

Research Article

Sofia Thunberg* and Tom Ziemke

User-centred design of humanoid robots' communication

<https://doi.org/10.1515/pjbr-2021-0003>

received February 27, 2020; accepted July 20, 2020

Abstract: Interaction between humans and robots will benefit if people have at least a rough mental model of what a robot knows about the world and what it plans to do. But how do we design human-robot interactions to facilitate this? Previous research has shown that one can change people's mental models of robots by manipulating the robots' physical appearance. However, this has mostly not been done in a user-centred way, i.e. without a focus on what users need and want. Starting from theories of how humans form and adapt mental models of others, we investigated how the participatory design method, PICTIVE, can be used to generate design ideas about how a humanoid robot could communicate. Five participants went through three phases based on eight scenarios from the state-of-the-art tasks in the RoboCup@Home social robotics competition. The results indicate that participatory design can be a suitable method to generate design concepts for robots' communication in human-robot interaction.

Keywords: user-centred design, participatory design, communication, human-robot interaction

1 Introduction

Communication between humans and interactive robots works better when people have a clear mental model of what robots can and cannot do [1]. Providing a correct first impression of the robot's knowledge and its interactive capacities seems especially important for service robots that interact with people in public environments,

i.e. where people might encounter a robot for the first time and where robots need to deal with users with very different backgrounds and levels of experience. Ideally, robot designs should take advantage of the kind of social cues that would enable people to make inferences about the robot based on elements of its appearance or behaviour [2,3]. In that case, the robot does not need to be explicit about the details of its capacities, its knowledge or its plans. Previous work has shown that it is possible to influence people's mental models of robots [1,4,5], but most of the research have not been carried out in a user-centred way. We would like to emphasise the importance of involving potential users as early as possible in the design process. This makes it easier to create intuitive user interfaces by identifying the users' goals, tasks and needs [6]. By employing a user-centred design methodology, we can reduce errors and improve productivity without requiring significant new technological capabilities. Otherwise, if robots are developed using an overly technology-centred design, there is a risk that the user will be unable to pay attention to the important signals and that information overflow will increase [6]. Previous research has shown that user-centred design creates safer, more effective, ethical and sustainable designs [7].

The aim of this study was to investigate how human-robot interaction could be developed using a participatory design method (PICTIVE). For this, we used an existing standard platform, Softbank's humanoid robot Pepper, and state-of-the-art tasks from the RoboCup@Home social robotics challenge [8], with a relevant use case: how to communicate the robot's knowledge of the situation and its plans for actions.¹ The present study was the first step in this investigation, and the research question was as follows:

* **Corresponding author: Sofia Thunberg**, Department of Computer and Information Science, Linköping University, 581 83, Linköping, Sweden, e-mail: sofia.thunberg@liu.se

Tom Ziemke: Department of Computer and Information Science, Linköping University, 581 83, Linköping, Sweden, e-mail: tom.ziemke@liu.se

¹ A shorter version of this paper has been presented at the workshop on Quality of Interaction in Socially Assistive Robots at the International Conference on Social Robotics in November 2019.

– How can the participatory design method PICTIVE be used to design communication between people and the Pepper robot?

The remainder of this article is structured as follows: first we discuss some background in Section 2, followed by the methodology used in Section 3. The results are divided into three sections (Sections 4–6), corresponding to the three phases (label, sketch and interview), with both method and result presented in each section. Section 7 presents a summary of the results, and Section 8 is an overall discussion of the results and the methodology, followed by conclusions in Section 9.

2 Background

In the following subsections, we discuss some of the cognitive-science and psychology background relevant to how people interpret robots (Section 2.1) as well as the participatory design method PICTIVE used in our study (Section 2.2). In Section 2.3, the RoboCup@Home social robotic competition and the Pepper robot are presented briefly.

2.1 Theoretical background: How I know that you know

To understand how people interpret robots, let us first have a quick look at how people interpret other people. To communicate effectively with others, one needs to have at least a rough idea – or a mental model – of what the other one knows. According to the cooperative principle, or the Gricean view, people do not convey information that others can be assumed to already have [9]. For example, a speaker who overestimates what a listener knows may talk over the listener's head, and a speaker who underestimates what a listener knows may be interpreted as talking down to the listener [10].

Nickerson proposes that one uses one's own knowledge as a basis for creating a model for what others know [10]. According to Nickerson, knowledge includes beliefs, opinions, suppositions, attitudes, and related states of mind. As a general rule, he argues, relatively accurate models of what other people know, believe and feel is preferred over inaccurate ones. However, it would not be difficult 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. There is also evidence indicating that a high

degree of empathic accuracy can endanger the survival of a relationship [11]. According to Eisenberg et al. [12], empathy (or empathic accuracy) has both emotional and cognitive aspects, where focus is on knowledge in the conventional sense. They refer to the cognitive aspects of empathy, the ability to understand others' internal states, as perspective taking.

According to Fussel and Krauss [13,14], in communication people tend to adapt the message to the background knowledge of the recipient. They showed, for example, that people's verbal descriptions of nonsense figures differ depending on who they thought would match the description and the figure later, i.e. themselves, friends or strangers. Figure 1 illustrates Nickerson's view of how one constructs a model of a specific other's knowledge by adapting a default model of what an unknown person knows, considering the information one has about the other person that differs from the default model. "Others" in this context are heterogeneous groups (e.g. people watching the 9 o'clock news), small groups with shared characteristics (e.g. members of an association) or single individuals (e.g. the cashier at the supermarket).

