D3.1 – Normative Behavior Report

Project acronym: *TERESA*

Project full title: Telepresence Reinforcement-Learning Social Agent

Grant agreement no.: 611153

Due-Date:	15
Delivery:	Month
Lead Partner:	University of Twente (UT)
Dissemination Level:	PU
Status:	Submitted
Version:	1.0

DOCUMENT INFO

Date and version number	Author	Comments
15.09.2014 v0.1	Jered Vroon	First version; literature overview and report on the observations we conducted during various activities and while members of the target group were controlling the robot.
24.02.2015 v0.2	Jered Vroon	Added description of the inductive user study.
26.02.2015 v0.3	Jered Vroon	Added general conclusion.
27.02.2015 v1.0	Jered Vroon, Gwenn Englebienne, Vanessa Evers	Final version.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

1 Executive summary

This deliverable is aimed to report on our investigations into what telepresence robot behaviors are experienced as socially normative.

First, to get a better impression of the context in which the robot will be to navigate itself, we conducted several passive observations at 3 locations. Specifically, we investigated different kinds of social situations, the applicability of proxemics to elderly (in interaction and in navigation), and other non-verbal social signals used by elderly. We focused our observations on social situations with various elderly in the same room, because those allowed us to study both one-on-one and group interactions. We observed a great diversity in (proxemics) behaviors between elderly. Some got very close to each other, others stayed further away. These differences often seem to be intended to accommodate for individual variables, such as hearing problems. In caregivers we saw this same variation and we expect that for them the additional closeness may be part of their motivation too.

In September and October 2014, the consortium partners conducted collaborative experiments in the Living Lab in Troyes. We designed these experiments such that they covered approach, conversation and retreat behavior with the telepresence robot in a wide range of situations (seated interaction target, standing interaction target, moving interaction target). The resulting data shows a variety of behaviors and challenges in using the robot that may be specific to our target group. Among other things, we observed with all of our participants some difficulty in steering the robot. The required training times varied strongly (20 minutes – 1 hour). And even after training, most of the participants seemed to be unable to control the robot and have a fluent conversation at the same time

In November 2014 a study was performed to collect a large amount of quantitative data on groups of students interacting. In this study, groups of three participants and the Teresa robot (controlled remotely by a fourth participant) solved a task together while we collected data, including tracking of positioning/orientation and ratings of the behaviors used. Hypotheses on patterns of potentially suitable telepresence robot behaviors were formed based on the collected data. These include the hypothesis that a robot can pass through a group when retreating, without this affecting how comfortable that retreat is, and the hypothesis that a group will rate the position/orientation of a robot as more comfortable when it is oriented more toward the center of that group.

Together with the conclusions from our observations, these hypotheses give insights in potentially socially normative robot behaviors for the Teresa robot.

2 Contributors

This document is the collaborative effort of Jered Vroon, Gwenn Englebienne and Vanessa Evers. Jered Vroon performed most of the work described in this work and provided the first drafts. Vanessa Evers and Gwenn Englebienne supervised these efforts and gave feedback on these drafts.

In addition, we wish to mention here that the inductive study here described was conducted as part of a larger, collaborative effort. This effort was the joint work of Jered Vroon, Michiel Joosse, Manja Lohse, Jan Kolkmeier, Jaebok Kim, Khiet Truong, Gwenn Englebienne, Dirk Heylen, and Vanessa Evers[1].

The controlled observations on elderly controlling the robot was conducted as part of experiment 1B by the Teresa consortium. Though all consortium partners contributed to this, we want to mention some people specifically. Together with Jered Vroon, Kyriacos Shiarli (UvA) was in large part responsible for the first version of the protocol. Raphaël Koster and Aly Chkeir (Madopa) took the role of experimenters and did most of the data collection (with the aid of the project partners).

[1] We declare that all of the hours invested in this joint effort have only been written once. More specifically, only Jered Vroon, Jaebok Kim and Gwenn Englebienne have written hours invested in this study on the Teresa project. Since the results of the study have been used for tasks in work packages 2 and 3, these hours have been divided accordingly.

3 Normative Behavior Report

The TERESA project aims to develop the modules required for a telepresence robot to display socially appropriate behavior in various interactions with groups of (and individual) elderly. These interactions entail both navigation in areas with people and conversation behavior, including the approach and retreat behaviors required to socially initiate or terminate a conversation.

