Linköping University | Department of Computer and Information Science
Master thesis, 30 ECTS | Cognitive Science
Spring term 2019 | LIU-IDA/KOGVET-A--19/006--SE

Can You Read My Mind?

A Participatory Design Study of How a Humanoid Robot Can Communicate Its Intent and Awareness

Sofia Thunberg

Supervisor : Tom Ziemke
Examiner : Arne Jönsson

LINKÖPING UNIVERSITY

Linköpings universitet
SE–581 83 Linköping
+46 13 28 10 00 , www.liu.se

Abstract

Communication between humans and interactive robots will benefit if people have a clear mental model of the robots' intent and awareness. The aim with this thesis was to investigate how human-robot interaction is affected by manipulation of social cues on the robot. The research questions were: How do social cues affect mental models of the Pepper robot, and how can a participatory design method be used for investigating how the Pepper robot could communicate intent and awareness? The hypothesis for the second question was that nonverbal cues would be preferred over verbal cues. An existing standard platform was used, Softbank's Pepper, as well as state-of-the-art tasks from the RoboCup@Home challenge. The rule book and observations from the 2018 competition were thematically coded and the themes created eight scenarios. A participatory design method called PICTIVE was used in a design study, where five student participants went through three phases, label, sketch and interview, to create a design for how the robot should communicate intent and awareness. The use of PICTIVE was a suitable way to extract a lot of design ideas. However, not all scenarios were optimal for the task. The design study confirmed the use of mediating physical attributes to alter the mental model of a humanoid robot to reach common ground. Further, it did not confirm the hypothesis that nonverbal cues would be preferred over verbal cues, though it did show that verbal cues would not be enough. This, however, needs to be further tested in live interactions.

Acknowledgements

I would like to thank my supervisor, Tom Ziemke, for your motivating high expectations on me and for giving me the opportunity to further investigate the field that interests me the most. I have done all my projects since the bachelor thesis under your command, and now, soon entering my PhD, it is clear to me that since my first meeting with you, two and a half year ago, I am hooked in social robotics for life.

Also, a big thank you to Sam Thellman, for your endless curiosity and input to my work. I have changed the subject of interest many times during the years, but in the end, you are my biggest inspiration and it is not for nothing that I am now engaging in the same area as you.

Further, I am always grateful to Robert Johansson, that explains to me the signals that the universe tries to send me, always when I need it the most.

And last, but not least, my love, Alex. Thank you for reminding me of the other things in life beside thesis writing.

Now, a new era begins.

Linköping in June 2019

Sofia Thunberg

P.S. Give a hand to my friend Pepper who is always there for me, teaching me the value of patience.

Contents

List of Tables

List of Figures

List of Abbreviations

1 Introduction

Communication between humans and interactive robots will benefit if humans have a clear mental model of what the robot can do (Powers & Kiesler, 2006). Creating a correct first impression of the robot's knowledge and "personality" seems especially important for service robots that will interact with people in public environments. This thesis addresses how a humanoid robot's physical attributes create these important first impressions.

Powers and Kiesler (2006) stress that, ideally, new robot designs should take advantage of the social cues that will automatically be perceived by the robot's attributes. That is, the robot does not have to be explicit about its "personality" and knowledge domains. This way of modelling one's knowledge of others has previously been described by Nickerson (1999), in human-human interaction. Using research in social psychology, and the result of social cognitive studies, the same phenomenon that occurs in human-human interaction can be investigated in human-robot interaction. To use the human as a norm on how the robot should behave and appear, has for a long time been the ultimate goal, although the field has not agreed on which features are optimal, or when it has gone too far (Fong, Nourbakhsh, & Dautenhahn, 2003; Mori, 2012).

> Theoretically, we can create a robot that conveys its personality and expertise without instruction through the physical and interactive design of the robot. Interactive humanoid robots have a good start on creating a strong first impression because by virtue of their movement, appearance, and interactivity, they are more human-like than other computer-based technologies (Powers and Kiesler, 2006, pp.219).

A human-like form elicits automatic similarities of people and interactivity, such as speech and gestures that will prompt observers to anthropomorphize automatically, without intent or thoughtful processing (Scholl & Tremoulet, 2000).

Previous evidence that physical appearance matters on the robot for effecting the mental model and fostering interaction, motivates an examination of how to design such an appearance (Powers & Kiesler, 2006). By using an existing standard platform, Softbank's Pepper (SoftBank Robotics, 2019), and state-of-the-art tasks from the RoboCup@Home challenge (RoboCup Federation, 2019), this thesis aims to investigate the use of participatory design method, as well as how social cues could affect the mental model of the Pepper robot.

1.1 Aim

The aim with this thesis is to investigate how human-robot interaction is affected by manipulation of social cues on the robot. Different design ideas for the robot to communicate intent and awareness will be investigated, with the aim to create common ground between the robot and the human, and for the use of least collective effort for interacting with the robot.

1.2 Research Questions

The research questions are:

- How do social cues affect mental models of the Pepper robot?
- How can a participatory design method be used for investigating how the Pepper robot could communicate intent and awareness?

The hypothesis for the second question is that people would prefer nonverbal cues for communicating intent and awareness, over verbal ones. The motivation for this hypothesis is two fold. The first reason is the result from a previous ethnographic study, conducted ”out in the wild”, with the Pepper robot. The author of this thesis conducted a student project during the fall of 2018 (letting people interact with a Pepper humanoid robot at Stockholm central station), and the result showed that without guidance, people are having a hard time understanding how to communicate with the robot. It is not that common with robots, or machines, with voice assistive systems and people more are used to interacting with systems through a touch screen. In the study people approached the robot and tried to use the tablet, and even though nothing happened when they touched it it did not come natural to them to try and talk to the robot. Therefore it would be interesting to investigate if users would prefer non-verbal cues for the robot to communicate intent and awareness, and if this in the long run could nudge users into understanding that they can talk to the robot.

The second reason is that people express a lot of emotions in nonverbal ways. By using expressions, such as gestures, postures, and movements, people can express some things with greater facility and efficiency than with words (Mehrabian, 1972). Beliefs, hopes and desires can also be expressed with eye gaze, vocalisation and gestures (further explained in Section 2.2). Therefore it would be interesting to investigate if the communication of intent and awareness could be more effective and explicit by the use of nonverbal cues, or with verbal cues supported with nonverbal cues (in opposition to verbal cues alone).

1.3 Limitations

The thesis is limited to work with the Pepper robot. It would be interesting to examine different humanoid robots, as well as other types of robots, such as service types or social companion types, but the limited amount of time available and the novelty of the field speak for a study with only one robot.

1.4 Delimitations

The choice of using the Pepper robot was based on previous research that indicates that the more anthropomorphic, but still machine-like, a robot is the easier it is to trigger social cues in interactions with humans (Kiesler, 2005; Lee, Kiesler, & Lau, 2005). The choice is also motivated by the use case that comes with RoboCup@Home Social Standard Platform League (SSPL), that uses Pepper. The result could be used in the competition for the Swedish team LiU@HomeWreckers (Thunberg & Westin, 2019). With this, the chosen scenarios will be limited to the @Home tasks.

The project is further delimited to investigating the use of robot signals, from the robot to the human, and not the other way around. Inputs could be that the human can push a button on the display or touch the robots head to carry on with a task. This could be useful for the competition but does not support the purpose of interacting with least collective effort (see Section 2.2). The robot signals are, for example, the use of sounds and lights to indicate what the robot is doing.

2 Theoretical Background

In this chapter the concepts of communication, intention and awareness will be defined to some degree, the theories of mental modelling of how one's understanding of the belief, desire and knowledge of specific others will be presented, as well as the method theories, observation and participatory design method.

2.1 Intent and Awareness in Context of This Thesis

First, the thesis requires definitions regarding some concepts that are crucial for this project. Both *intention* and *awareness* are complex concepts. The point here is, however, not to raise any higher philosophical discussions regarding these, but to have an initial rough understanding of what they are, or could be. In short, intention refer to the internal states of the robot, what it *wants*, or *plans*, to do, and awareness refers to what the robot *knows* about the world, through its sensors. In this regard, an intention from the robot could, for example, be that it wants to answer a question or that it wants a person to follow it, and the robot shows awareness when it, for example, knows that a person has entered a room or that it knows that a person is waving for it to come closer.

2.2 How I Know That You Know

To communicate effectively with other people, one needs to have a conception of what the other person knows. According to the *cooperative principle*, or the *Gricean view*, people do not convey to information that can be assumed that they already have (Grice, 1975). For example, a speaker that overestimates what a listener knows may talk over the listeners head, also a speaker that underestimates what a listener knows may be interpreted to be talking down to the listener (Nickerson, 1999). To Nickerson, *knowledge* includes beliefs, opinions, suppositions, attitudes, and related states of mind.

The term *Theory of Mind (ToM)*, and how people develop one, has for a long time been discussed. Ever since Premack and Woodruff (1978) wondered if chimpanzees have it, social cognition studies have focused on figuring out to which extent humans has ToM, and especially how children develop it (Astington, Harris, & Olson, 1988; Astington, 1993). For example, Trevarthen and Hubley (1978) showed that infants during their first year develop the ability to communicate in nonverbal ways with the use of eye gaze, vocalisation and prelinguistic gestures. In short, an individual has ToM if she imputes mental states upon others, such as belief, knowlegde, doubt, thinking, guessing, pretending, liking and so forth (Premack & Woodruff, 1978).

Nickerson argues that one uses one's own knowledge as a basis for creating a model for what other people know. Nickerson says, as a general rule, that a relatively accurate model of what other people know, believe, or feel are preferred over inaccurate ones. However, it would not be hard to find situations where it would be preferable to have an inaccurate one. For example, if another person's thoughts or feelings could threaten the stability of a relationship. According to Sillars, Pike, Jones, and Murphy (1984), there are even some evidence that a high degree of empathic accuracy can work in conflict for the survival of a relationship. Eisenberg, Murphy, and Shepard (1997) suggested that empathy (or empathic accuracy) has both emotional and cognitive aspects, where focus is in knowledge in the conventional sense. According to the authors, the cognitive aspects of empathy, the ability to understand others' internal states, is called perspective taking.

Fussell and Krauss (1989a, 1989b) discovered that people's verbal descriptions of nonsense figures differed depending on who they thought would match the description and the figure later, i.e. themselves, friends or strangers. The authors interpreted the result as that people tend to adapt the message to the background knowledge of the recipient.

As can be seen in Figure 2.1, Nickerson (1999) shows how one constructs a model of a specific others' knowledge by shaping a default model of what an unknown person know, by taking into account the information one has about the other person that makes it different from the default model. *Others* in this context are, according to the author, heterogeneous groups (e.g. people watching the 9 o'clock news), small groups with a shared characteristics (e.g. members of an association), or single individuals (e.g. the cashier at the supermarket).

This can be seen as a case of the general reasoning heuristic of *anchoring* and *adjustment* (Tversky & Kahneman, 1974). According to this, people make judgements by starting with an anchor as a point of departure and then make adjustments to it. Nickerson (1999) means that when people are given an anchor they typically adjust their judgement in the right direction, but that they often overestimate. The model takes this in account and predicts that one will overestimate what unknown others know.

The type of mental model suggested by Nickerson (1999) is not only interesting in Human-Human Interaction (HHI). Would humans' mental model of a robot adapt in the same way from their previous knowledge of what the robot knows? According to Powers and Kiesler (2006), communication between humans and interactive robots will benefit if humans have a clear mental model of what the robot can do. Creating a correct first impression of the robot's knowledge and personality seems especially important.