This is a case of the general reasoning heuristic of anchoring and adjustment [15]. According to this, people make judgements by starting with an anchor as a point of departure and then adjust to it. Nickerson means that when people are provided an anchor, they typically adjust

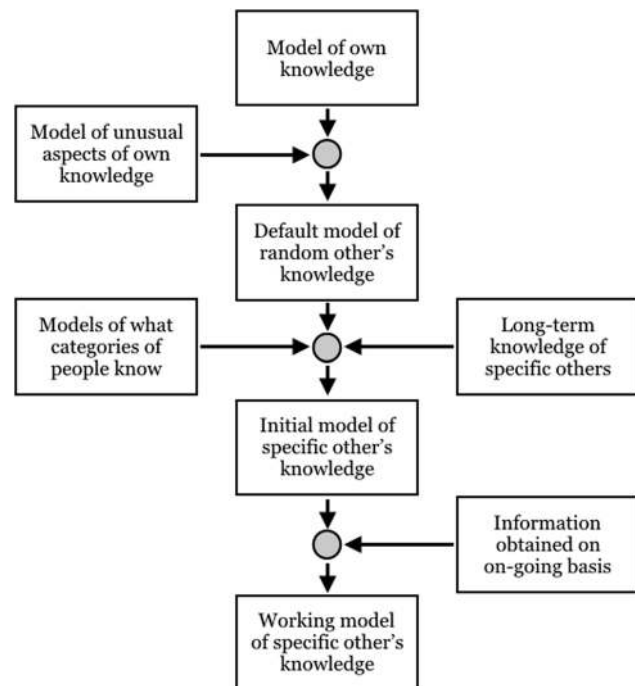


Figure 1: A working model of a specific other's knowledge, adapted from Nickerson (1999).

their judgement in the right direction, but they often overestimate. His model takes this into account and predicts that people tend to overestimate what unknown others know.

The type of mental model suggested by Nickerson is, in our opinion, also relevant to interaction between humans and robots. People's mental model of robots are likely to be formed in a similar way based on their previous knowledge of what robots in general know and how they might be expected to act as well as what specific robots know [16,17]. According to Powers and Kiesler, communication between humans and robots will benefit if people have a clear mental model of what the robot can do [1]. Creating a correct first impression of the robot's knowledge and capacities seems especially important. Nass and Moon [2] reviewed experimental studies providing evidence indicating that people relatively mindlessly apply social rules to computers, and that they have expectations on them. They also showed that people apply stereotypes and social heuristics automatically to interactive systems [18]. There is also evidence for this behaviour in HRI [19]. According to Powers and Kiesler [1], the underlying bottom-up processes are immediate and unconscious. They suggest that there is a parallel cognitive process using structural mapping. For example, when somebody sees a smiling robot, they might retrieve information from the long-term memory, like the way metaphors work. Hence, 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, belonging to the non-social category or as a nice person, in the social category. Together this creates a mental model of a sociable robot. Powers and Kiesler's experimental results show evidence that people create mental models of robots they encounter within the first 2 minutes. The result further shows that the robot's physical attributes affect people's mental model, which can change over time [1].

Common Ground Theory [20] was originally proposed by Clark and Brennan, as a framework for understanding communication between people. According to Kiesler, the main assumption is that communication between people requires coordination to reach mutual understanding [21]. To achieve coordination, for example, between two interaction partners or two ballroom dancers, the partners must have a large amount of shared knowledge or common ground. One of the key elements of the theory is the idea of least collaborative effort, which suggests that when people interact, they try to minimise the collective effort required for mutual understanding [20]. Kiesler believes that there is a

need for common ground also in HRI and that robots should have action plans for creating common ground with people [21]. She believes that the first step for this is for the robot to use social cues to help people create an appropriate mental model of the robot. The robot should also actively try to correct the person's mental model, where possible, and repair damage to the common ground.

The assumption is that people will overestimate what the robot knows and wants, but with the help of the physical attributes, this view can be altered; and the information that is obtained on an ongoing 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 between a user and a humanoid robot by altering the physical attributes of the robot. In this study, we make a first attempt to create design concepts for how the robot communicate to the user what it knows and plans to do, by using a participatory design method involving the potential user from the start. We believe that designs created that way would enable for the creation of a better mental model and common ground, thus facilitating the interaction.

2.2 PICTIVE

In this study, we chose to use the participatory design method called PICTIVE [22]. In the original method, participants reflect on interfaces 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 can provide input about, the target implementation technology. The original vision was that by using low-tech objects, in other words, non-computer representations of system functionality, all participants can contribute their ideas in an easy way. The sessions are usually recorded with video, which allows the session leader to be more engaged and not distance herself from the process by, for example, taking notes. The videos can also provide design documents.

The design sessions in PICTIVE are typically conducted with one participant and the session leader, but it could also be carried out with more participants at once. The two are usually sitting face to face in a secluded room with a table between them. On the centre of the table is the so-called shared design surface. This is a large paper that the session leader has designed beforehand. It should be depicting the intended environment that the product or system that is designed could be interacted with. For example, if the product is a web page, then a picture of a

user sitting in front of a computer could be on the shared design surface. The purpose of the shared design surface is to enable the participant to easily begin to create design ideas. This surface is also where most of the design happens. To the participants disposal there are plastic icons, coloured highlighters, coloured pens, labels (data fields), pop-up events and post-it notes. The plastic icons used could, for example, be icons for WiFi, calling and camera. The labels are also predefined by the session leader and these are adjusted to how the user can interact with the intended product. For the web page example, the labels could be, for example, “type,” “talk” and “scroll.” The participants can then place the predefined plastic icons and the labels on the shared design surface, and they can also create their own with post-it notes.

A PICTIVE session consists of three phases: label, sketch and interview. In the label phase, the participants begin by looking at the shared design surface and creating a first scenario of how the product should interact with the user. In this phase, the participant should explore every aspect and scenario that could happen with the product or in the system, and how the user would interact with it. The participants use the plastic icons and labels to depict how the system will work. When no more ideas are raised by the participant, the session moves on to the sketch phase.

In the sketch phase, the participants are asked to sketch at least three different design ideas for the product or the system (but it could also be three iterations of the same idea). The participants are asked to think aloud while they sketch and explain why they choose different design solutions. They can draw their ideas directly on the shared design surface or on white paper. At the end of the session, the participants are asked to sum up their design ideas; and in the next phase, they are to be tested.

In the interview phase, the participants are faced with typically eight scenarios. These scenarios have been created beforehand by the session leader and are problems that could occur with the product or the system and that the participants' design ideas could solve. The session leader reads the scenarios aloud and the participants should choose one of their ideas that could solve the problem and explain how. Then the participants should rank their ideas as best, fair or worst in handling each scenario, and they should propose changes to improve their designs.