To identify the behaviors that could be suitable in these interactions, we have conducted a literature study, observations with the target group (with and without the robot), and an inductive user study. Based on the literature study and the degrees of freedom of the robot – which can position itself and change its head height/tilt – we chose to focus ourselves on proxemic behaviors; varying the distance between agents by positioning, orientation and posture (Section 3.1). With our observations we attempted to specify the contexts in which these interactions occur, as well as to further investigate the proxemic behaviors specific to elderly (Section 3.2). With the inductive user study we performed, we aimed to get insight on the specific features of position, orientation and posture that are applicable to dynamic interactions between a telepresence robot and a group (Section 3.3). We discuss the implications and limitations of our findings in the general conclusion (Section 4).

3.1 Literature

If a robot is to position itself in an environment with people, it will have to consider how to take these people into account – both when navigating and when conversing. In other words; which approach angle and approach distance to use when interacting with a person. The field of proxemics, first coined by anthropologist Hall [1], describes these interaction distances in some detail. In his book, Hall defines some general numbers for different comfort zones based on familiarity with the interaction partner. He further suggests that the prime underlying reason for proxemics are the sensory experiences of the interacting partners, which seems to be confirmed by the influence of various environmental factors such as lightning conditions [2] on preferred distance.
These numbers and comfort zones have been used as a starting point for many robotic applications [3,4,5,6]. One approach has been to search for relevant factors that could influence these distances to some extent in human robot interaction. This way, (interacting) factors such as age, gender, history of pet ownership and height of the robot have been distinguished [3,4].

Another approach has been to use these preferred distances as parameter settings to adapt the behavior of the robot while in the proximity of humans. These findings suggest that, at least in social navigation, there is an optimal distance between too close and awkwardly far away [5].

Kristoffersson et al [6] combined proxemics, and the related theory of F-formations of Kendon [7], with elderly and telepresence in a user study. There they found a relation between different spatial formations and perceived presence and ease of use. Both perceived presence and ease of use are found as important non-technical requirements by various authors with various approaches [8,9,10,11]. Together with more technical functioning requirements, such as a stable wireless internet connection, these mediate the capacity to facilitate actual social contact. This is quite a relevant feature, as literature shows for example that elderly engaged in social activities have lower mortality risk [12].

To summarize, literature seems to suggest that it is feasible to use telepresence robots as a social device for the elderly. In doing so, several requirements need to be fulfilled regarding among others presence, functioning, ease of control and capacity to facilitate actual contact. The semi-automatic display of social behavior in general and proxemics behaviors in specific likely play a role in fulfilling those requirements. That said, even though proxemics seem to be quite applicable to social robots, they are not yet an out-of-the-box-solution; individual and environmental factors both have their effect on what kind of proxemics behavior is appropriate.

3.2 Observations

To get a better impression of our target group and the existing contexts in which the robot will be to navigate itself, we conducted some observations. Specifically, we investigated different kinds of social situations, the applicability of proxemics to elderly (in interaction and in navigation), and other social signals used by elderly. We focused our observations on social activities with various elderly in the same room, because those allowed us to study both one-on-one and group interactions. We also included observations on elderly using the Teresa robot to interact with their peers.

3.2.1 Observation goals

We wanted to make two kinds of observations. First, because the Teresa robot will be used to participate in social (group) activities, we aimed to get an impression of those activities and the behaviors that could be suitable for them – without the robot. Second, because interacting through the robot may pose its own challenges to the target group, we aimed to also investigate such interactions.

For this first kind of observations, we defined the observation goals listed below. The first and second question can be quite straightforwardly answered and aim to provide some intuition about the general context. The third and fourth question are the biggest, and aim to identify relevant (proxemics) behaviors used by elderly. The last question is intended as a safety net, for catching all other factors that might not fit with the other questions but could still be relevant. All these observation goals are aimed at the context of interactions between elderly, in pairs and in (small) groups.

1. What are the different social activities and how do they proceed?
2. How much space will be available for navigating the robot during social situations?
3. How do elderly handle each other's (and their own) comfort zones?
4. Which signals do elderly use to steer each other's social behavior, and what is the apparent goal of these signals?
5. Which other signals and factors could be relevant for social behavior with elderly?

The second kind of observations is somewhat apart from the first kind, since it explicitly involved the use of the Teresa robot;

6. What are the challenges that using the Teresa telepresence robot could present to our target group.

Together, these observations will give us further insights in the contexts in which the Teresa robot will have to function.