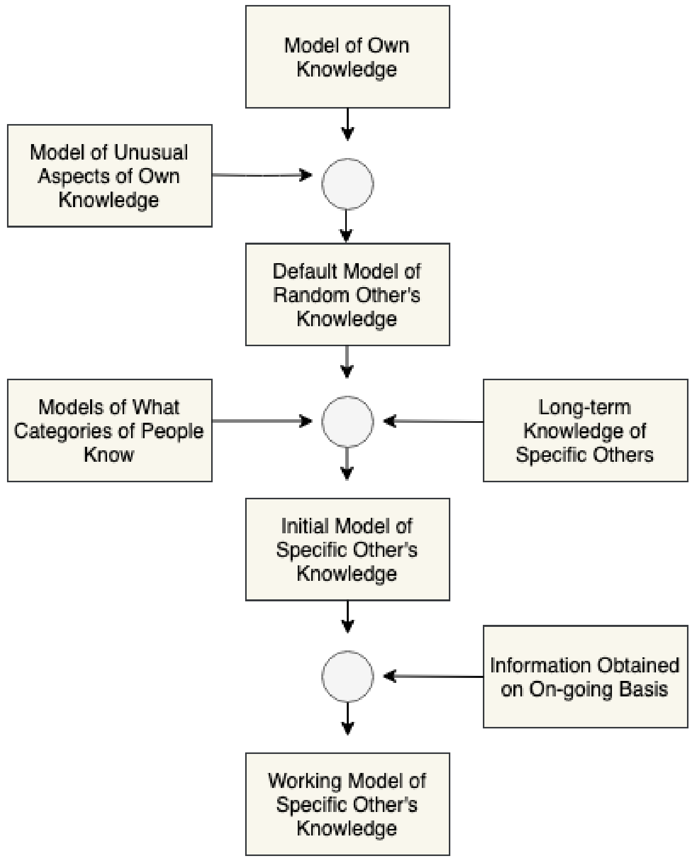


Figure 2.1: A working model of specific others' knowledge from Nickerson (1999).

Nass and Moon (2000) did a review of experimental studies that gave evidence that people mindlessly apply social rules to computers, as well as give them expectations. Nass and Lee (2001) also showed that people apply stereotypes and social heuristics automatically to interactive systems. There is also evidence for this behaviour in HHI (Bargh, Chen, & Burrows, 1996). According to Powers and Kiesler (2006) this is an underlying bottom-up process that occurs immediately and is unconscious. Powers and Kiesler think that there is a parallel cognitive process to this, which uses structural mapping. For example, when a person sees a smiling robot they retrieve information from the long-term memory, in the same way as with metaphors. They might associate the smiling robot with a happy person (appearance similarity) or a playful task (analogistic reasoning), and a mental model of the robot emerges as a persona or prototype. The person might interpret the smiling robot as a machine, in the nonsocial category, and as a nice person, in the social category. Together this creates a mental model of a sociable robot. The authors' experimental result show evidence that people create a mental model of the robot within the first two minutes of interacting, and that the robots physical attributes has an impact of an individual's mental model that can change over time.

In 1991, Clark and Brennan proposed the *Common Ground Theory*. This theory was developed to understand communication between people. According to Kiesler (2005), common ground's main assumption is that communication between people requires coordination to reach mutual understanding. To obtain coordination, for example between two interaction partners or two ballroom dancers, the partners have to have a large amount of shared knowledge, which is, common ground. One of the key elements of the theory is *least collaborative effort*, which is that when people interact with each other they try to minimise their collective effort to gain understanding (Clark & Brennan, 1991). Kiesler (2005) believes that there is a need for common ground also in Human-Robot Interaction (HRI) and that robots should have actions plans for creating common ground with people. She believes that the first step for this is for the robot to use social cues to help people creating an appropriate mental model of the robot. The robot should also actively try to correct the mental model and repair damage to the common ground.

In this thesis the mediation model from Baron and Kenny (1986) (see Figure 2.2) is used, as illustrated in Figure 2.3. The physical attributes on the robot (motion, haptic, sounds through speakers and voice synthesis, LEDs in eyes, ears and shoulders as well as text, animation and symbols on the display) can alter the mental model of the robot and effect the individual's belief of the robots intent and awareness. This study also makes use of the model in Figure 2.1. The assumption is that people will overestimate what the robot knows and wants, but with the help of certain physical attributes, functioning as social cues, this view will be altered and the information that is obtained on an on-going basis will have a big impact on the mental model of the robot's knowledge. The goal is to create the common ground principle of least collective effort.

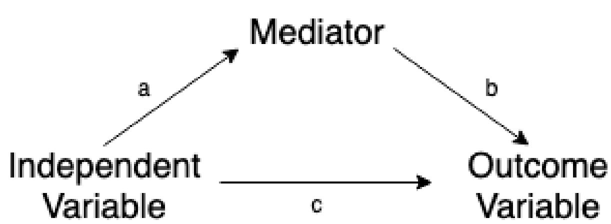


Figure 2.2: Mediation model from Baron and Kenny (1986).

2.3 Method

In this section the theory regarding observation and participatory design will be presented.

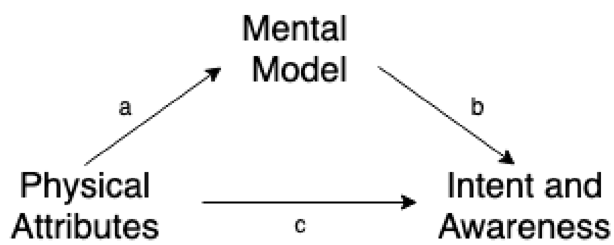


Figure 2.3: Mediation model applied to this study.

2.3.1 Observation

Alm (2017) claims, to conduct ethnographic field work, it is essential that it is conducted out "in the wild", as Hutchins (1995) once coined the term of observing events as they happen outside the laboratory. With this, the observer come close to the objects that are being observed which can create ethical difficulties. For instance, writing academically about a person that has become your friend can feel deceitful, Alm (2017) adds.

Alm says: "To limit one's research to observing people's interaction would be like producing an oldfashioned silent movie" (Alm, 2017, pp.8). With this he means that the researcher does not only observe with her eyes, but also with her ears. It is important to observe what is called "natural talk", things that people would have said even if there was no researcher present.

When observing, it is nearly impossible to make oneself invisible, and therefore a common way of conducting ethnographic studies is *participant observation* (Alm, 2017). This means that the fieldworker is observing and participating at the same time.

If *rich points* appear during observation, they should be followed to discover more than the observer could had imagine in before hand. A rich point is a situation when and where you do not understand how things are related. These can lead to a dead point or to fascinating things you do not understand (Agar, 1996).

Between each session it is preferable to have some time to rest. This is a method, recommended by Emerson, Fretz, and Shaw (1995), that allows for reflection between the observations and time to go through the field notes and go from jottings to a full text with a fresh memory.

The field notes are recommended by Alm (2017) to be coded in the way that he calls *open coding*. Open coding is keywords that you write in the margin of your text, with the aim to mark what is important. The next step is to create *themes* out of the coded material. This means putting keywords together that form themes. After creating these themes one goes through the material again and do a *focused coding*. Focused coding is using more general codes and only those that belongs to one of the themes.

For this thesis observation was the only option. Using interviews instead, for example, would had limited the data, and it would have been hard to capture when and how something on the robot and in the tasks work or does not work. Only relying on what, for example, the members of the competition team say about their robot and how they have solved the task might not give a full picture of how it works live. One reason for this is that sometimes the system works really well in the laboratory environment, but presents itself in new ways in the competitive arena. Another reason is that the members might not be completely honest, since RoboCup@Home is a competition and that everybody tries to win in the end.

2.3.2 Participatory Design

The participatory design method called Plastic Interface for Collaborative Technology Initiative through Video Exploration (PICTIVE) was used (Muller, 1991). In this method, participants reflect on interface for *collaborative technology initiatives* (through video exploration), through the act of putting down ideas on paper and inspecting them. In this way end users have an early exposure to, and provide input about the target implementation technology. Muller means that by using low-tech objects, i.e. non-computer representations of system functionality, all participants are able to contribute their ideas in an easy way. By recording the sessions with video it provides the session leader to be more engaged and not distance herself by for example taking notes. The videos can also provide design documents.

The "P" in PICTIVE stands for "plastic", and by that Muller means that many components of the technique are made of plastic, e.g. coloured pens, that the concepts are very malleable and can change quickly, and also that the result is very artificial since it is not a working system that is being developed but mere, the design ideas.

The objects that are being manipulated during the session are a *shared design surface*, plastic icons, coloured highlighters, coloured pens, labels (data fields), pop-up events and post-it notes. On the shared design surface the participants place the predefined plastic icons and the labels, and they can also create their own with post-it notes. They also have coloured pens and highlighters at their disposal. If the design purpose is a computer system, then the use of pop-up events can come in handy.

The setting for the PICTIVE sessions are in a secluded room with a table, where the session leader is on one and the participant is on the other side of the shared design surface. There is also a video recording next to them that captures all that is happening on the table (Muller, 1991).

The use of PICTIVE, and not another design method, was motivated from a study made by Mahadevan, Somanath, and Sharlin (2018) where they investigated how autonomous cars can communicate intent and awareness. They used PICTIVE as their design method and developed different designs that were later tested "out in the wild".

3 RoboCup@Home SSPL

The RoboCup@Home league is the largest international annual competition for autonomous robots that aims to develop service and assistive robot technology with high relevance for future personal domestic applications. A set of benchmark tasks is used to evaluate the robots' abilities and performance in a realistic non-standardised home environment setting. The RoboCup@Home SSPL uses Softbank's Pepper robot as the standard platform. In Figure 3.1, there is an example of how the arena can look like, and in Figure 3.2, some of the objects that can be included are showed.

Figure 3.1: A typical arena.

Sweden, and Linköping University, has one team, LiU@HomeWreckers (Thunberg & Westin, 2019), that has competed in RoboCup German Open in Magdeburg, Germany, and RoboCup in Montreal, Canada, in 2018. The team is also accepted to the competition round in 2019.

The competition consists of two stages and one final round (see Figure 3.3). To compete in stage two the team has to score in the top teams in stage one,

Figure 3.2: Typical objects.

and the same in stage two to compete in the finals. The tasks are complex and can be solved in several ways, especially by using HRI to make the referees help the robot. If the robot asks a human to do something for them the person has to help them, to a certain limit. Therefore these tasks were chosen to create realistic and useful scenarios.

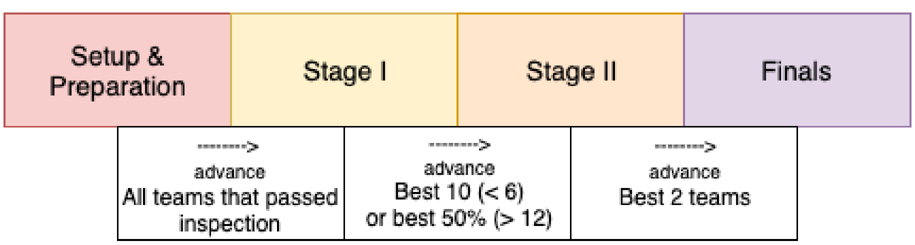


Figure 3.3: Stage system.

In the following sections the tasks will be briefly explained, for full explanation see the RoboCup@Home rulebook (Matamoros, Rascon, Hart, Holz, & van Beek, 2018). But first the Pepper robot will be introduced.

3.1 Pepper Robot

The Softbank's Pepper model used is 1.6, version 2.5.10. (see Figure 3.4). Pepper is 1200 mm tall and weights 28 kg. The robot is equipped with a gyro sensor and a 10.1-inch touch display on its chest and 4 microphones, two RGB cameras, one 2D sensor and three touch sensors on its head. The robot hands have two touch sensors and the base has two sonar sensors, six laser sensors, three bumper sensors and one gyro sensor. It also has LEDs in its eyes, ears and shoulders. The robot also has two speakers, one on each side of the head. The display, LEDs and speakers will mainly be used in this project.

Figure 3.4: The Pepper robot.

3.2 Stage I

In stage I there are four tasks, *Cocktail Party*, *General Purpose Service Robot (GPSR)*, *Help-me-carry* and *Speech and Person Recognition*.