The role of the session leader is to help the participant with the design process but not to come up with design ideas oneself. If the participant gets stuck on one idea the session leader could guide the participant to explore other options, or to ask questions about the tasks the product could solve for the user. The session leader should not tell participants if an idea is impossible to

build or implement, because they should keep an open mind to all ideas that come up.

The reason for using PICTIVE in this study, instead of any other user-centred design method, is the advantage of using pre-defined labels. These can be seen as building blocks that have limited the design space of what is possible before the session even starts. This is especially useful when the participants are not designers or have previous knowledge of design processes. In comparison, methods like *cuta* [23], building scrappy prototypes or role playing, may be a larger creativity step for the participants to take than what they are used to. With PICTIVE, the participants are guided by the researchers' shared design surface, the labels and the scenarios to provide room for creating a suitable design. In this study, the participants were potential future users of the robot and therefore PICTIVE was chosen. PICTIVE has also previously been used in HRI in a study investigating how autonomous cars can communicate intent and awareness [24]. Mahadevan, Somanath and Sharlin used PICTIVE as their design method and developed different designs that were later tested “out in the wild” [25].

Others have used participatory design methods, and especially focus groups [26,27], in the development of new humanoid robots. But it is not as common to use as a method for behaviours on existing platforms, such as the Pepper robot. It is more limited to work with the communication modalities available on a standard platform than it is to develop new ones. Some studies have been carried out with displaying emotional non-verbal cues on Pepper [28], but the cues (e.g. happiness, fun and joy) were not tested in a user-centred way but grouped as closest in valence and arousal. The researchers had the Pepper robot recommending films and at the same time displaying a coherent or incoherent emotion in relation to the genre. The study did not find a significant difference in Pepper's coherent and incoherent emotional behaviour, and the authors argued that this was because the participants had already seen all movies. Another interpretation could be that the cues were not displaying the right emotion, which could possibly be corrected by testing which communication modality should be used with which emotion in a user-centred way, instead of using Softbank's pre-programmed alternatives.

2.3 RoboCup@Home SSPL

RoboCup@Home is the largest international annual competition for autonomous robots, aiming to develop service and assistive robot technology with high relevance for future



Figure 2: A typical arena used in RoboCup@Home.

personal domestic applications [29]. 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 2, there is an example of what the arena can look like, and in Figure 3, some of the objects that can be included are shown.

The SSPL competition consists of two stages and one final round. The tasks are complex and can be solved in several ways, especially by using HRI to make the referees help the robot. For example, if the robot asks a person to do something for it (put a bag in the robot's hand), the person must help the robot, to a certain limit. This creates situations where people need to cooperate with the robot and the robot can make use of that to get a higher score. Therefore, these tasks were chosen to create realistic and useful scenarios. For a full explanation of the tasks, see the RoboCup@Home rulebook [8].



Figure 3: Objects typically manipulated by the robot.



Figure 4: Softbank's Pepper robot.

2.3.1 Pepper robot

Since the launch of the Softbank's Pepper robot in 2014, 12,000 of them have been sold [30]. The humanoid robot is one of the most commercialised humanoid robots worldwide and it is used in a number of different areas, e.g. guiding passengers in airports, selling coffee at Nescafé stores [31] or activating elderly at residential homes [32]. The robot is described as a social service robot and it is the social standard platform in the RoboCup@Home competition [29], where it is supposed to perform tasks such as taking out the garbage, working in a restaurant as a waiter and acting as a party host [33].

The Pepper model used in this study was the 1.6, version 2.5.10. (see Figure 4). The robot is 120 cm tall and weighs 28 kg. It is equipped with a gyro sensor and a 10.1-inch touch display on its chest and four 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. Furthermore, the robot has two speakers, one on each side of the head. The display, LEDs and speakers were used the most in this study.

3 Method

Five participants (*mean age* = 26 years, 40% female) were recruited to the participatory design study, based on a convenience sampling strategy. They were all students at Linköping University in Sweden, with different curricula:

teaching, cognitive science, environmental science, theoretical philosophy as well as applied physics and electrical engineering. Two of them had previous experience of interaction with the Pepper robot.

As discussed in Section 2.2, the specific method used was PICTIVE [22]. A sketch showing the Pepper robot standing in a typical living room was used as the shared design surface (see Figure 5) presented to the participants on an A3 paper. To facilitate participants' creation of interfaces, ten labels depicting communication modalities were created. The labels were inspired by Mahadevan et al. [24] and were customised according to the robot's limitations. The predefined 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. In addition to these labels, the participants were provided with coloured pens, post-it notes, glue, pencil sharpener, eraser, tape, scissors and white paper.

The study sessions, approximately 60–90 minutes each, took place in a secluded room. The setup (see Figure 6), was that one participant at a time, and the session leader, sat at opposite sides of a table. A Pepper robot was also physically present to provide the participant a better view of how the different functions could work on the robot. The robot was turned on but did not move. The session started with a short presentation of the robot and the purpose of the study. Furthermore, age, gender, study background and experience of autonomous robots were collected. The study procedure contained three phases: the label phase, the sketch phase and the interview phase. The specific procedures used in each phase are further described in Sections 4–6, where both the method and the results are presented.

From each study session, participants' designs and video recordings of the session were collected. The video



Figure 5: The shared design surface.



Figure 6: The session setup. In the middle of the table facing the participant is the shared design surface. On the left side are the predefined labels. On the right side are creative tools. On the top of the table is the camera facing the shared design surface. On the side of the table is a Pepper robot.

dialogues were transcribed and the material was open coded [34], which means that one writes keywords in the margin to mark what is important (e.g. aware of presence, greeting, blinking). In the next step, the keywords were put together to identify themes (e.g. meeting the robot for the first time). After creating these themes, the first author did a focused coding by going through the material again using more general codes and only those that belong to one of the themes (e.g. first encounter).