3.2.2 Methods

Since we are interested in two kinds of observations, we used two methods to study them. To investigate the first five observation goals, we observed a variety of social activities involving groups of elderly 'as they are' (Section 3.2.2.1). To investigate the last observation goal, we observed members of the target group as they controlled and interacted with the Teresa robot during the joint consortium data collection (Section 3.2.2.2).

3.2.2.1 Observing social activities

To observe the social situations ‘as they are’, we performed passive observations. To this end, we created some forms in which we could easily indicate the phases in different social events and the navigation space available during those phases, and the different kinds of social signals including their frequency.

During our observations we openly sat down on a small distance from the social event and started taking notes on the forms. For organized events, we always discussed our observations beforehand with the organizers. Though most of the time we did not identify ourselves beforehand to the elderly, we always explained our intent and purpose to anyone who had questions.

Observations were conducted at three locations; retirement home ‘De Polbeek’, ‘het Alzheimercafé Oldenzaal’, and ‘de Ariënsstaete’. A detailed overview of the different activities we observed and their attendance can be found in Table 1.

De Polbeek

‘De Polbeek’ is a retirement home located in Zutphen, the Netherlands, that provides various levels of care. Some of the elderly attending the activities are elderly living independently in the neighbourhood, others live in assisted living. They provide care to clients with various cognitive and physical impairments as long as they do not require constant care. Central in the retirement home is the big common area, which is open all day. It provides ample place to sit around tables and is used for most social events, both organized and spontaneous. A central counter in front of the kitchen serves as a combined reception, café, and shop. Around noon a meal is served, primarily to elderly that eat there every day. Small snacks are served all day long. The common area is mainly used by elderly regulars. Two smaller ‘living rooms’ in a more private location provide day care to two groups of elderly with cognitive impairments (mostly dementia). On the 15th of May 2014 one observer spend the day in both the common area and one of the ‘living rooms’ for our observations.

Het Alzheimercafé Oldenzaal

‘Het Alzheimercafé Oldenzaal’ is a monthly meeting in Oldenzaal, the Netherlands, on the various practical and emotional concerns that come with Alzheimer and dementia in general. It is open to anyone who wants to attend, but mainly attended by elderly with some form of dementia and their caregivers. Commonly meetings are attended by 40 to 50 people. The observed meeting was a celebration of the 10 year anniversary, attended by about 100 people. The main speaker was the Dutch singer Marga Bult, who performed some nostalgic songs. On the 26th of June 2014 one observer made observations during this meeting.

De Ariënsstaete

'De Ariënsstaete' is a retirement home located in Enschede, the Netherlands, much like the Polbeek – but slightly larger. Though it also has a central common area, we did our observations in an open meeting space at one of the far ends of the building. On the 17th of July 2014 one observer made observations during a meeting of the 'geheugentraining' (memory training), where six elderly (all female) performed various activities to benefit their memories.

3.2.2.2 Observing interactions with the Giraff robot

In addition to the observations described above, we also did some observations on situations where members of the target group were controlling the Teresa telepresence robot. For practical and safety reasons, these observations were more controlled and happened under the supervision of one or more experimenters. These observations were aimed at getting insight into the challenges that using the robot could present to our target group.

To get a first intuition, we performed two informal observations. The first took place in the Living lab in Troyes during the first integration week. On the 8th of July we had several people control the Teresa robot, among which one member of the target group. In the second three members of the target group likewise controlled the Teresa robot using the Giraff interface.

During the joint consortium data collection, we investigated further similar situations where elderly controlled the robot. From September 8th to September 15th and from October 14th to October 17th, at total of 9 members of the target group participated. All these participants used the Teresa robot to interact with one or two of their peers, making for a total of 19 participants. In these interactions, we investigated both situations in which they controlled the robot themselves and situations in which a Wizard of Oz controlled the robot to display autonomous behavior, both good and bad. A more detailed description of the joint consortium data collection can be found in Deliverable 5.1 (though the focus there is different, a full description is included in the annex of that deliverable).

3.2.3 Results

We will here report our results per observation goal. Implications that these results could have on the design of socially normative behavior for a telepresence robot will be discussed in our conclusions and discussion (Section 3.2.4).