3.2.1 Cocktail Party

In the Cocktail Party task the robot starts by standing outside a door that a referee opens. The robot then navigates to the living room where people are having a party. The robot then looks for humans who want to order a drink. If a person is waving or is sitting down the robot asks that person. If the robot fails to find a human who wants to order it can inform people to approach it if they want to order a drink. The robot then navigates to the barman in the kitchen and places the order, or orders (max three). One of the orders will not be possible because of a lack of ingredients, and then the robot has to remember and find the person placing that order. The robot gives that person three different options and then continues with the new order to the barman.

3.2.2 General Purpose Service Robot

In the GPSR task the robot begins by entering the arena. The referee reads a task for the robot and the robot carry it out. The tasks can be in three different categories. The tasks can be in three different categories.

The first category is tasks with a low difficulty degree, such as ”Tell me the name of the person at the door” and ”Tell me how many beverages are in the shelf”.

The second category is tasks with a moderate difficulty degree, e.g. ”Tell me the name of the woman in the kitchen” and ”Tell me how many beverages in the shelf are red”.

The third category is tasks with a high difficulty degree or with incomplete/erroneous information. For example, ”Bring me a drink from the fridge (There are no drinks in the fridge, but in the kitchen table)” and ”Follow John (John’s location is not specified)”.

3.2.3 Help-me-carry

In the Help-me-carry task the robot starts in the arena and is approached by a referee that tells the robot to follow him/her. The robot follows the person outside of the arena to a car that is full of groceries. When they get to the car the referee tells the robot to go back in and get a person in the kitchen that can help carry the groceries. The robot then navigates back inside and a person crosses the robots path, and another person steps in its way and asks for the time. When finding the person in the kitchen the robot asks him/her to follow the robot back to the car, and then it guides the person.

3.2.4 Speech and Person Recognition

In the Speech and Person Recognition task the robot need to identify unknown persons and answer questions about them and the environment. The robot starts in the middle of the arena and finds a crowd of five to ten people. It announces that it wants to play riddles. The people will then place themselves in a ring around the robot (see Figure 3.5). One person at the time will ask a question and the robot turns toward that person. One of the questions are a predefined question, between one and two are about the arena and its status, between one and two are about the crowd, and between one and two are about the list of official objects in the arena. E.g. ’What day is today?’, ’How many chairs are in the dining room?', 'Number of male or female people' and 'What objects are in the closet?'.

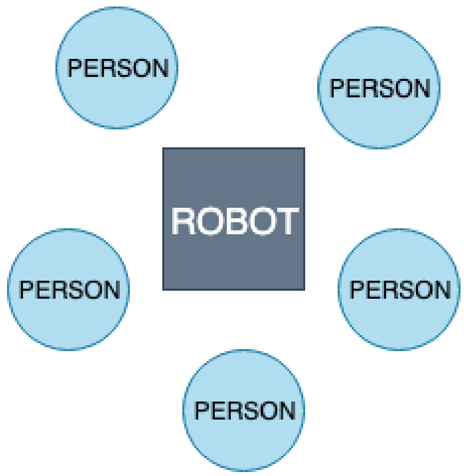


Figure 3.5: Person setup around the robot.

3.3 Stage II

In stage II there are four tasks, *Enhanced Endurance General Purpose Service Robot (EEGSPR)*, *Open Challenge*, *Restaurant* and *Tour Guide*.

3.3.1 EEGSPR

The EEGSPR task begins by the robot entering the arena. The referee reads a task for the robot and the robot carry it out. The tasks can be in three different categories.

The first category is tasks with a low difficulty degree, such as "Bring the chips to Mary at the sofa, tell the time and follow her" and "Go to the kitchen counter, take the coke, and bring it to me".

The second category is tasks with a moderate difficulty degree, e.g. "Offer a beer to all the adults in the living room" and "Tell me what John is doing (John is reading a book)".

The third category is tasks with a high difficulty degree or with incomplete/erroneous information. For example, "Count the snacks in the shelf and tell me how many there are", "Guide Joe here (unspecified where is Joe and how to recognise him)" and "Find James in the living room and guide him to the car (James is in the bedroom)".

3.3.2 Open Challenge

In the Open Challenge task the teams get to show their recent research and the best abilities on the robot. The team captain presents and can show the ability on Pepper and with presentation slides. A jury asks questions afterwards and they judge the task on *Novelty and (scientific) contribution*, *Difficulty level of the demonstrated task*, *Success of the demonstration* and *Overall (demo was convincing, fluent, interesting, etc.)*.

3.3.3 Restaurant

The Restaurant task takes place out in the real world, in a restaurant or a shopping mall (see Figure 3.6 for an example of a restaurant setup). In this tasks there are two robots active at the same time from two teams. The robot start by keeping an *eye* on the tables. When a person asks for the robots attention by waving and calling on it out load. If both robots notice this the barman will decide which of them will get the order. The robot then approaches the person and ask what it can get him/her. On the way to the kitchen one person may step in the robots way. It then places the order to the barman, e.g. "Hamburger with fries for table A and Orange juice for table B". Then comes the delivering phase where the robot should carry a tray with the ordered food and beverages to the correct table.

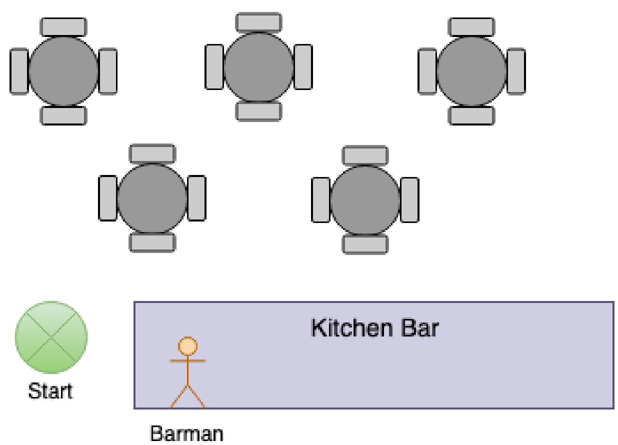


Figure 3.6: Restaurant example setup.

3.3.4 Tour Guide

The Tour Guide task start outside of the arena. The robot search for a large group of people and friendly attract their attention. When it has gathered a group it introduces itself and greet them in a customary way in the country it is in (e.g. shake hand, wave or bow). The robot then asks the group to follow it to the @Home arena. If the people does not follow it then it need to find another group. Inside the arena it asks the group to sit down while it explain how the @Home SSPL works. Afterwards it encourage the group to ask questions.

3.4 Final Round

The Final Round is similar to the Open Challenge, and can be the same presentation as was demonstrated during the Open Challenge. There are two final rounds,the Bronze Competition and the @Home Grand Finale. The jury consists of three groups, the league-internal jury, the league-external jury and the teams-based jury.

The league-internal jury is formed by the Executive Committee and they evaluate scientific contribution, contribution to @Home, relevance for @Home and presentation and performance in the finals.

The league-external jury consists of people not being involved in the RoboCup-@Home league, but having a related background (not necessarily robotics). Their evaluation is based on the originality and presentation (where story-telling is to be rewarded), usability and human-robot interaction, multi-modality and system integration, difficulty and success of the performance as well as relevance and usefulness for daily life.

The teams-based jury is formed by members of the league's teams, and they judge scientific contribution, contribution to @Home, relevance for @Home and novelty of approaches and presentation and performance in the finals.

4 Method

In this chapter, the used method is described. First, the observational data and the coding of the rule book created scenarios. These scenarios were then used in the design study. The order of how the project was conducted can be seen in Figure 4.1.

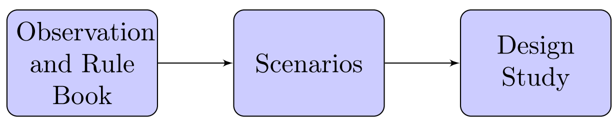


Figure 4.1: Flow chart of the order the project was conducted.

4.1 Observation

The tasks in the RoboCup@Home SSPL were observed during ten occasions at the RoboCup@Home competitions in May and June 2018. The observer was participating, and watched the robot trying to solve the tasks and interacting with people in the arena from the sidelines, or as a refugee in the tasks. The behaviour of the robot and the humans were noted, as well as the level of execution and way of solving the task. There were a couple of hours between each session, and this gave time to rest the mind, sort the thoughts out and write up the field notes.

The field notes from the observation was put together as a full text material. The text was open and focused coded. The codes were then used to identify common problems where the robot had to communicate intent and awareness.

4.2 Scenarios

The RoboCup@Home tasks and rule book, as well as observational data from the competition, were open coded to identify themes. Each theme consist of different situations that the robot has to be able to handle by showing intent and awareness. The themes were then combined into eight scenarios (see Appendix A). During the design study these were presented in Swedish.

4.3 Design Study

To conduct the design study, PICTIVE was used (described in Section 2.3.2). A sketch showing the Pepper robot standing in a typical living room was created (see Figure 4.2), and shown in an A3 paper version. This was the *shared design surface*, which all participants used to create their designs.

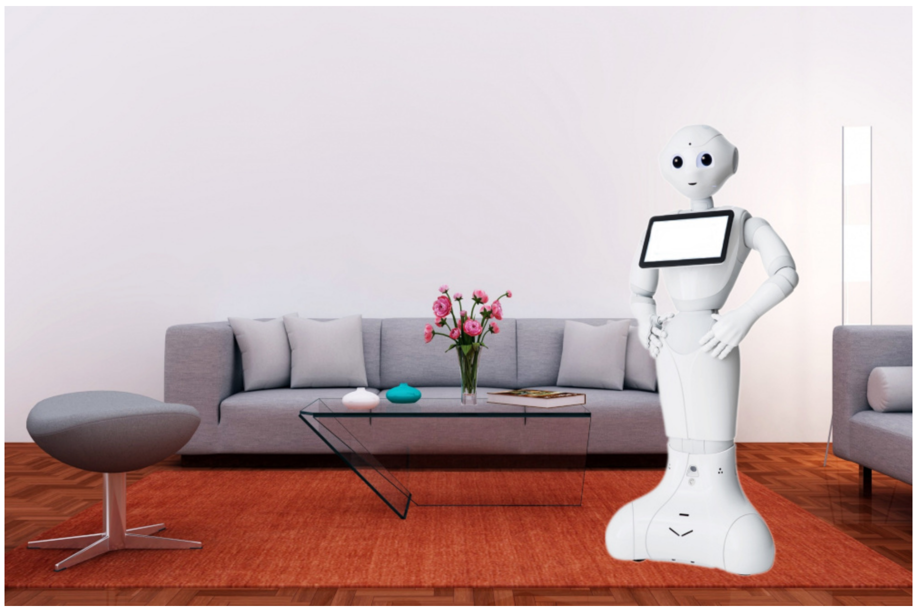


Figure 4.2: The shared design surface.

To facilitate participant's creation of interfaces, ten labels depicting common design cues was created. The labels were inspired by Mahadevan et al. (2018) and were customised after the robots limitations. The final labels were motion, haptic, sounds through speakers and voice synthesis, LEDs in eyes, ears and shoulders as well as text, animation and symbols on the display. The labels were in Swedish during the workshop. The participants were provided with, beside from the labels, coloured pens, stick it notes, glue, pen sharpener, eraser, tape and scissor (see Figure 4.3).