4 Label phase

In the label phase, participants were encouraged to use the predefined labels (motion, haptic feedback, sounds, voice synthesis, LED's in eyes, ears and shoulders, text, animation and symbols on the display) on the shared design surface to map different design solutions (see Figure 7). To start the creative session, they were provided with a scenario where the participant was in a living room. While starting with that very open scenario, the participants continued with brainstorming possible interactions with the robot and which actions the possible user (themselves) would like the robot to be able to carry out. For every scenario they created, the session leader encouraged them to describe how they wanted the Pepper robot to communicate what the robot knows and what it plans to do. When the participants said that they had explored every option they could think of, the session moved on to the sketch phase. The data analysis from the label phase identified four themes that will be described below: the first encounter, on a mission, screensaver mode and need of assistance.



Figure 7: Example of what the shared design surface can look like during the label phase.

4.1 The first encounter

Some participants' first idea was that they wanted the robot to move, just a little so they would know that it is on and could be interacted with. It was also considered important that the robot is aware of the person's presence in the room for a natural interaction. If a person needs to draw the robot's attention in the first encounter, it might cause a feeling of confusion about the robot's purpose and could lead to doubts about its usefulness. All participants therefore suggested that the robot should verbally introduce itself and its capabilities. This would also indicate to the user that the robot is aware of the presence of the person. By explicitly pointing out that it sees the user, the user knows that the robot knows that a person is in the room. It was further suggested that the robot should blink its eyes when it sees a person for the first time.

4.2 On a mission

When the robot is on a mission (engaged in a task), two participants wanted the robot to show pictographs on the display, which would serve as the robot's own language. They wanted the robot to communicate its plans in a sequence, for example, if the task is to greet someone at the door, the pictographs could, for example, display "go," "person," "door" and "say hello." The robot should reuse familiar symbols, to make the user feel more comfortable with the pictograph system and the robot interface. Four participants said that there also should be a menu system on the robot's display with different tasks marked with both symbols and text. It should also have a depth, so when the user chooses one task it jumps to

different alternatives. They also thought that the user could ask the robot by voice or by touching the display. In contrast, one participant said that the user should only be allowed to use the display, since it is "unnatural" to talk to machines.

4.3 Screensaver mode

There were some different ideas from the participants on how the robot should behave when the robot is not involved in a task and the user does not need it. The robot should not talk spontaneously after the introduction, and some 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 considered an option that the robot would move and look around by itself, because that could create an unpleasant situation which might distract the user. One of the participants suggested that the robot should move to the corner of the room, or to an adjacent room, stand still and look at the floor.

4.4 Need of assistance

If the robot gets stuck and needs help, such as if there is something in its way, all participants wanted the robot to verbally ask the user for assistance. They thought that the robot should verbally say once that something has happened and that it needs help, but thereafter it should only blink so it will not disturb the user too much. They further thought the robot should blink with blue lights if it gets stuck in a task. One participant suggested that it should send a notification to the user's phone, so that the user can choose to help it when convenient. One participant suggested the use of pictographs here; that the robot does not need to say anything but show for example boxes on the display and a symbol for aid, and in that way communicate to the user that it needs to help the robot move boxes obstructing its path.

5 Sketch phase

In the sketch phase, participants were asked to create at least three unique interface designs, and they could use the shared design surface or blank paper. The predefined labels were at hand, and they could also use their own if

they had created others during the label phase. Their sketches could be iterative and based on the same design idea but in different versions. During this phase, they were encouraged to describe their thoughts using a think-aloud protocol; and when they felt done with their sketches, they were asked to describe and summarise their sketched designs.

The participants sketched in total 25 different design ideas that could be grouped into three themes: menu interface, pictographs and map (see Figure 8 for examples). The participants wanted the menu system to show which tasks the robot know how to perform. The menu interface could be on the display, showing different tasks that the robot could help the user with. These task should have a depth in the system, so the user has several options, for example, choosing between channels on the TV. For one of the participants, who did not like to talk to machines and therefore preferred using the display, the menu system was the main way of interacting with the robot. The others thought of the menu system as a backup and a visualisation of what the robot said. For example, if the robot told the user that it could turn on the TV the first time they met, it could be a good reminder to have the task on the display for next time.

The pictograph system, where the robot shows the different steps of a task with symbols on the display, is a design idea for how the robot could show what it knows in a situation and its plans. The pictographs would be



Figure 8: Four examples of sketched designs made by participants. From left to right, top to bottom: a menu system on the display with different actions, the robot could help the user with (e.g. turn on the TV), together with sub-categories (e.g. Netflix, YouTube). A pictograph system where the robot shows pictures on the display of the sub-tasks in a task that it plans to perform. A map system on the display that shows the robot's position and goal, here with the use of the label voice synthesis. A combination of a map system on the display with the robots position and goal, and pictographs at the bottom of what the robots' mission is.

pictures used as a “robot language” where the robot broke down the task into sub-tasks and showed its intentions through pictures. Instead of expressing verbally what it was planning to do, the robot could easily show it on the display. For example, if the task was to pick up a box in the bedroom and bring it to the user, then the display would show the main task “robot hand over box to user” with pictures “robot, hand over, box, person.” Below the main task this would be split up into pictures of sub-task, for example, “robot, walk, bedroom, box, pick up, walk, hand over, person.” The participants described that the reason to have a robot language was that the user could easily see what the robot was planning to do and either let it execute the plan or stop it. It would also be easy for the user to see where in the sub-tasks the robot was, and to get an idea of when the robot is finished and is available for a new command. They thought that this would make the interaction more effective.

If the task included moving somewhere, the robot would show a map on the display with its current position and that it knows the tasks goal position. For example, if the robot needed to charge, it could show the charging station's position on the map, a planned pathway there from its current position and then start moving toward its goal. The participants thought this would also make it easy for the user to see what the robot is planning and why it is moving. The next step with the map could be to display knowledge of different things that are happening in the robot's surroundings. For example, if it sees that somebody is standing in the way, it can show this on the map, with the hope that the person notices and understands that the robot has the plan of moving in her direction, and then they could move out of the way. But if the person does not notice this, then the robot can take the time to plan a route to go around the person and display this plan on the map.

These themes could also be combined. For example, if a user asks the robot to “go get water” from the menu system, then the robot could show a map on the display with the kitchen sub-goal, and the main goal of the position of the user. It could also have pictographs at the bottom of the screen that illustrates the robot's plan for how to perform the task.