3.2.3.1 What are the different social activities and how do they proceed?

In Table 1 we have listed the various social events during which we have done our observations. The last column indicates the extent to which the participants had to follow a

	# attendants	duration	
Shared meal in the restaurant	~40 elderly	12-13h	Schedule
Living room	5-3 elderly	Ongoing	Free
Table of Memories (social art project)	5 elderly	14-16h	Guided by organizer
Sitting together in the café	6 elderly	Ongoing	Free
Alzheimer café	~100 elderly and caregivers	19-22h	Guided by organizer and main speaker
Memory training	6 elderly	1h 30m	Guided by organizer

Table 1 - Overview of the observed social events

fixed schedule. The activities that followed a fixed schedule or were guided by the organizer in general all had roughly the same phases; arrival, welcome/introduction, go through the steps of the activity, departure. It is noteworthy that there usually was a rather large time window for both arrival and departure (15-30 minutes).

3.2.3.2 How much space will be available for navigating the robot during social situations?

We observed that people actively tried to keep open a lot of available navigation space. For example, during the shared meal several caregivers and elderly without walking problems actively cleared away all walking aids from the main paths (and returned them to their owners at the end of the meal). Though the space was a bit more cramped at the Alzheimer café, we there observed similar behavior.

3.2.3.3 How do elderly handle each other’s (and their own) comfort zones?

Part of the observed population seems to maintain a reasonable distance to their communication partners. However, we also commonly observed that people were very close to each other during communication.
Among people with hearing problems (identifiable by their hearing aids and utterances), ‘leaning’ behavior was very commonly observed. During conversation, the person with hearing problems would turn her upper body and face to the person she was talking with and lean towards her conversation partner. The conversation partner commonly returned the

leaning behavior. We observed several times that a conversation partner relayed what others in a group had said. In most cases, the ‘leaning’ behavior was only used during conversation and the interacting parties kept more distance when not talking. In one case we saw someone displaying this ‘leaning’ behavior towards an interaction partner some meters away at the far end of a table – which suggests that it is also used as a social signal.
We observed one exception; an older lady who had severe hearing problems did not display any leaning behavior. One of the caregivers compensated for this by moving real close and showing strong leaning behaviors. In this particular case it took the observer well over 30 minutes to realize that the older lady had severe hearing problems, which seems to indicate that the leaning behavior also provides an important social cue.
In interactions between standing and sitting people, the standing people often used the chairs of the sitting people for support. When not only navigating, but also conversing, they commonly lean over the sitting people to establish some eye contact. As a result, people were very close in these interactions. Possibly this behavior is also intended to compensate for hearing problems, as it was commonly observed in interactions where one of the interaction partners seemed to have hearing problems.

Some of the caregivers also got very close to their interaction partners. For example, we often saw caregivers crouch to be on the same height as the sitting people they were interacting with. These behaviors may have partly been for practical reasons, such as hearing problems. However, since they were also displayed in communication with people without hearing problems, such as the observer, it is likely that these behaviors are also used as a social signal. It may be interesting to note that these behaviors were observed more in caregivers that have to guide social activities than in caregivers that do more physical work, such as medication, helping people stand, cleaning, etcetera.

The few people we observed who were severely restricted in their freedom of movement and used a wheelchair all did not seem to interact much with the people in their environment. Their wheelchair may have been a cause, but they also seemed to be at a somewhat larger distance than commonly observed.

3.2.3.4 Which signals do elderly use to steer each other’s social behavior, and what is the apparent goal of these signals?

Some of the observed elderly use rather large gestures. For example, during the meal, when one group wanted to ask for an extra serving, the whole group started to gesture wildly (including partly standing up) to get the attention of the waiter. Another example is that of an older lady, who became a little angry and in expressing that started leaning and pointing towards the person she was angry at.

Navigation-wise we saw surprisingly little social signals; most of the (potential) ‘conflicts’ were very effectively resolved by either waiting, forming a line or choosing another direction altogether. This should not be taken to suggest that navigation was not social, but rather that, as far as we observed, there were little explicit gestures involved.

3.2.3.5 Which other signals and factors could be relevant for social behavior with elderly?

Some of our observations did not really fit the other observation goals, but could still be relevant for social behavior with elderly. We have here listed a few of the most salient;

- The great majority of interactions took place between sitting elderly. Interactions between standing elderly were uncommon, interactions between moving elderly even more so.
- Some elderly did not seem to pro-actively participate in conversations and other interactions; they did react when spoken to, but did not really take initiative. This behavior seems to be related to dementia, since we observed it more often in the ‘living room’ and during the Alzheimer café.
- Many elderly seemed to have a select group of contacts, even in a bigger crowd. Their interaction mostly seemed to be limited to people in this group, though caregivers sometimes interjected.
- Most of the observed elderly seemed to enjoy their social interactions.