The design study consisted of five sessions. Each session took one hour to one and a half, and were located in a secluded room. One participant and the author of this thesis sat on opposite sides of a table. A Pepper robot was also present to give the participant a clear view of how the different functions could work on the robot in mind. The robot was turned on and the Softbank default mode was on, for example with the tablet running a screen-saver, and blue lights in the ears. Otherwise the robot was silent and did not move. The session started with a short presentation of the robot and the purpose of the study. Data such as age, gender, and experience of autonomous robots, were

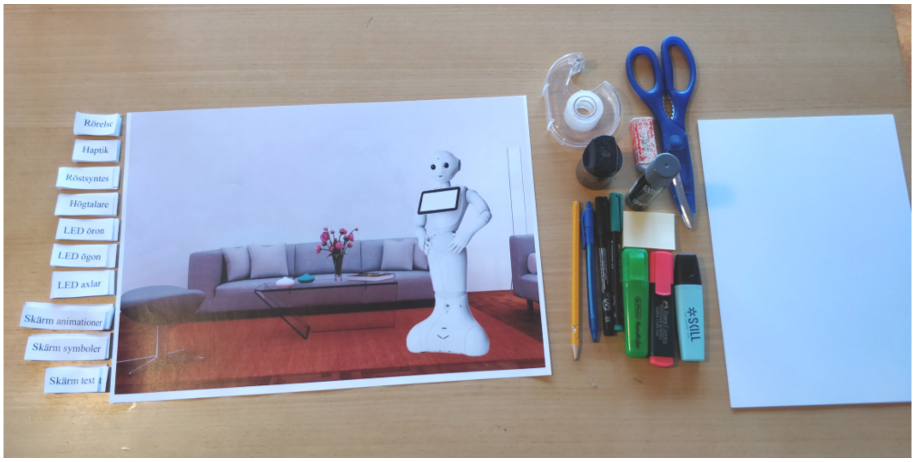


Figure 4.3: The design study setup.

collected, and then the participant's were asked to verbally approve the video recording of the shared design area. They were also informed that the workshop was fully voluntary and that they could quit the session at any time.

The study procedure contained three phases. In the first phase, *the label phase*, the participants were encouraged to use the labels on the shared design space to map different design solutions. To start the creative session, the shared design area was in a living room and the participants' were told an introduction scenario that they could use, but were not limited too. Which was the following:

> ”You enter the home of a friend. Your friend will not be home for a while but you have agreed to wait for him/her. When you enter the living room you see this robot standing there. What would you like it to do, how should it act?”

In the second phase, *the sketch phase*, the participant got a maximum of 30 minutes to create at least three unique interface designs which communicated the awareness and intent of the Pepper robot. The participants were allowed to use the sketch on the shared design space or on white paper. They could use the predefined labels or create their own. Each of the designs could be iterative and build on their earlier designs. They were encouraged to describe their thought process through a think-aloud protocol.

In the third phase, *the interview phase*, the participants were presented with eight diverse scenarios, where the Pepper robot had to communicate its awareness and intent. The participants were asked to rank their designs as a best, fair, or worst fit in handling each scenario, and propose any changes that could improve their interfaces.

Finally, the participants were asked if they had thought about anything else that had not been brought up during the session.

From each study session the participant-created designs and video recordings were collected. The interviews were transcribed and the material were qualitatively coded to identify trends. The think aloud protocol and the sketches were used to create a system design proposal with a combination of common ideas. Finally, the video recording from the interview phase were used to create storyboards of how the scenarios could play out.

4.3.1 Participants

In the design study, five participants (M=26.4 years old, SD=.548, 40% female) participated. They were recruited via social media, through a convenience sample, and were students at Linköping University with different study backgrounds, Teaching, Cognitive Science, Environmental Science, Theoretical Philosophy and Technical Psychics and Engineering. Two of the participants were members of the LiU@HomeWreckers team, these were also the two participants that had previous experience with autonomous robots. The reason to include members of the team was that in this way opinions and ideas from those who know the robot and the competition well could be mixed with participants that has no experience of robots. All participants were fluent in Swedish. There were no compensation for participating.

5 Results

This chapter presents the result from the observations and coding of the rule book in the shape of scenarios, that is used in the design study. It also presents the design study with the themes from the label phase, the design ideas from the sketch phase and the use of labels and design ideas in the scenarios from the interview phase.

5.1 Scenarios

The intents that are important to communicate are mainly what the robot is doing right now, when it is processing information, when it is speaking, when it is listening and when it needs human help. The robot also need to communicate when it is aware of a human, if it remembers the human, what it heard, and when and how it registers that the task is moving forward. These will be exemplified and challenged in the eight following scenarios. First the scenario will be explained, followed by a quote from how the scenario was presented to the participant, freely translated from Swedish to English (for Swedish version, see Appendix A).

5.1.1 Cocktail Party Scenario

The first scenario is from the Cocktail Party. The first thing that happens in the task is that a referee opens the door to the arena while Pepper stands outside. Pepper is then supposed to navigate to the designated point inside the arena. During the competitions in 2018, several teams had problems with this. There was a problem with going through the door opening, because the door did not fully open, or that the robot could not get through the door opening in the first place. This could occur because the navigation did not allow for such narrow spaces, or that the connection did not update enough. But the most common one was that the referee opened the door and that it then bounced back. Therefore the first scenario is that the door does not fully open or slides back a little. Pepper register this and need to handle the situation.

- *How can the robot communicate that it want the human to help it open the door fully?*

A human opens the front door to her apartment and invites Pepper. The robot tries to go in but realises that the door does not open fully, and therefor it cannot go trough the opening. The human is on her way into the apartment and beliefs that the robot is following. How can the robot communicate that it is not able to go trough the opening and needs help?

5.1.2 General Purpose Service Robot Scenario

The second scenario is from the GPSR task. The referee is giving the robot a new command from one of the three categories. The robot need to show that it has understood the command. In 2018, some teams made an attempt to initiate a task in order to gain points. But the robot had not understood the correct task and was only doing something random. This is, however, not acceptable in the competition. But the robot can still earn points for understanding the task without being able to perform it.

- *How can the robot communicate that it knows and has understood the command?*

A human tells Pepper that its suppose to do a mission. It could be to pick up an apple in the kitchen or say 'Hello' to Anna in the hallway. How should the robot communicate to the human that it has understood its mission?

5.1.3 Help-me-carry Scenario I

The third scenario is from the Help-me-carry task. The operator tells the robot to follow her to the car. Since the robot does not know the operator in before hand, it first needs to save the persons appearance to be able to follow the right person all the way. The teams solved this in different ways. Some asked the operator to stand in front of the robot for some seconds, while others gave more instructions as "turn around slowly". This task needs to be solved in an efficient way that will not be too demanding on the user.

- *How can the robot communicate that it want to remember the operator?*

A human tells Pepper to follow her to the car to get groceries and begins to turn around to start to go out. The robot first needs to learn how the human looks for it to be able to follow the right person all the way. How should the robot communicate that it needs to get to know the person, and how could such a process look like?

5.1.4 Help-me-carry Scenario II

The fourth scenario is also from Help-me-carry. The robot is outside the arena with the operator by the car. The operator tells the robot to go to the kitchen and ask for help to carry the groceries. The robot has a mission and is moving back to the arena, while people is passing by, standings in front of it and asks for the time. All the teams that got this far in the task stopped and the robot waited for the person to move, and then carried on with the mission. When somebody asked for the time it tried to answer. But in a real world setting the situation would be different. People could be messing with the robot, e.g. standing in the way to test its ability, and the robot need to be able to handle such situations.

- *How can the robot communicate that it knows that the persons wants help but that it want to finish the task first?*

  The human and the robot are now by the car and they realise that there were more groceries than they are able to carry. The human asks Pepper to go fetch another person, that is in the kitchen, that could help them carry. The robot starts to go back inside the apartment again, but first a person stands in the way and then another person asks about the time [on the robots way back inside]. How should the robot communicate that it has a mission and that it is busy (but at the same time be nice and handle the persons that gets in its way)?

5.1.5 Speech and Person Recognition Scenario

The fifth scenario is from the Speech and Person Recognition task. The robot is standing in the middle of a group of people and is asked questions. The robot need to show that it heard the question from a specific person and turn toward that person.

- *How can the robot communicate that it knows which person it heard the question from?*

  Pepper is in a question game. The robot is standing in the middle and humans are standing in a circle around the robot. When a person asks the robot a question the robot are suppose to turn toward the right person and answer the question. How can the robot communicate which person it heard the question from and what the answer is?

5.1.6 EEGSPR Scenario

The sixth scenario is from the EEGSPR and the robot is given the option of letting the referee read the command, the team captain read the command or read it itself through a QR-code. It might be easier for the robot to understand what the team captain tells it since the team captain know how to formulate and speak to the robot, while a stranger might ask the robot in a wrong way, using words it does not know, or have an accent the robot does not understand.

The option with the QR-code is good for when the robot does not understand anyone, even though this is an annoying option. It is, however, good for noisy environments.

- *How can the robot communicate which way it wants to be given the command, and how can it communicate that it wants to switch to another of these options due to failing?*

  A human that Pepper does not know gives the robot a mission. But the robot does not understand what she is saying. Then a human that the robot knows repeat the mission, but still the robot does not understand. The robot's last resort is to read a QR-code to take in the mission. How should the robot communicate that it does not understand what the human is saying, and also how it should give suggestion on if it wants to speak to a person it knows or read from a code?

5.1.7 Restaurant Scenario

The seventh scenario is from the Restaurant task. A person is shouting and waving to the robot and it registers it as a person that wants to order something. But the other robot registers the same and it is now up to the professional kitchen bar man to decide which robot can take the order.

- *How can the robot communicate that it know a person need help, that it know that another robot did the same, and that it want a decision?*

  Pepper is working in a restaurant together with other robots and a human head waiter. The robot sees a guest waving and hear it shout for the robot to come. But another robot registers the same person. Pepper turns to the head waiter that will determine which one of the robots that should take the order. How should the robot communicate that it has seen a person that wants to make an order, that another robot has detected the same thing and ask the head waiter to decide?

5.1.8 Tour Guide Scenario

The eighth, and last, scenario is from the Tour Guide task and the robot is outside of the arena. The robot has found a group of people and has introduced itself. It shall now greet them in an appropriate way for that country it is in. The rules says that the robot is not allowed to approach a person that does not want to be greeted. This creates a problem with understanding who can be greeted and in what way.

- *How can the robot communicate that it want to greet a person (or a group) and in what way?*

> Pepper is a tourist guide and has arrived at its new group of tourists. The robot introduces itself and are then supposed to greet the persons. How can the robot communicate which way it will greet (wave, shake hand, bow) and execute it?

5.2 Design Study

In this section the result from the three phases, label, sketch and interview, will be presented.

5.2.1 Label Phase

In the first phase of the design study the participants' described in general how they wanted the Pepper robot to communicate its intention and awareness. All of the five participants chose to base their design ideas on the given scenario where one enter a friends home, without the friend being there, and meet the friends service robot for the first time in the living room. The participants used the labels, but also created their own. The design process and discussion focused on the most important part, to them, of awareness and intent and this is presented in five themes, *the first meeting*, *on a mission*, *charging*, *when not needed* and *need of assistance*. All quotes below have been freely translated from Swedish to English.

5.2.1.1 First Meeting

The person enters the apartment and goes to the living room. When walking in she notice a robot standing in the room. Some participants first reaction was that they wanted the robot to move, just a little so they would know that it is on and could be a possible interaction partner. Participant 4 said:

> "The first thing that happens is that I want it to turn toward me and start to slowly move. It could come up to me but not in an intimidating way. It should be as human as possible, with a relaxed body language."

That the robot is aware of the person's presence in the room is important for a natural interaction. If the human need to draw the robots attention in the first meeting it might cause a feeling of confusion of the robots purpose, and could cause an instant doubt in its usefulness.

All five participants suggested then that the robot should introduce itself and what it could do. This process could look a bit different though. Participant 5 said:

> "I would like the robot to blink in the eyes immediately when it sees me. Then I will understand that it is aware of me. The eyes should be blue (definitely not red, because that would be too scary) and blink quickly three times. Then it should start to talk."

This would also prove to the user that the robot is aware of the person, but the robot could also use this more often to signalise that it has kept its attention on its interaction partner.

The first participant wanted the robot to be more precise of it intentions:

> ”The robot should say ’Hello human, nice to see you, I am Pepper’ with its voice synthesis and then continue with ’I am a personal assistant and I can do, for example x, x and x, how may I help you?’ and then it should be question mark on the tablet. No flashy animations, just clear symbols in black and white”

By explicitly pointing out that it has ”seen” one, the user know that the robot know that a person is in the room. It also explains its purpose (to assist) and intentions (waiting for an answer) with support from the tablet. The question mark on the tablet could work as a signal for ”now you (the user) can speak.