6 Interview phase

In the interview phase, participants were presented with eight scenarios: Cocktail Party, General Purpose Service Robot, Help-me-carry I, Help-me-carry II, Speech and Person Recognition, Enhanced Endurance General Purpose

Service Robot, Restaurant and Tour Guide. In these scenarios, the Pepper robot had to communicate what it knew of a situation and how it would plan actions ahead. The scenarios were created through observations from the RoboCup@Home competitions 2018 as well as from the rule book [8]. Coding of the notes from the observations and the rule book was carried out the same way as previously described for the video dialogues. The scenarios were parts of some of the different challenges in the competition and were chosen to represent different tasks that are difficult for robots to perform and that teams in the competition struggled with. The purpose of this was to challenge the participants sketched design solutions with state-of-the-art task that a communicative humanoid robot would be able to perform in everyday situations.

The scenarios will be further described in the following section, but it might be worth noting here that they dealt with different aspects of the robot communication. What the robot planned to do in a situation was addressed in the Cocktail Party, the Help-me-carry I and the Tour Guide scenarios, and the scenarios dealing with what the robot knew about a situation where the General Purpose Service Robot, the Speech and Person Recognition and the Enhanced Endurance General Purpose Service Robot scenarios. The scenarios Help-me-carry II and Restaurant addressed both what the robot knew about the situation and what it planned to do.

In the interview phase, the participants were asked to choose one of their sketched designs that could solve the problem in the scenario and rank their designs from the sketch phase as a best, fair or worst fit in handling each scenario and propose any changes that could improve their interfaces.

6.1 Scenarios

The eight scenarios will be exemplified below, in the way they were read by the session leader to the participants (freely translated from Swedish to English), followed by a short explanation of what the scenario represents and examples of the participants' suggestions to solve the scenario problems. The participants were free to ask follow-up questions to clarify the situation in the scenarios.

6.1.1 Cocktail party

The problem presented in this scenario is that the robot has discovered an obstacle that it cannot move itself and

that it needs help from a person. The robot needs to communicate to the user its plan to enter through a door that is not fully opened.

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

Examples of the participants' suggestions 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 LEDs. 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.”

Notice that already in the first scenario, most of the participants do not suggest their own sketched design. One participant wants to use the pictograph design, but the others preferred the robot to be more embodied in its communication to the user. They want the LEDs to blink in different alarming ways or show symbols on the display of the robots' emotional state. It should also make different sounds, for example, a beep or asking verbally for help.

Another interesting observation is that it was the participants without a technical background who envisioned more advanced technology and less natural interaction. They wanted the robot to use its LEDs and send notifications to the users' smartphone, while the other two participants were more interested in simple cues, such as a beep and a voice command.

6.1.2 General purpose service robot

The main purpose of a social service robot is that the robot performs tasks for the user. This scenario represents the

issue of how the robot should communicate to the user that it has understood the command for the task correctly.

A human tells Pepper that it is supposed to carry out 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?

Examples of the participants' suggestions 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 provides 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 an 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."

In this scenario, two of the participants wanted the suggested sketched theme – pictographs – but also a new use of the display was suggested. One participant wanted the command that the robot heard to be displayed in text on the display for the user to read and correct if necessary, and one wanted the robot to search on the internet for common emojis or icons that could represent the task on the display, for example, show an apple.

6.1.3 Help-me-carry I

A common situation for a service robot can be to follow a person to perform a task at a location the robot does not know. What the user might not be aware of is that for the robot to be able to do this it needs to learn what the person looks like, which is a process that could take some time. This scenario represents such a situation and how the robot should communicate this need.

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 what could such a process look like)?

Examples of the participants' suggestions were as follows:

- "It should be symbolised with a pictograph on the tablet. The robot 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 goes 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."

In two of the suggested solutions the participants mentioned some kind of tracking device, which had not come up before in the previous phases. Also, one of the participants wanted the robot to use technical terms to describe its problem (e.g. "scan"), while two others wanted the robot to use folk psychological terms (e.g. "need to get to know") and display human emotions. One observation was that there seems to be a difference between the people who want the robot to act humanlike and those who want it to act machine-like.

6.1.4 Help-me-carry II

In situations when the robot is performing a task for a user and other people asks it to help them, it needs to handle it in an appropriate manner. In this scenario, the robot needs to communicate that it is busy and that it could help perhaps later or not at all.

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, who is in the kitchen, to 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 robot's 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 person that gets in its way)?

Examples of the participants' suggestions 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.”

One participant suggested the map solution from the sketch phase to show on the tablet, showing the goal with the mission, which implicit could indicate to the user that the robot is busy and when the it is done with the task. The other participants had different solutions for how the robot could verbally communicate with people and in combination with the pictograph solution.

6.1.5 Speech and Person recognition

When a robot is in a crowded place, it needs to react to the person talking to it and communicate that it is listening. If not, the user is unable to understand when it can give a command. This kind of situation is represented in this scenario in the form of a game.

Pepper is in a question game. The robot is standing in the middle and people are standing in a circle around the robot. When a person asks the robot a question, the robot is supposed to answer the question. How can the robot communicate which person it heard the question from?

Examples of the participants' suggestions were:

- “The robot seeks 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 person’s 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 had asked the question with combinations of verbally saying the person’s name and seeking eye contact, which is similar to how people communicate that they are listening to each other. This had not come up in the sketch phase; and as with the first scenario, the participants are suggesting an embodied solution.

6.1.6 Enhanced endurance general purpose service robot

There might be a difference in how experienced users of a service robot express themselves compared with the first-time users. In this scenario, it is explored how the robot could handle this situation and how the robot would communicate that it does not understand what the person is saying.

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 repeats 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 how it should provide suggestions if it wants to speak to a person it knows or read from a code?

Examples of the participants' suggestions were:

- “There should be a pictograph and a text on the tablet with 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 thinks 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.”

The most interesting observation for this scenario was that some of the participants created a scenario by themselves by adding how the robot should communicate that it cannot hear because of a noisy environment. The suggestions were to communicate to the user using text on the tablet that it cannot hear and that the users could generate, for example, a QR code for the robot to read instead.