3.2.3.6 What are the challenges that using the Teresa telepresence robot could present to our target group.

The observed elderly required 20-60 minutes for training. These longer times for training seems to be partly caused by a lack of experience with computers in general, since those that required shorter training usually report being more familiar with computers. After training, they often drove slow and carefully (with some exceptions). One of our participants had polyarthritis, causing additional difficulties in using her hands to control the robot.

The conversations were strongly influenced by being mediated by a telepresence robot. Many of the elderly were concentrating strongly on controlling the robot and as a result seemed to be less available for conversation. Or, as one participant remarked; “I can’t do everything at the same time”. One of the participants even made a similar remark while not he, but the Wizard of Oz was controlling the robot. Though this may have been a novelty effect, much of the conversations were about the robot. In addition, some of those interacting with the robot tended to give it orders (such as “follow me”, “sit down”). They do however seem to feel presence; some even remarked that they saw no difference between talking with the robot or without it, describing it as much more ‘present’ than talking through a phone or Skype.

Afterwards, many of the participants clearly indicated that they enjoyed the experience. Despite the Wizard of Oz also displaying ‘bad’ behaviors, some of them still indicated that they liked the autonomous behavior of the robot. As one participant remarked; “It is my husband that I don’t trust [to control the robot], not the robot!”

3.2.4 Conclusions and Discussion

We have observed a variety of different social activities in which the Teresa robot could play a role. We will here list some of the implications these observations could have for the design of the Teresa robot and its (socially normative) behaviors.

- The Teresa robot will probably not have to navigate in cluttered locations, since most areas in which elderly interact are actively kept free of obstacles.
- As one would expect, elderly respect each other’s personal zones. However, elderly with hearing problems commonly use ‘leaning behavior’ where those involved in an interaction actively lean in to intimate distances, presumably to hear each other better. This finding is consistent with those of Webb & Weber [13], and rather relevant given the prevalence Presbycusis (age related hearing loss) – which affects over half of the population aged 75 and above [14]. The Teresa robot will probably have to take this into account as a requirement, even though the capacity to change the volume settings provides alternative ways of handling these situations.
- Caregivers often get very close to their interaction partners, also those without hearing problems. This suggests that it is a social signal (probably indicating figurative closeness), which could also be used by the Teresa robot.
- Though some elderly use rather large gestures in communication, we saw very little social signals in navigating. Potential conflicts in navigating are often effectively resolved by either waiting, forming a line or choosing another direction altogether. The Teresa robot should probably incorporate these or similar strategies.
- Controlling the robot is hard to learn and even after training requires a lot of effort, reducing the quality of the conversation. Introducing semi-autonomous navigation, as the Teresa project aims to do, thus could well help make the robot more usable

- Elderly seem to enjoy their social interactions, both with and without the robot.

In addition, it should be noted that our target group is complex; various internal, highly personal, variables strongly influence their behavior in interactions. One of the limitations of the generalized conclusions above is that they focus on the majority, without taking these individual differences into account. Examples include the one person we observed with hearing problems that did *not* show leaning behavior and the person whose polyarthritis hindered her ability to control the robot.

Though it does not really fit the scope of the Teresa project, it would thus be interesting and relevant as well to try and take these individual differences into account. At the very least, the Teresa robot should allow the person controlling the robot to modify the semi-autonomous behaviors of the robot to do so.

Overall, when allowing for these individual differences, the conclusions here discussed can help guide the general design of the Teresa robot and its (socially normative) behaviors.

3.3 Inductive user study

To get insight on the specific features of position, orientation and posture that are applicable to dynamic interactions between a telepresence robot and a group, we performed an inductive user study.

This user study has also been submitted to the RoMan 2015 conference[2]. The description of the study below is for a large part copied from that submission.

3.3.1 Introduction

Since the Teresa robot has the capacity to move around, one important aspect of being social are the behaviors required to be positioned and oriented in a social way (**social positioning**). In this section we will focus on social positioning with the Teresa robot for conversation with a group of people, including approaching a group with that aim and retreating from it.

A wide range of robotics research concerns various aspects of social positioning (e.g. [3,15,16]). Most of this work is based on two theories from sociology on social positioning in humans, proxemics [1] (distancing) and F-formations [7] (spatial arrangements).