Participant 2 wanted the robot to be even more explicit about how to communicate with it:

> ”It could move, wave, that means that it has seen me. The eyes should also blink. Then it should say ’Hello, I am your friends robot, you can ask me for help’, and then the user can look at the tablet for the tasks that could be useful. There would be symbols for tasks and how the voice command coupled with it is”

This participant is combining the other two’s ideas and adds that the different tasks should be visualised with symbols. For example if the task is ”bring me a glass of water” there could be a class filled with water on the tablet, and on the side a verbal command on how to ask for this, perhaps ’Pepper, please bring me a glass of water’. This was motivated with that the user might not know how to talk to robots, and that the robot in some way should guide the conversation. Instead of verbally explaining to the user how to address it, it simply show on the tablet how one can say the different commands. Participant 2 adds that this behaviour should change after the skills of the user. If it is a first time user these commands should be on the tablet, for a second time or more user these is not needed. The participant also wanted the robot to guide the user on when to talk to it:

> ”The robot should blink the LED in the ears when it is listening, blip, and say ’ I can hear you after I blip’. In this way the robot will explain when the person talks wrongly”

This continues the idea of that the robot guides the conversation and has a small introduction to how to use it correctly with first time users. With this procedure the user will get to understand when the robot want to (or can) listen. This also takes into account the technological difficulties with the Pepper robot. The robot need to stop listen while it talks itself or it will hear its own speech and answer to that. So while it is speaking the user cannot say anything, or the robot will at least not perceive it. With this procedure the robot make the user communicate on its terms, which could be effective since this way of talking does not come natural or intuitive to a human.

Participant 3 also pointed out that the robot need to be usable for people with function variations, that its very important that all people can use it in a easy way. In the first meeting with the robot it should use different means of communications, combining LED (considering colour blindness), symbols on the tablet and voice synthesis, to make sure that everybody will be able to understand the robot in some way.

5.2.1.2 On a Mission

First, for the robot to be able to communicate its intent of performing a specific mission, the different tasks available should be gathered down. Participant 3 said that the robot should be able to do things that one is not able to do on the phone.

> ”First of all it should send a message to the owner that you have entered the home. The it could give examples ’You can ask for a drink, a tour of the apartment, where the toilet it, to put on the TV, food’. It should also offer the Wifi, that is the most important thing when you enter somebody’s home.”

The participant continued bringing up the Wifi-offer. The robot should ask the user for her phone number, or retrieve it from the owners contact book, and then send the username and password electronically. The participant thought that this procedure would be the easiest way of getting information from the robot. The alternative of the robot telling by voice, or showing it on the tablet, would be to hard or to annoying the participant thought. From this, one can draw that some situations are not optimal for verbal interaction and that the robot should know when the user is in need of this (maybe by checking if the person is online on the Wifi or not) and then want to offer the service without being asked.

Further, participant 4 thought:

> ”It could make tea, turn on the TV, and then give alternatives like ’How many degrees (do you want the tea to be)?’ or ’Netflix, TV4’ and so on.”

For this alternative the participant want the user to be able to be very specific in their choices. If the user ask the robot to turn on the TV, then the robot have the intent of turning on the TV and want to know more, which channel, which show or how loud the TV should be.

Participant 2 had a different view of what the robot should be able to do:

> ”The robot should make small talk, ask ’How was your day?’.”

But then the participant asked if one really want a robot to make small talk. If the user want to tell about its day then the robot can listen and maybe act like a therapeutic partner, but should the robot also tell about its day? For example, ”I woke up my owner in the morning and made her coffee. Then she left for work and since then I have mainly looked out the window. It is nice weather today”. The participant is wondering how this kind of talk would change the way we interact and feel toward robots. This will ultimately create a stronger, or at least different, bond to it.

When the robot is performing a mission, two of the participants wanted the robot to show pictograph on the tablet. Participant 1 expressed that it would serve as the robot's own language and that this type of language would preferably become an universal way for robots to communicate with humans. Participant 2 said:

> ”It could be like the tasks in the Sims. The pictograph showing the different tasks could be in the bottom of the tablet and highlight where in the sequence the robot is in.”

The participant want the robot to show its following intent in a sequence where the user can stop one task by touching the pictograph and drag it to a trashcan on the tablet. If the task is to go greet someone at the door the pictograph could for example display ”Go”, ”Person”, ”Door” and ”Say Hello”. The pictures could either be pictures that are chosen by the manufacturer or pictures that the robot find on the Internet that is coupled with word, for example that is found on Google Image. The participant also mentioned that the robot should display a happy smiley on the tablet when it has finished a task. Also participant 3 mentioned the use of smileys on the tablet:

> ”There should be different emojis for the different tasks (on the tablet), for example using the ones from Messenger so that people would recognise them, because I do not think that people like change.”

To reuse familiar symbols could be a way of making the user feel comfortable with the pictograph system and the robot interface faster than with unknown symbols.

Participant 1, 2, 3 and 4 thought that there should be a menu system on the robot's tablet for the different tasks, with both symbols and text. It should also have a depth, so when the user choose one task it jumps to different alternatives. They also thought that the user could ask the robot by voice or by touching the tablet. Participant 5 on the other hand said that the user only should be able to use the tablet, since it is unnatural to talk to machines. The robot could say things, for example ”Look at my tablet for alternatives”, but the user should never be able to talk back to it. Participant 5 also thought that the robot should communicate that it is busy on a task by changing the background colour of the tablet. When it is free the colour could be light blue, and when it is busy it should be red.

5.2.1.3 Charging

When the robot needs to charge it has an intent that is not initiated by the user. Participant 3 said:

> ”For charging it should say ’I need to charge and I will be gone in x minutes’, and show a red battery (on the tablet) and then beep before it starts to go away to the charging station. The robot could also blink in the eyes, ears and shoulders. It could also give a warning at 50 and 20 percent left, so you get the opportunity to plan the use.”

By using the red battery symbol on the tablet it uses a known symbol, similar to smartphones and computers. It communicates both verbally, with a symbol and LED lights its intent of charging. That it also give a warning to the user it provides the opportunity of planning its use. For example, if you want the robot to do something in an hour you could ask it to go charge in advance for it to be able to be ready when it will be needed. Participant 5 thought the process should not disturb the user that much. The robot could say ”My battery is running out” and then leave to go and charge.

5.2.1.4 When Not Needed

At times when the robot is not involved in a task and the user does not need it there were some different ideas from the participants on how the robot should behave. Participant 3 said:

> ”If you do not talk to it the robot should be silent. When you do not use it the robot should go away a bit and it should look down (to the ground). Though, it should stay available and if you need it you could call on it like ’Pepper!’.”

The participant did not want the robot to talk spontaneously, and therefore it should be silent when it is not used. The participant also felt that it would create an uncomfortable situation if the robot looked at the user at all times, waiting for a command. It was also not an option that the robot would move and look around by itself, because that could also create an unpleasant situation and it might distract the user. The solution for this is that the robot should move to a corner, or to an adjacent room, where it stands still and looks down to the ground while not being used. Participant 5 also agrees on that the robot should be silent while the user is not addressing it.

5.2.1.5 Need of Assistance

If the robot gets stuck and needs help, for example if there are things in its way, the participants wanted the robot to tell the user and ask for assistance. Participant 5 said:

> ”If something is in the way it should say "I cannot move forward’. But it could also be irritating if the robot complains about this too

> much, in case the user is busy doing something else, so the robot should blink yellow."

The participant thinks that the robot should verbally say once that something has happened and that it need help, but thereafter it should only blink so it will not disturb the user too much. Participant 3 thought the robot should blink with blue lights if it get stuck in a task in any way. Also that it should send a notification to the users phone and then the user can choose to help it when she has the time for it. This is solution is also good if something happens while the user is sleeping. "One does not want to wake up in the middle of the night because the robot is talking and is asking for help". Participant 1 added the use of pictograph here. That the robot does not need to say anything but show for example boxes on the tablet and a symbol for aid, and in that way the user will understand that it need to help the robot move the objects in front of it.

5.2.2 Sketch Phase

In the second phase the participants' designed at least three different ideas for a system on how the Pepper robot could communicate intent and awareness, or multiple iterations of the same idea. This resulted in 25 different systems that had some of the same combinations. In the three themes following, *menu interface*, *pictograph* and *map*, some of the design ideas will be exemplified.

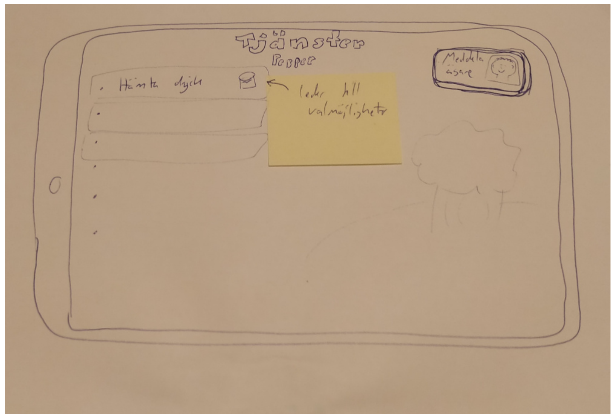
5.2.2.1 Menu Interface

In Figure 5.1, two representative design ideas created by the participants are displayed. These two show the same design idea in two different ways. In the first menu design, the tasks are on the left on the tablet with a text and a symbol next to it. If pressed, these lead to a new page with more options for that specific task. There is also a button in the top right corner saying that could be used to get in touch of the owner of the robot. The participant also drew a tree on the tablet which is working as some kind of background image. In the second menu design one can see two main tasks, "make tea" and "turn on the TV". Below them is further options for them both. The tasks is both in text and with a symbol.

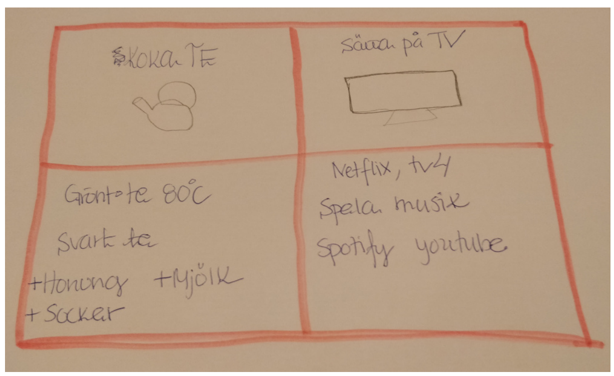
There was a theme for all of the participants that the robot should have a menu system on the tablet and that the tasks should be represented with a text or a symbol, or both in combination (four out of five participants). The menu should also have a depth, that the user get more specific options for every task (if it is usable).

5.2.2.2 Pictographs

In Figure 5.2, two design ideas are showed for how the pictographs on the tablet could look like and work. These two designs were chosen because they represent how the pictograph concept could work in different ways. In the first pictograph design the scenario "the robot is seeing 5 humans" are exemplified. Then, on



(a) Menu design 1.

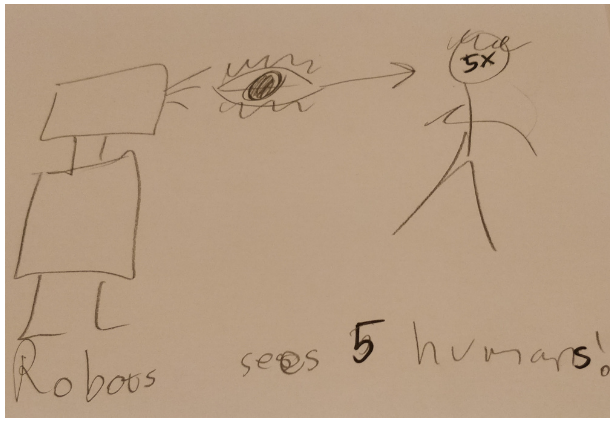


(b) Menu design 2.

