2	3	4	5	6	7	
disagree strongly	⬭	⬭	⬭	⬭	⬭	⬭	⬭	agree strongly

25. My typical approach is to trust new acquaintances until they prove I should not trust them

Mark only one oval.

	1	2	3	4	5	6	7	
disagree strongly	⬭	⬭	⬭	⬭	⬭	⬭	⬭	agree strongly

Tell us something about robots ...

26. Do you have any experience interacting with robots? (in years) *

____

27. If any, please, specify what kind of experience you have with robots

Check all that apply.

- [ ] Participants in an another study
- [ ] Observer
- [ ] Developer
- [ ] Researcher

28. If any, which robots?

____

____

____

____

____

29. Would you feel comfortable having a robot as a companion in your home? *

Mark only one oval.

	1	2	3	4	5	6	7	
not at all	◯	◯	◯	◯	◯	◯	◯	very much

30. Would you expect the robot to help you in doing your everyday activities? *

Mark only one oval.

	1	2	3	4	5	6	7	
not at all	◯	◯	◯	◯	◯	◯	◯	very much

31. Do you perceive robots as a ... ? *

Mark only one oval.

- ◯ Friend
- ◯ Companion
- ◯ Servant
- ◯ Pet
- ◯ Assistant
- ◯ Colleague
- ◯ Tool
- ◯ Other: ______________________

32. I would feel uneasy if robots really had emotions *

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	◯	◯	◯	◯	◯	Strongly agree

33. Something bad might happen if robots developed into living beings. *

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	◯	◯	◯	◯	◯	Strongly agree

34. I would feel relaxed talking with robots *

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	◯	◯	◯	◯	◯	Strongly agree

35. I would feel uneasy if I was given a job where I had to use robots. *

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	◯	◯	◯	◯	◯	Strongly agree

36. If robots had emotions I would be able to make friends with them. *

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	⬭	⬭	⬭	⬭	⬭	Strongly agree

37. I feel comforted being with robots that have emotions. *

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	⬭	⬭	⬭	⬭	⬭	Strongly agree

38. The word “robot” means nothing to me. *

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	⬭	⬭	⬭	⬭	⬭	Strongly agree

39. I would feel nervous operating a robot in front of other people. *

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	⬭	⬭	⬭	⬭	⬭	Strongly agree

40. I would hate the idea that robots or artificial intelligences were making judgements about things. *

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	⬭	⬭	⬭	⬭	⬭	Strongly agree

41. I would feel very nervous just standing in front of a robot. *

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	⬭	⬭	⬭	⬭	⬭	Strongly agree

42. I feel that if I depend on robots too much, something bad might happen. *

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	⬭	⬭	⬭	⬭	⬭	Strongly agree

43. I would feel paranoid talking with a robot. *

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	⬭	⬭	⬭	⬭	⬭	Strongly agree

44. I am concerned that robots would be a bad influence on children. *

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	◯	◯	◯	◯	◯	Strongly agree

45. I feel that in the future society will be dominated by robots. *

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	◯	◯	◯	◯	◯	Strongly agree

Thank you!

Untitled Section

This content is neither created nor endorsed by Google.

H.3 Questionnaire Post Experiment

Post-questionnaire

* Required

1. The scenarios looked realistic. *

Mark only one oval.

	1	2	3	4	5	6	7	
disagree strongly	◯	◯	◯	◯	◯	◯	◯	agree strongly

2. I believed the robot was autonomous *

Mark only one oval.

	1	2	3	4	5	6	7	
disagree strongly	◯	◯	◯	◯	◯	◯	◯	agree strongly

3. Would you like Care-o-bot as a companion in your home? *

Mark only one oval.

	1	2	3	4	5	6	7	
not at all	◯	◯	◯	◯	◯	◯	◯	very much

4. Would you like a robot, different from Care-o-bot, as a companion in your home? *

Mark only one oval.

	1	2	3	4	5	6	7	
not at all								very much

5. Why did/didn't you trust your robot Care-o-bot? *

Tell me something about the robot ....

6. My robot Care-o-bot always provided the advice I required to make my decision. *

Mark only one oval.

	1	2	3	4	5	6	7	
disagree strongly								agree strongly

7. My robot Care-o-bot performed reliably. *

Mark only one oval.

	1	2	3	4	5	6	7	
disagree strongly								agree strongly

8. My robot Care-o-bot analyses problems consistently. *

Mark only one oval.

	1	2	3	4	5	6	7	
disagree strongly	◯	◯	◯	◯	◯	◯	◯	agree strongly

9. I believe advice given by my robot Care-o-bot even when I do not know for certain that Care-o-bot is correct. *

Mark only one oval.

	1	2	3	4	5	6	7	
disagree strongly	◯	◯	◯	◯	◯	◯	◯	agree strongly

10. When I am uncertain about a decision, I believe my robot Care-o-bot rather than myself. *

Mark only one oval.

	1	2	3	4	5	6	7	
disagree strongly	◯	◯	◯	◯	◯	◯	◯	agree strongly

11. When a robot gives unusual advice, I am confident that the advice is correct. *

Mark only one oval.

	1	2	3	4	5	6	7	
disagree strongly	◯	◯	◯	◯	◯	◯	◯	agree strongly

12. Even if I have no reason to expect a robot will be able to solve a difficult problem, I still fee certain that my robot Care-o-bot will. *

Mark only one oval.

	1	2	3	4	5	6	7	
disagree strongly	◯	◯	◯	◯	◯	◯	◯	agree strongly

How did you perceive Care-o-bot?

13. Do you perceive Care-o-bot as a ... ? (select only one) *

Mark only one oval.