6.1.7 Restaurant

This scenario represents a situation where the robot is working as a waiter in a restaurant. When a customer gives the signal for the robot to come to their table, this could be picked up by several robots working there. The problem here is how the robot should communicate that it will go take the order.

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

Examples of the participants' suggestions were the following:

- “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.”

In this scenario, the participants did not agree that the situation actually could play out as described in the scenario description. In the real world, the robots should talk to each other or have an ordering system that makes all the decisions. Therefore, the ideas for how the robots

should communicate to the main waiter regarding which robot should take the order were felt unnecessary to the participants, and accordingly the result here is not as thought out as for the other scenarios.

6.1.8 Tour guide

The final scenario represents an issue of how the robot should be able to communicate that it wants to greet people in a culturally appropriate way.

Pepper is a tourist guide and meets a new group of tourists. The robot introduces itself and is then supposed to greet people. How can the robot communicate which way it will greet (wave, shake hand, bow)?

Examples of the participants' suggestions were as follows:

- “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.”

Also in this scenario the participants had a hard time to come up with ideas since this scenario was considered a bit unnatural. One participant thought that it was a good thing that the robot would be sufficiently considerate to not offend anyone with the wrong type of greeting. Another suggested that the robot should not adapt to the human way of greeting at all but instead develop a robot way of greeting. The others thought that the robot should “just do it” (greet in any way without considering the norm). None of these suggestions is really solving the problem in the scenario though. It might be the case that the task itself is developed in an unnatural way and should be revised, or that the participants did not see the value of a robot adjusting for different cultures.

7 Summary of results

The labels used by participants are shown in Table 1 along with the frequency of their use across the eight

Table 1: The design labels used in the interview phase, by the number of participants, across the eight scenarios

Label	1	2	3	4	5	6	7	8
Motion			1	3	5			2
Haptic	1		2					
Display text	1	1		2	1	4	2	
Display animation			1	2				
Display symbols	2	3	1	3		2	2	
Speaker	2		1					
Voice synthesis	4	3	2	4	5	3	4	4
LED eyes	3				1			1
LED ears	3							
LED shoulders	3							

scenarios. There is a great diversity across the different scenarios in how people want the robot to communicate its plans and knowledge. This kind of customisation suggests that it might be difficult to find a universal “golden way” design solution. For example, in scenario five, all participants wanted the robot to use motion and voice synthesis (i.e. speech), but also in this scenario two other labels were used. It was not surprising that voice synthesis was the most used label, 29 times, and for comparison, the second most used label was display symbols, which were used 13 times. For people, these are two of the most common ways of interacting. Speech is the most natural way for people to communicate with each other. However, previous research of ours with the Pepper robot in a public environment showed that people did not find it natural to talk to a robot [35]. In this study, people preferred to use the display for interacting with the robot even if the display did not show any symbols, text or animation. This might indicate a difference between how users think they want to interact with robots and how they try to interact with them “in the wild.”

Furthermore, we found that the label voice synthesis was never chosen alone; for all the scenarios, the label was used in combination with other labels, and it was often used as a way of making it more explicit to the user what the robot planned or knew. This indicates that voice synthesis is seen as the main way of communicating what the robot knows and plans, but that it should be used with support from other communication modalities.

Although the participants had a broad view of when to use different labels in the label phase, their sketched designs created three themes: menu system, pictograph and map. Most of the sketched design suggestions were focused to the display and exclude other suggestions of, for example, movements and haptic feedback. When the participants later described the sketched designs, they included the robot’s embodiment with, for example,

blinking lights, but they found it difficult to sketch these details. The reason for this might be that the participants felt insecure with sketching and found it easier to keep working with the display.

The diversity of labels for the scenarios shows that the Pepper robot could make use of all of its communication modalities when interacting with the user in different scenarios. The four themes that the participants created for the Pepper robot – the first encounter, on a mission, screensaver mode and need of assistance – could be new scenarios or be combined with the existing scenarios. For example, need of assistance is closely related to the cocktail party scenario, and on a mission is part of general purpose service robot and help-me-carry II. For the next iteration, the restaurant and tour guide scenario could be removed and replaced with scenarios involving the first encounter and screensaver mode, which could contribute to better integrating the potential users’ wants and needs in the design process.

The next step of the design process would be to create prototypes of the robots’ communicative behaviour. We identify four different prototypes that would be of interest to test against each other: the use of a silent pictograph language on the display, the use of a map system on the display, the use of a menu system on the display, and a more embodied version making use of LEDs, the speakers, motion and haptic feedback in situations suggested by the participants. This would also facilitate the option of testing whether verbal or non-verbal cues are the most successful way of interacting, given that the participants had very different views regarding whether the robot should talk or not.

8 Discussion

The study illustrates the use of a participatory design method to investigate how a humanoid robot can communicate what it knows and what it plans to do. It shows how a user-centred design approach engages participants to think about how they might want to interact with robots. There was a high degree of agreement among participants regarding some of the proposed design solutions, and the study generated a large amount of data that could be used to create prototypes. The number of five participants was, in our opinion, sufficient for the task because the same kind of design ideas and solutions were mentioned by several of the participants. The use of five participants in design studies is also in accordance with previous research stating that additional subjects

are less likely to give new information [36,37]. At the same time, the differences in some of the solutions to the scenarios highlight the challenges for HRI designs aiming for a general user. In fact, the participants pointed out that it was important that the robot could be adjusted to its main user; for example, if the main user is having a hearing problem, it should not talk but use other ways of communicating. In the following, the results and the method are discussed in more detail.

8.1 Results

The label phase generated four themes: the first encounter, on a mission, the screensaver mode and the need of assistance. In the first encounter, the participants focused on a robot that should show awareness that the potential user is in the room. This means that when someone enters 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 in line with Powers and Kiesler's arguments that the first impression that the robot creates is especially important. This is the robot's chance to make the person's 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 [1]. 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 listens 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 this out as especially important to reach least collective effort for interaction [21].

When the robot is on a mission, and showing its plan using either the pictograph system or the map system (or a combination) it is trying to help the user create a correct mental model of the robot's knowledge and plan [10]. Initially, however, the user may overestimate what the robot knows but by illustrating the task step by step, or showing the goal point on the map (and also including its knowledge of its environment on the map) the robot can help users to adjust their mental models on an ongoing basis.