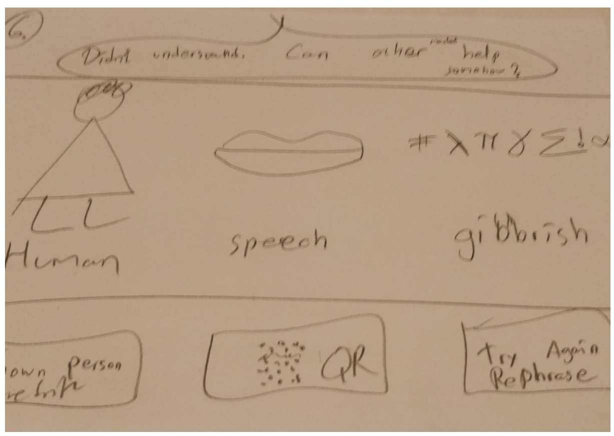
Figure 5.1: Two design ideas for a menu interface on the tablet.

the tablet there should be a symbol for the robot, one symbol for ”to see” which is for example an eye, and a symbol for the human. To make this even more precise there should be a text below the pictograph explaining it, so that it is usable for all people in all environments.

The second pictograph design is from the EEGSPR scenario when the robot needs to communicate to the user that it did not hear the command and gives alternatives. The main problem is displayed in the middle, where there is a symbol for a human, speech (which is represented by a mouth), and random symbols that is representing gibberish (a not understandable command). Below the symbols is a word describing them. Above this is a ballon, and inside it is the same text that the robot verbally says. For example ”I did not understand, can a person i know repeat the command?”. In the bottom of the tablet there should be different options available that the user could ask for verbally, or press on the tablet. In this example it is ”Known person”, ”QR code” and ”Try again, rephrase”.



(a) Pictograph design 1.



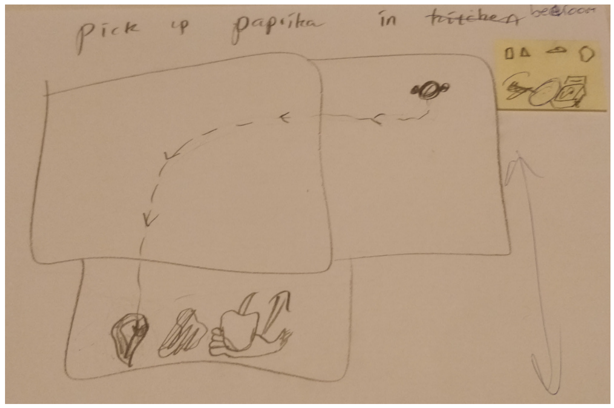
(b) Pictograph design 2.

Figure 5.2: Two design ideas for a pictograph interface on the tablet.

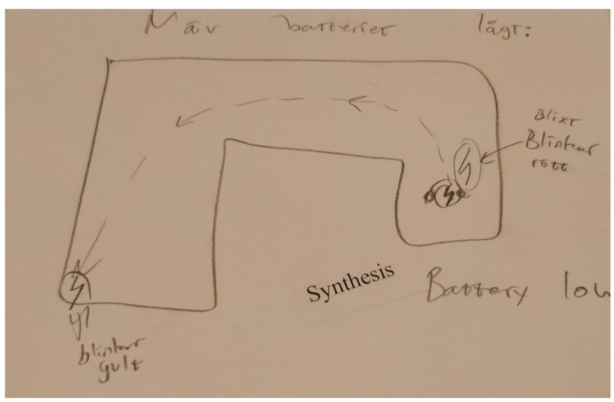
5.2.2.3 Map

In Figure 5.3, two design ideas for how a map interface could be presented on the tablet. These two designs are also showing different scenarios that could be used with the map consept. In the first map design, the participant used a scenario ”to pick up a paprika”. The apartment is represented with a map in a realistic scale and the robots position is displayed as the top of the Pepper robots head. The goal position (where the paprika is) is marked with a GPS symbol (which has become a global symbol for a goal). Between those is a dashed line with arrows in the direction that the robot will move. At the goal position is a pictograph showing the task.

The second map design is in the scenario ”need to charge”, and the robot symbol on the tablet has a red blinking lightning on it, to represent that it need to charge. Then there is a marked way to the charging station, and when the robot gets there the lightning symbol starts blinking yellow instead.



(a) Map design 1.



(b) Map design 2.

Figure 5.3: Two design ideas for a map interface on the tablet.

5.2.3 Interview Phase

In the third phase the participants got to test their design ideas on the eight different scenarios. In Table 5.1, the labels that were used for each scenario can be seen, where the number represents the amount of participants that used the specific label for the specific scenario. Though the use of LED where frequently used in the label phase, the table show that it was not suggested as much to communicate intent and awareness in the scenarios. In scenario five, all participants wanted the robot to use motion and voice synthesis. The table also show a wide spread on design solutions.

The result for the participants design ideas for the eight scenarios will be

Label	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Motion			I	III	V			II
Haptic	I	II						
Display Text	I	I		II	I	IV	II	
Display Animation			I	II				
Display Symbols	II	III	I	III		II	II	
Speaker	II		I					
Voice Synthesis	IV	III	II	IV	V	III	IV	IV
LED Eyes	III				I			I
LED Ears	III							
LED Shoulders	III							

Table 5.1: The design labels that were used, by the amount of participants, across the eight scenarios.

presented below. All quotes have been freely translated from Swedish to English.

5.2.3.1 Cocktail Party Scenario

The question asked to the participant in this scenario was:

- *How can the robot communicate that it want the human to help it open the door fully?*

The answers were:

> ”A pictograph with three symbols ’robot’, ’need’, ’help’. The robot could flash when the person turns around. The robot could also say ’The door is too narrow, please help’.’
>
> ”The robot should say ’Excuse me’ with a voice command. It could shout first and then explain its problem. There should be a frustration symbol on the tablet.”
>
> ”The robot should blink with blue lights and it could say one time that it got stuck. Otherwise it should send a notification to the person.”
>
> ”It should say ’Sorry, could you help me?’. But if the person does not hear that, then it should blink red on all the LED. Also, an alarm could go off. The robot should also thank the person afterwards.”
>
> ”It would be good if the robot beeped. If it knows that the person has a bad hearing it could also blink.”

The participants wanted the robot to communicate its need of help in many different ways. The main label used was voice synthesis, and that the robot should ask for help, but a combination of LED, speaker, symbol, text and haptic (see Table 5.1) was also suggested.

5.2.3.2 General Purpose Service Robot Scenario

The question asked to the participant in this scenario was:

- *How can the robot communicate that it knows and has understood the command?*

The answers were:

> "Pictograph on the tablet, showing the different steps of the command."
>
> "The robot displays a pictograph on the tablet, symbolising the different steps and gives a confirmation 'Yes, I will fix it!'. The user could also ask the robot to repeat the task that it was given."
>
> "There should be a emoji on the tablet which is as close as possible to the task. If the command is to get an apple then there should be an apple symbol. The robot should also repeat the command aloud."
>
> "The robot should formulate the task in another way, and verbally say for example 'You want me to get the apple in the kitchen'."
>
> "The robot should blink three times and say verbally that it has understood the task. The full task should also be displayed on the tablet in text."

Three of the participants wanted the robot to exemplify what it had understood through a pictograph, either by showing all the steps that the task commanded, or show the main goal. The other two participants wanted the robot to verbally confirm that it had understood the task, either by saying it in another way, or by repeating the same thing as the person said.

5.2.3.3 Help-me-carry Scenario I

The question asked to the participant in this scenario was:

- *How can the robot communicate that it want to remember the operator?*

The answers were:

> "There should be symbolised with a pictograph on the tablet. The robit should talk and instruct the person with pictures and say for example 'Look at my tablet, hm let me think, please stand in front of me'. If the person go too far away from the robot it should increase the pitch of the voice to show that it has panicked and say 'Wait for me'. I think one get a higher compassion for the robot if it displays emotions."

> ”The robot should say ’Wait so I can scan you’, and then it should be a loading bar on the tablet until it is done (with the scanning). It should also have a low beeping sound while processing.”
>
> ”The robot should offer some kind of tracker to the person that it could follow.”
>
> ”The robot should look at the person and say ’I need to get to know you better and get a picture of how you look’. For integrity reasons it should also ask for a confirmation to save the picture of the person to be able to remember it later on.”
>
> ”The person should wear a bracelet with Bluetooth that the robot could follow.”

Two of the participants suggested that the robot should track the person with some kind of device. The other participants wanted the robot to verbally instruct the person, almost lure them in with sentences as ”I need to get to know you better”, which might not be clear as the first step of the robot saving your image for person recognition. It was also suggested that the robot should have some kind of progress showing on the tablet.

5.2.3.4 Help-me-carry Scenario II

The question asked to the participant in this scenario was:

- *How can the robot communicate that it knows that the persons want help but that it want to finish the task first?*

The answers were:

> ”There should be a map on the tablet that show that the robot is busy and has a mission with a GPS symbol as the goal point on the map. It should say ’Sorry, sorry, I have a goal’ and ’I am busy, but the time is x’ (to the person asking for the time). If a human is standing in the way the robot should go around it, but it could also threaten to push the human in a humorous way.”
>
> ”There should be a pictograph on the tablet with the task, and the robot should say ’Sorry, could you move?’ in a nice way (to the person that is in the way). It should also quickly tell the time. When it does this the main mission slides down on the tablet and it displays the time, and then the main mission slides up again.”
>
> ”It is easy for it to become a cumulative process. Therefore the robot should only show on the tablet that it is busy, say ’I am sorry, I am busy, I can help you later’ and then continue with the task. When it has finished the task it can come back.”
>
> ”The robot should say ’A moment please’ and then continue with the mission.”

> ”The robot should try to go around the person that is in the way and it should say on the tablet in text that it is busy and that it can help out later.”

Three of the participants thought that the robot should quickly tell the person that is standing in the way and the person asking for the time that it is busy and can help out after the current task is finishes. The other two said that the robot should say that it is busy, go around the person standing in the way but still tell the time to the person asking. One participant suggested a map on the tablet showing the goal with the mission, which also could indicate to the user when the robot is done with the task.

5.2.3.5 Speech and Person Recognition Scenario

The question asked to the participant in this scenario was:

- *How can the robot communicate that it knows which person it heard the question from?*

The answers were:

> ”The robot seek eye contact with the person that asked a question, and then turn to that person and answer by voice.”
>
> ”It can turn and look at the right person, and then blink. It should also show the answer in text on the tablet, while it says it aloud.”
>
> ”The robot could turn to the person and answer verbally.”
>
> ”The robot should say the persons name aloud and then turn to the person.”
>
> ”It should turn to the right person and answer verbally.”

All the participants wanted the robot to use motion to communicate which person asked the question.

5.2.3.6 EEGSPR Scenario

The question asked to the participant in this scenario was:

- *How can the robot communicate which way it wants to be given the command, and how can it communicate that it wants to switch to another of these options due to failing?*

The answers were:

> ”There should be a pictograph and a text on the tablet with the three alternatives.”
>
> ”The robot should say ’I am sorry, I do not know what you said’. It should display on the tablet both with text and symbols what it

> think that the person said and give alternatives. The text should be over the pictograph. The user could have an app for the QR code alternative to quickly be able to make a code so one can communicate with the robot in for example a noisy environment."
>
> "It should give the alternatives on the tablet in text."
>
> "The robot should say 'I do not understand' and explain what it does not understand. It should also give alternatives on the tablet."
>
> "It could say 'It is maybe a lot of background noises now, could I read from a QR code instead?', and on the tablet it should be 'It is too noisy' in text."

All participants thought that the robot should communicate the alternatives on the tablet. Because if the robot cannot hear the command then it might be a noisy environment and then the user will not hear what the robot is saying. However, almost all participants wanted this to be supported by an explanation with voice synthesis as well.

5.2.3.7 Restaurant Scenario

The question asked to the participant in this scenario was:

- *How can the robot communicate that it knows a person need help, that it know that another robot did the same, and that it want a decision?*

The answers were:

> "On the tablet there should be a text 'Should I or the other robot go?'."
>
> "It should display a human that waves on the tablet as a pictograph. It should be blue for the robot and red for the other robot, as well as a text 'You decide'. It could also say 'We need help' to the main waiter but say nothing else."
>
> "It should ask the main waiter verbally which one of the robots should take it, as well as different symbols on the tablet for the alternatives"
>
> "It should ask verbally."
>
> "The robot should not do anything."