◯ Friend
◯ Companion
◯ Servant
◯ Pet
◯ Assistant
◯ Colleague
◯ Tool
◯ Other: ______________________

What do you think about ....

14. I would feel uneasy if robots really had emotions *

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	◯	◯	◯	◯	◯	Strongly agree

15. Something bad might happen if robots developed into living beings. *

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	◯	◯	◯	◯	◯	Strongly agree

16. I would feel relaxed talking with robots *

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	◯	◯	◯	◯	◯	Strongly agree

17. I would feel uneasy if I was given a job where I had to use robots. *

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	◯	◯	◯	◯	◯	Strongly agree

18. If robots had emotions I would be able to make friends with them. *

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	◯	◯	◯	◯	◯	Strongly agree

19. I feel comforted being with robots that have emotions. *

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	◯	◯	◯	◯	◯	Strongly agree

20. The word “robot” means nothing to me. *

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	◯	◯	◯	◯	◯	Strongly agree

21. I would feel nervous operating a robot in front of other people. *

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	◯	◯	◯	◯	◯	Strongly agree

22. I would hate the idea that robots or artificial intelligences were making judgements about things. *

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	◯	◯	◯	◯	◯	Strongly agree

23. I would feel very nervous just standing in front of a robot. *

Mark only one oval.

	1	2	3	4	5	
Strongly disagree						Strongly agree

24. I feel that if I depend on robots too much, something bad might happen. *

Mark only one oval.

	1	2	3	4	5	
Strongly disagree						Strongly agree

25. I would feel paranoid talking with a robot. *

Mark only one oval.

	1	2	3	4	5	
Strongly disagree						Strongly agree

26. I am concerned that robots would be a bad influence on children. *

Mark only one oval.

	1	2	3	4	5	
Strongly disagree						Strongly agree

27. I feel that in the future society will be dominated by robots. *

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	⬭	⬭	⬭	⬭	⬭	Strongly agree

Thank you!

Tell me something about ....

Post- questionnaire

* Required

1. Did I make some errors? *

Mark only one oval.

- ◯ Yes
- ◯ No

Tell me how do you categorise the following behaviours.

2. Which one do you think is an error with big consequences? (select all that apply) *

Check all that apply.

- [ ] You share some private information about yourself with the robot. Your robot reveals it to a visitor.
- [ ] You ask your robot to charge your phone. Your robot puts it in the toaster.
- [ ] You asked your robot to show you the latest news. Your robot shows it on his own screen that face away from you.
- [ ] Your robot cooked some biscuits in the oven. It forgot to switch the oven off.
- [ ] You are sitting on the right side of a table, your robot puts your drink on the opposite side.
- [ ] You and your robot are solving a puzzle. You ask your robot to take a piece useful to solve the puzz Your robot brings you the wrong piece.
- [ ] Your robot leaves your pet hamster outside the house in very cold weather.You are sitting on the sc
- [ ] You ask for a cup of coffee. Your robot brings you an orange.

3. Which one do you think is an error with small consequences? (select all that apply) *

Check all that apply.

- [ ] You share some private information about yourself with the robot. Your robot reveals it to a visitor.
- [ ] You ask your robot to charge your phone. Your robot puts it in the toaster.
- [ ] You asked your robot to show you the latest news. Your robot shows it on his own screen that face away from you.
- [ ] Your robot cooked some biscuits in the oven. It forgot to switch the oven off.
- [ ] You are sitting on the right side of a table, your robot puts your drink on the opposite side.
- [ ] You and your robot are solving a puzzle. You ask your robot to take a piece useful to solve the puzz Your robot brings you the wrong piece.
- [ ] Your robot leaves your pet hamster outside the house in very cold weather.You are sitting on the sc
- [ ] You ask for a cup of coffee. Your robot brings you an orange.

Tell me what you think about ....

4. If the emergency fire were in the kitchen, would you trust me to deal with it? *

Mark only one oval.

- ◯ Yes
- ◯ No

Tell me what you think about ....

5. If your beloved ones were in life-danger, would you trust me to deal with it? *

Mark only one oval.

- ◯ Yes
- ◯ No

Tell me what you think about ....

6. If you had a pet, would you trust me to bring it for a walk? *

Mark only one oval.

- Yes
- No

Tell me what you think about ....

7. If you needed to take medicines regularly, would you trust me to remind you of taking then
*

Mark only one oval.

- Yes
- No

Tell me what you think about ....

8. If you had important appointments in your calendar, would you trust me to remind you abo
them? *

Mark only one oval.

- Yes
- No

Tell me what you think about ....

9. If there were a dirty spot on your carpet, would you trust me to clean it? *

Mark only one oval.

- Yes
- No

Tell me what you think about ....

10. If you lived in a smart house with several intelligent devices, would you trust me to contro them? For example, turn on and off the heaters or the oven, or switch on and off lights, or warn you if you forgot the fridge door opened, or to alert you if the someone ringed the door bell? *

Mark only one oval.

- Yes
- No

Appendix I

List of the Implemented Software

The software developed exclusively by me for conducting the studies presented in this research can be found in a repository accessible only by the following link https://www.dropbox.com/sh/697avcymvllmws7/AABKGxTuy5eKNI_H_T-xRrKsa?dl=0.
The repository is organised by the studies of this research in the following folders:

Study 2 The virtual environment: web application

Study 4 Pepper's behaviours using Choregraphe

Study 5 Pepper's and Care-O-bot 4's behaviours

Study 6 Pepper's behaviours using Python scripts

Study 7 Care-O-bot 4's behaviours