At times when the robot is not needed (in screensaver mode), the participants wanted the robot to stay out of sight, but still nearby, and that it should look down to the ground. This could be considered 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 of the user. As Nass and Moon showed, people apply social rules to computers, as well as robots, and have expectations regarding behaviour [2]. When the robot is not needed, the user might feel uncomfortable if the robot stayed close to them and looked at them. This would be roughly the same in human-human interaction. Therefore, the expectation is that the robot should move away to not break any social contract.

The main purpose of the robot is that it should help people and assist the user in different ways. But sometimes there will be situations when the robot needs 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 or move. Then the robot needs to call for help, communicate its own plans and what the human could do to help it fulfil its task. This is a situation that really requires a correct mental model of the robot. If the person hears or sees it calling for help and believes that the robot knows how to solve the situation by itself, then the mental model is incorrect. The user in that case has overestimated the robot's knowledge and the robot will not be able to continue (and eventually run out of battery) [10]. For this situation common ground is especially important, because it would enable the user to understand the robot's problem and help it solve the problem [21].

8.2 Method

This study was an attempt to use participatory design to meet the HRI problem of how to communicate a humanoid robot's knowledge and plans to the user. The specific method used, PICTIVE, in our opinion, 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, four concepts are used to draw from to build prototypes: pictographs, maps, menu system and an embodied version. By going through three phases, label, sketch and interview, the participants got a deeper understanding of the issues that could occur if the robot cannot communicate its plans and knowledge in each situation. Using this setup, they were all engaging in the task and gave 60–90 min of their time voluntarily, which indicates that they all could see the value of their input. During the label phase, all participants were at first confused about 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 occur. In the sketch phase, it was challenging for some of the participants to produce at least three different designs (some made three iterations instead of one), and most of them got stuck on sketching interface designs for the display. But there were also participants who produced more than three designs, so in the end there were a lot of different ideas. However, for the use of PICTIVE in the future, it would be recommended to clarify the participants the use of the designs and that they could have a great variety. The interview phase really tested the participants' ideas and view of how the robot could communicate, and it was a good way of investigating when and why they would like to correct their design to solve the scenario better. In sum, this was a method that worked well for the purpose and could be used in the future studies, with different humanoid robots or other social entities.

The method of combining coding of the RoboCup@Home rule book and observations from the competition to create scenarios, was, in our opinion, a good way of identifying realistic tasks that social service robot researchers are struggling with all over the world. However, some of the scenarios failed to create realistic stories for the potential users. Some of the participants had complaints about the restaurant scenario and the tourist guide scenario. The difficulties appeared for the participants when the scenarios did not seem applicable in a real-world setting, which illustrates the importance of involving the user as early as possible when developing a robot to create realistic scenarios to work with the design process. For future studies, we recommend that researchers should pilot test the scenarios used to early on to discover the difficulties in the situations that are presented to participants. What could be considered a realistic and natural scenario is not necessarily the same for robot developers and for potential users.

9 Conclusion

The aim of this study was to investigate whether and how user-centred design – and PICTIVE in particular – could be used to develop human-robot interaction. We specifically used the robot's communication of its knowledge and plans in a situation as use cases, which is highly relevant even for the simplest interactions between humans and robots.

In sum, participatory design was found to be a suitable method to generate design concepts in HRI, and we have shown step by step how other researchers and practitioners could use PICTIVE when developing robot communication in a user-centred way. The three phases of PICTIVE played

different roles in the development of design ideas; the label phase made it possible for the participants to be creative, the sketch phase required them to be more concrete about their ideas and, finally, the interview phase challenged their sketched designs. This resulted in four prototype concepts ready to be implemented. The frequent use of the communication modalities through all phases, and the fact that the participants designed for the robot to communicate what it knows and plans, indicate that the physical attributes of the robot could help the users to adjust mental model of the robot's knowledge and plans, which is in line with the theories discussed in Section 2.1. This was the first step of investigating how to design this communication on a Pepper robot, and for future work, the design ideas and scenarios could be implemented on the Pepper robot and tested with live interactions. We encourage others to further investigate whether and how humanoid robots' physical attributes and behaviour influence users' mental model of the robot's knowledge and plans.

References

- [1] A. Powers and S. Kiesler, "The advisor robot: tracing people's mental model from a robot's physical attributes," in *Proceedings of HRI '06 – ACM SIGCHI/SIGART Conference on Human-Robot Interaction*, 2006, pp. 218–225.
- [2] C. Nass and Y. Moon, "Machines and mindlessness: Social responses to computers," *J. Soc. Issues*, vol. 56, no. 1, pp. 81–103, 2000.
- [3] B. J. Scholl and P. D. Tremoulet, "Perceptual causality and animacy," *Trends Cognit. Sci.*, vol. 4, no. 8, pp. 299–309, 2000.
- [4] A. Powers, A. D. I. Kramer, S. Lim, J. Kuo, S. L. Lee, and S. Kiesler, "Eliciting information from people with a gendered humanoid robot," in *IEEE International Workshop on Robot and Human Interactive Communication*, 2005, pp. 158–163.
- [5] A. Powers, S. Kiesler, S. Fussell, and C. Torrey, "Comparing a computer agent with a humanoid robot," in *ACM/IEEE International Conference on Human-Robot Interaction*, 2007, pp. 145–152.
- [6] M. R. Endsley, "Theoretical underpinnings of situation awareness: A critical review," in *Situation Awareness Analysis and Measurement*, M. R. Endsley and D. J. Garland, Eds., Mahwah, NJ: Lawrence Erlbaum Associates, 2000, pp. 3–32, DOI: 10.1016/j.jom.2007.01.015.
- [7] D. Benyon, *Designing Interactive Systems: A comprehensive guide to HCI, UX and Interaction Design (3rd Edition)*, Edinburgh: Pearson, 2014.
- [8] M. Matamoros, C. Rascon, J. Hart, D. Holz, and L. van Beek, *RoboCup@Home 2018: Rules & Regulations*, 2018, <http://www.robocupathome.org/rules>.
- [9] H. P. Grice, "Logic and Conversation", in: *Syntax and Semantics, Vol. 3, Speech Acts*, P. Cole and J. L. Morgan, Eds., New York: Academic Press, 1975, pp. 41–58.