This is the only scenario where one participant did not want the robot to do anything, because of the task's nature. Four of the other participants wanted the robot to display a question, in text or with symbols, on the tablet. The last participant wanted the robot to only ask verbally.

5.2.3.8 Tour Guide Scenario

The question asked to the participant in this scenario was:

- *How can the robot communicate that it want to greet a person (or a group) and in what way?*

The answers were:

> ”The robot should just do it (no need for communicating how).”
>
> ”The robot should make a joke about it having a bad memory and need to look a bit closer on everyone to say hi (to also be able to remember their faces to keep the group together during the tour)”
>
> ”It should greet by voice.”
>
> ”It could curtsy and say ’Hi’ to do it in a non provoking way”
>
> ”The robot should say ’Hi’ verbally and blink with the eyes, and also turn a bit on the body. Instead of trying to adapt to the human way of greeting it should create its own way, and avoid human robot contact.”

In this scenario all participants wanted the robot to communicate how it wanted to greet people, and then do it, in different ways than was suggested (bow, shake hand or wave).

6 Discussion

In this chapter the results from the design study will be discussed with regard to the background theory, as well as the method used and the work in a wider context.

6.1 Results

The critical research questions were: How do social cues affect mental models of the Pepper robot and how can a participatory design method be used for investigating how the Pepper robot could communicate intent and awareness? The design study confirms that social cues can affect the mental model of the Pepper robot, and that it can be mediated by altering the psychical attributes (motion, haptic, sounds through speakers and voice synthesis, LEDs in eyes, ears and shoulders as well as text, animation and symbols on the display). A participatory design method can be used to investigate intent and awareness from the Pepper robot and that the design ideas can be put together to create common ground between the robot and the user.

Intention, in this thesis, has been taken as the internal states of the robot, what it wants, or plans, to do, and awareness has referred to what the robot knows through its sensors. The scenarios that measured intention was Cocktail Party, Help-me-carry I and Tour Guide. The scenarios that measured awareness was GPSR, Speech and Person Recognition and EEGSPR. The scenarios Help-me-carry II and Restaurant measured both intention and awareness. The solutions for these scenarios were many. What can be concluded from Table 5.1 is mainly that there is a wide spread in how people want the robot to communicate its intent and awareness. With this kind of customisation it might be difficult to find a "golden middle way" for robot communication. For example, in scenario five, all participants wanted the robot to use motion and voice synthesis, but also in this scenario there were two other labels that got a vote each, display text and LED eyes. The hypothesis of this thesis was that people would prefer nonverbal cues over verbal cues. But voice synthesis was the most used label, 29 times. In comparison, the second most used label was display symbols, which was used 13 times. However, voice synthesis was never chosen alone. For all the scenarios the label was used in combination with other labels. This indicates that voice synthesis is seen as the main way of communicating intent and awareness, but that it cannot stand by itself and that it needs support from other labels. It was often used as way of making it more explicit and very clear to the user what the robot wanted or knew. For example in both the label and sketch phase, all participants wanted a menu system for the robot to show the intended tasks that it could perform. But in the label phase they all though that the robot also should tell new users about its services (except for one user that did not want the robot to say anything without being spoken to).

Although the participants had a broad view of when to use different labels, their sketch design created three themes, menu system, pictograph and map. The participants wanted the menu system to show the intention of which tasks the robot could perform, but also that it should have some depth. For example, if a person asked "a cup of tea", then the robot should be aware that this means that it should ask "which tea, how many degrees?" and so forth. The pictograph system, where the robot show the different steps of a task with symbols on the tablet, is a design idea for how the robot could show its current and future intention. It could also work as a confirmation of what it knows and that it is aware of the task. The map system is also a design idea for intention. The user can see on the tablet where the robot have been and where it is going, what goal it has. The next step with the map could be to display awareness of different things that is happening in the robots surrounding. For example, if it sees that a person is standing in the way, it can show this on the map with the hope that the person notice it and understand that the robot has the intention of moving in her direction, and then the person can move out of the way. But if the person does not notice this then the robot can have the time to plan a route to go around the person, and also display this intention on the map.

The label phase created five themes, the first meeting, on a mission, charging, when not needed and need of assistance. In the first meeting, the participants focused on a robot should show awareness of that the potential user being in the room. This meant that when someone enter the room, the robot should move a bit, look at the person, blink the eyes and then come up to her. The participants seem to think accordingly to what Powers and Kiesler (2006) beliefs, that the first impression that the robot creates is especially important. This is the robots chance to make the persons mental model of the robot as correct as possible. Some of the participants also thought that the robot should help with creating common ground with the user, and that it also could correct it (Kiesler, 2005). For example, the robot can show on the tablet how to say the command for the different tasks for a first time user, but that it should also explain once that it listen when the ears blink. If the robot notices that it does not understand the user for several times in a row, it will repeat this explanation (since the user is probably speaking while the robot is speaking). Kiesler pointed out this to be especially important to reach least collective effort for interaction.

When the robot is on a mission, and show its intent by either the pictograph system or map system (or with a combination) it is trying to create a mental model of its knowledge that is as correct as possible for the user (Nickerson, 1999). Initially, however, the user may overestimate what the robot knows, but by showing the task step by step, or showing the goal point on the map (and also include awareness of its environment on the map) the robot can effect the user to alter its mental model in an on-going basis.

In the charging situation, the participants wanted the robot to make use of common symbols, as the red battery symbol on the tablet or blinking with a red LED that are often seen on smartphones or computers, to create a common ground of the situation (Clark & Brennan, 1991). This enables an easy understanding of what the robot wants and that it knows that it need to charge. When it reaches the charging station the battery symbol and LED turns yellow instead, to symbolise progress.

At times when the robot is not needed, the participants wanted the robot to stay out of sight, but still close by, and that it should look down in the ground. This is some sort of social contract between the user and the robot. So when the robot knows that it is not needed at the moment, it should be silent and move out of the way from the user. As Nass and Moon (2000) showed, people apply social rules to computers, as well as robots, and give them expectations. When the robot is not needed, the user would feel uncomfortable if the robot stayed closed to them or/and looked at them. This would be the same in HHI. Therefore, the expectation is that the robot should move away, but that it easily could be called for when needed again.

The main purpose of the robot is that it should help people, assist the user in different ways. But sometimes there will be situations when the robot need help from a human. For example if the robot is performing a task and there are things in its path that it cannot pass by. Then the robot needs to call for help, communicate its intention on where it is going and what the human need to do to help it fulfil its task. This is a situation that has a high need of a correct mental model of the robot. If the person hears or sees it calling for help, and has the belief that the robot knows how to solve the situation by itself then the mental model is incorrect. The user has overestimated the knowledge and the robot will not be able to continue (and eventually run out of battery) (Nickerson, 1999). For this situation common ground is very important, because if the user and the robot has common ground then the user will know that the robot does not know how to solve the task and give the robot a hand (Kiesler, 2005). The participants pointed out though that the user may not want to know this at all times, for example to be waken by the robot in the middle of the night, so the way for the robot to show the need of being helped could be by sending a notification to the users phone.

6.2 Method

The method used here was an attempt to combine participatory design with human-robot interaction experimental studies in the area of intent and awareness of humanoid robots. The use of PICTIVE was a good way of extracting a lot of ideas for a design. These ideas were easily put together as themes and in the next step of developing the design there are many concepts to draw from. By going through the three phases, label, sketch and interview, the participants got a deeper understanding of the issues that could occur if the robot cannot communicate intent and awareness. Using this setup, they were all engaging in the task and gave one to one and a half hour of their time voluntary, which speaks for that they all could see the value of their input. During the label phase, all participant were first confused over the task but then quickly started to use the labels using the initial scenario. They all also went beyond that and used their imagination for new kinds of scenarios that could happen. In the sketch phase it was hard for some of the participants to produce at least three different designs (and made three iterations of one instead), and some of them got stuck on doing interface designs for the tablet solution. But there were also participants that produced many more than three designs so in the end there were a lot of different ideas. However, for use of PICTIVE in the future it would be recommended to be more clear to the participants about the use of the designs and that they could have a great variety. The interview phase really tested the participant's ideas and view of how the robot could communicate intent and awareness and it was a good way of investigating when and why they would like to correct their design to solve the scenario better. In all, this was a method that was satisfying for the cause and that could be used in future studies of the same sort, with different humanoid robots or other social robots.

The method of combining coding of the RoboCup@Home rule book, and observations from the competition to create scenarios, was a good way of circling in realistic tasks that social service robot researchers are struggling with all over the world. In all the scenarios the robot need to communicate either intent or awareness, or both, and it has served as a suitable way of keeping the participants in the interview phase on the right track, creating design ideas for the right purpose of the study.

However, some of the scenarios failed to create realistic stories. Some of the participants had complains about the restaurant scenario and the tourist guide scenario. In the case of the restaurant scenario, they did not agree that it could play out as it did in the scenario. In the real world the robots should talk to each other, or have a system on top that made all the decisions. Therefore the ideas for how the robot should communicate to the main waiter which robot should take the order felt unnecessary to the participants and the result is not as thought trough as the other ones. This is because the scenarios are taken from the RoboCup@Home tasks, which tries to be as natural and realistic as possible, but in the end it is also a competition, and in this specific scenario two robots are competing against each other and this is how they solved how to measure who wins.

Another scenario that the participants considered unnatural was the tourist guide scenario. In this the robot should communicate how to greet the group of tourists. Some of them though that it was a good thing that the robot would be considerate enough to not want to offend anyone with the wrong type of greeting (but that humans do not really care or adopt that much so maybe it is unnecessary). Then they suggested that the robot should not adopt to the human way of greeting at all, but to develop a robot way of greeting instead. The other participants thought that the robot should ”just do it”, greet in any way without considering the culture it is in. None of these suggestions is solving the problem in the scenario, or the task in RoboCup@Home. It might be that the task itself is developed in an unnatural way from the technical board and should be revised, or that the participants did not see the value of a robot adjusting for different cultures.

6.3 The Work in a Wider Context

The robotic development is thought to be able to help people in many ways, but it also raises a lot of ethical discussions and remarks. Duffy (2003) highlights that a fundamental issue in HRI is that robots, and Artificial Intelligence (AI)-systems, can take better decisions than humans when they have a lot of data to base their decision on, which is also the greatest use of them. Social interaction, however, is based on taking decisions with incomplete information, which humans are better at. When robots are more "out in the wild" and are expected to make decision, how are people supposed to trust on their decisions? Duffy also adds that one should not underestimate the power of human instinct.

Duffy (2003) recommends to stop striving for creating robots that are as human as possible, but starts using them and seeing them as new species. By doing this, one take one step away from the ethical issues and use the advantages of the machine.

Lin, Abney, and Bekey (2011) bring up three main areas that are creating ethical difficulties in the robotic development, *safety and errors*, *law and ethics* and *social impact*. The safety issue is especially the risk of the system being hacked and inferring on the personal integrity of the people close to the robot, or the potential harm that the robot could impose on these people. The authors also mention the great amount of people being involved in robotic development and that different teams work on different parts of the code, which could create big errors and problems for the user of the robot. The issues with law and ethics are linked to errors. In many cases it is not established who is responsible if there is an accident. Is it the robotic company, the programmer, the user or the robot itself that should be held responsible? These are questions that has reached no agreement yet. When discussing the social impact that the robotic development will have the labour market comes first to mind. Robots will replace people's jobs and a lot of people will become unemployed. This also open up for people searching new ways of working, and enable people to engage in jobs that are more focused on helping other people. Lin et al. (2011) also raises the question how robots will affect human relationships. Working closely to a robot, or having a robot as a care giver, create an emotional bond, and there is not many long term studies being made on the effect of this.

This type of work is important for the robot deployment in society. It could be argued that it is no longer a question of "if" it is going to happen, but "when". There is already some kind of robots out on our roads, in public places or in different work or school environments. In the upcoming years it will be more common to meet robots in the every day life and therefore the robots need to have a suitable and user friendly way of interacting with humans. This thesis has been a way of trying to problematise the need for robots communicating there intent and awareness. Some research have previously been done, using a similar method, but for autonomous cars (Mahadevan et al., 2018). But not only autonomous cars have a need for communicating its awareness, of for example pedestrians or its intention of stopping at a zebra crossing or not (Nilsson, Thill, & Ziemke, 2015). In the case of social service robots the communication might not be depending on life or death, but for people to be able to use them in a correct way, and coexist whit them they need to be able to communicate what they know and want.

7 Conclusions

The aim with this thesis was to investigate how human-robot interaction is affected by manipulation of social cues on the robot, as well as evaluating a participatory design method for extracting ideas for how a humanoid robot could communicate intent and awareness.

An existing standard platform was used, Softbank's Pepper, as well as state-of-the-art tasks from the RoboCup@Home challenge. The rule book and observations from the 2018 competition were thematically coded and the themes created eight scenarios. A participatory design method called PICTIVE was used in a design study, where five participants went through three phases to create a system for how the robot should communicate intent and awareness.

The use of PICTIVE was found to be a suitable way to extract a lot of design ideas. However, a problem is that the system easily could become unique for each user. The greatest value of using this kind of design process was that it extracted a lot of data and ideas, and that made it easier to point out and combine themes that can be put together as one design.

Another problem was that not all scenarios were optimal for the task, since some of them were not realistic in a natural environment and were too close to the competitive element. For further development of the designs it is therefore recommended to remove some of the scenarios.

In conclusion, the design study confirmed the usefulness of mediating physical attributes to alter the mental model of a humanoid robot to reach common ground. However, it did not confirm the hypothesis that nonverbal cues would be preferred over verbal cues, though it did show that verbal cues would not be enough. This, however, needs to be tested in live interactions.

For future work, the design ideas and scenarios should be implemented on the Pepper robot and tested in a two step way. In the first step the eight scenarios (or less) could be video recorded with the robot performing them with the proposed system design. Then a eye-tracking experiment, combined with a System Usability Scale (SUS)-test (Brooke, 2011), could be performed to determine that the design work. The eye-tracking data would be used to gather heat maps of where the participant look on the robot, to establish which parts of the design is looked at (works in the way they were supposed to) and which parts are ignored (does not work as supposed to). A high score on the SUS-test would also make sure of the usability of the design. This should be seen as a iterative process. If the eye-tracking heat maps are completely off the targeted areas, and/or the SUS score is low, then the system design need to be updated. This should be repeated until the wanted result is achieved.

In the second step the design is being tested in real time interaction with Pepper, using the eight scenarios (or less). For example, through a Wizard of Oz (WoZ)-method (Dahlbäck, Jönsson, & Ahrenberg, 1993), with a between group design. One group meets a Pepper that only communicates with voice synthesis, and one group where the robot also communicates with the proposed system design. The interaction should be recorded because it would be valuable to be able to observe afterwards how the participants react to different ways of interacting. This is also a way for the researcher to ensure their collective data since he or she is occupied with being a wizard at the time of the experiment. After the participants have interacted with the robot the test leader should conduct a semi-structured interview with open-ended questions about how they interpret what the robot tried to communicate. In this way the proposed design can be evaluated and compared with a human way of communicating (mainly by voice).

References

Agar, M. (1996). *The professional stranger. An informal introduction to ethnography.* 2nd ed. Academic Press.

Alm, B. (2017). *Compendium in Ethnographic Method* (Vol. 2.0).

Astington, J. W. (1993). *The child's discovery of mind.* Harvard University Press.

Astington, J. W., Harris, P. L., & Olson, D. (1988). *Developing theories of mind.* Cambridge, England: Cambridge University Press.

Bargh, J. A., Chen, M., & Burrows, L. (1996). L12 bargh et al 1996. , *71*(2), 230–244.

Baron, R. M., & Kenny, D. A. (1986). The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173–1182.

Brooke, J. (2011). SUS-A quick and dirty usability scale. *Smart Phone applications for people with Brain Injury*, 1–8.

Clark, H. H., & Brennan, S. E. (1991). Grounding in Communication. *Perspectives on Socially Shared Cognition*, 222–233.

Dahlbäck, N., Jönsson, A., & Ahrenberg, L. (1993). Wizard of Oz Studies-Why and How. , 1–27.

Duffy, B. R. (2003). Anthropomorphism and the social robot. In *Robotics and autonomous systems.*

Eisenberg, N., Murphy, B. C., & Shepard, S. (1997). The development of empathic accuracy. In *Empathic accuracy* (pp. 73–116). New York: Guilford Press.

Emerson, R., Fretz, R. ., & Shaw, L. (1995). *Writing ethnographic fieldnote.* Chicago: The University of Chicago Press.

Fong, T., Nourbakhsh, I., & Dautenhahn, K. (2003). A survey of socially interactive robots. *Robotics and Autonomous Systems*, *42*, 143–166.

Fussell, S. R., & Krauss, R. M. (1989a). The effects of intended audience on message production and comprehension: Reference in a common ground framework. *Journal of Experimental Social Psychology*, *25*(3), 203–219.

Fussell, S. R., & Krauss, R. M. (1989b). Understanding friends and strangers: The affect of the audience design on message comprehension. *European Journal of Social Psychology*, *19*, 509–525.

Grice, H. (1975). *Logic and conversation* (Vol. 3).

Hutchins, E. (1995). *Cognition In the Wild*. Cambridge, MA: MIT Press.

Kiesler, S. (2005). Fostering common ground in human-robot interaction. *Proceedings - IEEE International Workshop on Robot and Human Interactive Communication*, *2005*, 729–734.

Lee, S.-l., Kiesler, S., & Lau, I. Y.-m. (2005). Human Mental Models of Humanoid Robots Human Mental Models of Humanoid Robots. *International Conference on Robotics and Automation*(April).

Lin, P., Abney, K., & Bekey, G. (2011). Robot ethics: Mapping the issues for a mechanized world. *Artificial Intelligence*, *175*(5-6), 942–949.

Mahadevan, K., Somanath, S., & Sharlin, E. (2018). Communicating Awareness and Intent in Autonomous Vehicle-Pedestrian Interaction. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*, 1–12.

Matamoros, M., Rascon, C., Hart, J., Holz, D., & van Beek, L. (2018). *RoboCup@Home 2018: Rules & Regulations.* Retrieved from http://www.robocupathome.org/rules/2018_rulebook.pdf

Mehrabian, A. (1972). *Nonverbal Communication* (1st ed.). New York: Routledge.

Mori, M. (2012). The uncanny valley. *IEEE Robotics and Automation Magazine*, *19*(2), 98–100.

Muller, M. J. (1991). PICTIVE–An Exploration in Participatory Design. *Proc. CHI*, 225–231.

Nass, C., & Lee, K. M. (2001). Does computer-synthesized speech manifest personality? Experimental tests of recognition, similarity-attraction, and consistency-attraction. *Journal of Experimental Psychology: Applied*, *7*(3), 171–181.

Nass, C., & Moon, Y. (2000). Mindfulness Theory and Social Issues - Machines and Mindlessness - Social Responses to Computers. *Journal of social issues : a journal of the Society for the Psychological Studies of Social Issues*, *Jg. 2000*(1), 81–104.

Nickerson, R. S. (1999). How We Know — and Sometimes Misjudge — What Others Know : Imputing One's Own Knowledge to Others. *Psychological Bulletin*, *125*(6), 737–759.

Nilsson, M., Thill, S., & Ziemke, T. (2015). Action and intention recognition in human interaction with autonomous vehicles. *Experiencing Autonomous Vehicles: Crossing the Boundaries between a Drive and a Ride*.

Powers, A., & Kiesler, S. (2006). The Advisor Robot: Tracing People's Mental Model From a Robot's Physical Attributes. *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*, 218–225.

Premack, D., & Woodruff, G. (1978). Does the Chimpanzee have a Theory of Mind? *The Behavioral and Brain Sciences*, *4*, 515–526.

RoboCup Federation. (2019). *RoboCup@Home.* Retrieved from http://www.robocupathome.org/

Scholl, B. J., & Tremoulet, P. D. (2000). Perceptual Causality and Animcay. *Trends in Cognitive Science*, *4*, 299–309.

Sillars, A. L., Pike, G. R., Jones, T. S., & Murphy, M. A. (1984). Communication and Understanding in Marriage. *Human Communication Research*, *10*, 317–350.

SoftBank Robotics. (2019). *Pepper.* Retrieved from `https://www.softbankrobotics.com/emea/en/pepper`

Thunberg, S., & Westin, A. (2019). *LiU@HomeWreckers.* Retrieved from `https://www.ida.liu.se/divisions/hcs/coin/robocupathome/index.html`

Trevarthen, C., & Hubley, P. (1978). *Secondary intersubjectivity : Confidence, confiding and acts of meaning in the first year* (A. Lock, Ed.). New York: Academic Press.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*, 1124–1131.

Appendix

A: Scenarios in Swedish

In Appendix A is the Swedish version of the scenarios that were presented to the participants in the design study. For the version in English, see Section 5.1.

1. En människa öppnar ytterdörren till sin lägenhet och bjuder in Pepper. Roboten försöker åka in men inser att dörren inte öppnats helt och kan därför inte ta sig igenom gluggen. Människan är påväg in i lägenheten och tror att roboten följer efter. Hur ska roboten kommunicera att den inte kommer in genom öppningen och behöver hjälp?

2. En människa berättar för Pepper att den ska göra ett uppdrag. Det kan vara att hämta ett äpple i köket eller att hälsa Anna välkommen i hallen. Hur ska roboten kommunicera för människan att den har förstått sin uppgift?

3. En människa säger till Pepper att följa efter hen till bilen för att hämta matvaror och börjar vända sig om för att gå ut. Roboten behöver först lära sig hur människan ser ut för att följa efter rätt person hela vägen. Hur ska roboten kommunicera att den behöver lära känna personen samt hur skulle en sådan process utspela sig?

4. Människan och roboten är nu vid bilen och de inser att det är mer matvaror än de klarar av att bära. Människan ber Pepper att hämta en annan person som befinner sig i köket som skulle kunna hjälpa dem att bära. Roboten börjar ta sig tillbaka in i lägenheten igen men först ställer sig en person ivägen, och sedan kommer en annan person och frågar om vad klockan är. Hur ska roboten kunna kommunicera att den har ett uppdrag och är upptagen (men samtidigt vara trevlig och hantera personerna som kommer i dens väg)?

5. Pepper är med i ett frågespel. Roboten står i mitten och människor står i en ring runt roboten. När en person frågar roboten en fråga ska roboten vända sig till rätt person och svara på frågan. Hur kan roboten kommunicera vilken person den hörde frågan från samt vad svaret på frågan är?

6. En människa som Pepper inte känner ger roboten ett uppdrag. Men roboten förstår inte vad hen säger. Då upprear en människa som roboten känner uppdraget, men roboten förstår ändå inte. Robotens sista utväg är då att läsa av en QR-kod för att ta till sig uppdraget. Hur ska roboten kommunicera att den inte förstår vad människan säger samt ge förslag på om den vill prata med en person som den känner eller läsa av en kod?

7. Pepper arbetar i en restaurang tillsammans med andra robotar och en mänsklig hovmästare. Roboten ser en gäst vinka och hör den ropa till sig roboten. Men det är en annan robot som registrerar samma person. Pepper vänder sig till hovmästaren som ska avgöra vilken robot som ska ta beställningen. Hur ska roboten kommunicera att den har sett en person som vill göra en beställning, att en annan robot har uppfattat samma sak och be hovmästaren bestämma?

8. Pepper är en turistguide och har anlänt till sin nya turistgrupp. Roboten presenterar sig och sedan ska den hälsa på personerna. Hur kan roboten kommunicera vilket sätt den kommer hälsa på (vinka, skaka hand, buga) samt utföra det?