- [10] R. S. Nickerson, "How we know – and sometimes misjudge – what others know: imputing one's own knowledge to others," *Psychol. Bull.*, vol. 125, no. 6, pp. 737–759, 1999.
- [11] A. L. Sillars, G. R. Pike, T. S. Jones, and M. A. Murphy, "Communication and understanding in marriage," *Hum. Commun. Res.*, vol. 10, no. 3, pp. 317–350, 1984.
- [12] N. Eisenberg, B. C. Murphy, and S. Shepard, "The development of empathic accuracy," in *Empathic Accuracy*, Guilford Press: New York, 1997, pp. 73–116.
- [13] S. R. Fussell and R. M. Krauss, "The effects of intended audience on message production and comprehension: Reference in a common ground framework," *J. Exp. Soc. Psychol.*, vol. 25, no. 3, pp. 203–219, 1989.
- [14] S. R. Fussell and R. M. Krauss, "Understanding friends and strangers: The affect of the audience design on message comprehension," *Eur. J. Soc. Psychol.*, vol. 19, pp. 509–525, 1989.
- [15] A. Tversky and D. Kahneman, "Judgment under uncertainty: Heuristics and biases," *Science*, vol. 185, pp. 1124–1131, 1974.
- [16] X. Zhao, C. Cusimano, and B. F. Malle, "Do people spontaneously take a robot's visual perspective?," in *HRI'15 Extended Abstracts: Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction Extended Abstracts*, 2015, pp. 133–134, DOI: 10.1145/2701973.2702044.
- [17] S. Thellman, A. Silvervarg, and T. Ziemke, "Folk-psychological interpretation of human vs. humanoid robot behavior: Exploring the intentional stance toward robots," *Front. Psychol.*, vol. 8, art. 1962, 2017, DOI: 10.3389/fpsyg.2017.01962.
- [18] C. Nass and K. M. Lee, "Does computer-synthesized speech manifest personality? Experimental tests of recognition, similarity-attraction, and consistency-attraction," *J. Exp. Psychol.: Appl.*, vol. 7, no. 3, pp. 171–181, 2001.
- [19] J. A. Bargh, M. Chen, and L. Burrows, "Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action," *J. Personality Soc. Psychol.*, vol. 71, no. 2, pp. 230–244, 1996.
- [20] H. H. Clark and S. E. Brennan, "Grounding in Communication," *Perspect. Socially Shar. Cognition*, 1991, pp. 222–233.
- [21] S. Kiesler, "Fostering common ground in human-robot interaction," *Proceedings of the IEEE International Workshop on Robot and Human Interactive Communication*, vol. 2005, pp. 729–734, 2005.
- [22] M. J. Muller, "PICTIVE – an exploration in participatory design," *SIGCHI Conference on Human Factors in Computing Systems*, 1991, pp. 225–231.
- [23] D. Lafrenière, "CUTA: A simple, practical, low-cost approach to task analysis," *Interactions*, vol. 3, no. 5, pp. 35–39, 1996.
- [24] K. Mahadevan, S. Somanath, and E. Sharlin, "Communicating awareness and intent in autonomous vehicle-pedestrian interaction," in *CHI Conference on Human Factors in Computing Systems*, 2018, pp. 1–12.
- [25] E. Hutchins, *Cognition in the Wild*, Cambridge, MA: MIT Press, 1995.
- [26] F. Förster, A. Weiss, and M. Tscheligi, "Anthropomorphic design for an interactive urban robot – The right design approach?," *HRI 2011 – Proceedings of the 6th ACM/IEEE International Conference on Human-Robot Interaction*, 2011, pp. 137–138, DOI: 10.1145/1957656.1957699.
- [27] J. Osada, S. Ohnaka, and M. Sato, "The scenario and design process of childcare robot, PaPeRo," *International Conference on Advances in Computer Entertainment Technology*, pp. 80–86, DOI: 10.1145/1178823.1178917.
- [28] S. Rossi, T. Cimmino, M. Matarese, and M. Raiano, "Coherent and incoherent robot emotional behavior for humorous and engaging recommendations," *2019 28th IEEE International Conference on Robot and Human Interactive Communication, RO-MAN 2019*, DOI: 10.1109/RO-MAN46459.2019.8956260.
- [29] RoboCup Federation, RoboCup@Home, 2019, <http://www.robocupathome.org/>.
- [30] P. Olson, "Softbank's robotics business prepares scale up," 2018, <https://www.forbes.com/sites/parmyolson/2018/05/30/softbank-robotics-business-pepperboston-dynamics/>.
- [31] Nestlé, "Nestlé to use humanoid robot to sell Nescafé in Japan," 29 October 2014, <https://www.nestle.com/media/news/nestle-humanoid-robot-nescafe-japan>.
- [32] M. Foster, "Aging Japan: Robots may have role in future of elder care," 2018, <https://www.reuters.com/article/us-japan-ageing-robots-widerimage-idUSKBN1H33AB>.
- [33] M. Matamoros, C. Rascon, S. Wachsmuth, A. W. Moriarty, J. Kummert, J. Hart, et al., *RoboCup@Home 2019: Rules and Regulations*, 2019, http://www.robocupathome.org/rules/2019_rulebook.pdf.
- [34] B. Alm, *Compendium in Ethnographic Method*, vol. 2.0, 2017.
- [35] S. Thunberg and T. Ziemke, "Are people ready for social robots in public spaces?," in *HRI '20: ACM/IEEE International Conference on Human-Robot Interaction*, March 23–26, 2020, Cambridge, United Kingdom, ACM, New York, NY, USA, 2020, pp. 482–484, DOI: 10.1145/3371382.3378294.
- [36] R. A. Virzi, "Refining the test phase of usability evaluation: How many subjects is enough?," *Hum. Factors*, vol. 34, no. 4, pp. 457–468, 1992, DOI: 10.1177/001872089203400407.
- [37] J. Nielsen, "Why you only need to test with 5 users," 2000, www.useit.com/alertbox/20000319.html.