[2] Submission title; “Dynamics of Social Positioning Patterns in Group-Robot Interactions: An Inductive Method”, authors are Jered Vroon, Michiel Joosse, Manja Lohse, Jan Kolkmeier, Jaebok Kim, Khiet Truong, Gwenn Englebienne, Dirk Heylen, and Vanessa Evers

However, one factor that makes it hard to study social positioning in interactions is that several dynamics are involved. First, interactions take place over time, which implies that requirements for social positioning can change over time and that movements should be taken into account as well. Second, participants in an interaction respond and adapt to each other. This means on the one hand that people could adapt to the robot and on the other that they might expect the robot to adapt to them. We will refer to these different dynamics as the **interaction dynamics**.

In the study we introduce here, there were repeated trials where one participant (**Visitor**) controlled a telepresence robot to have various conversations with a group of three other participants (**Interaction Targets**) (Figure 1). Participants could use any social positioning behavior they found suitable, to allow for the interaction dynamics to arise. We collected a range of data on the behaviors used.

Our methodology is aimed at generating hypotheses on (dynamic) features that could be taken into account when designing social positioning (telepresence) robot behavior for conversation with a group. Therefore, we apply inductive reasoning to go from patterns that can be qualitatively and quantitatively observed in the collected data to hypotheses for more general situations.

We will focus on introducting the inductive methodology and on describing the found

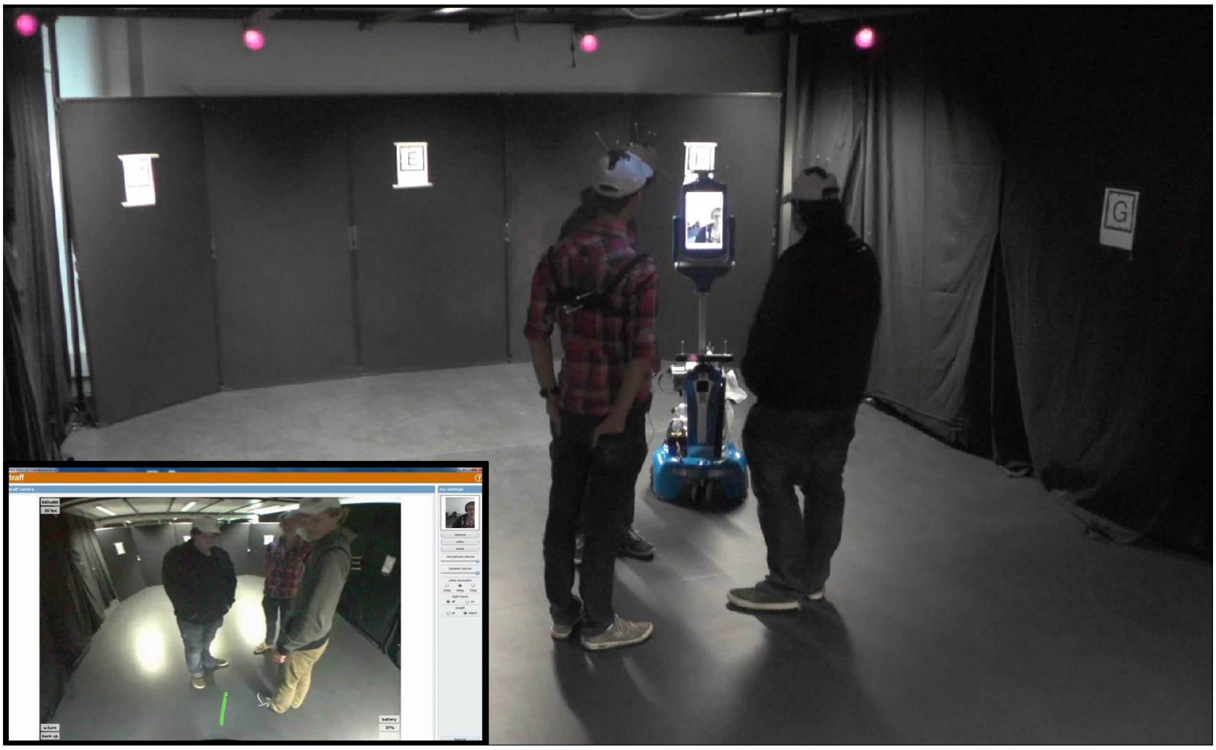

Figure 1 - Example of the interactions used in the study. A group of four participants discuss a murder mystery. One of them is remotely present through a robot, and has to go through several approach/converse/retreat cycles. The inset shows the interface as seen by this participant.

hypotheses on social positioning in a dynamic interaction between a (telepresence) robot and a group. We will discuss the theoretical background underlying our method (Section 3.3.2) and further specify it (Section 3.3.3). Furthermore, we will present our findings (Section 3.3.4), and discuss the implications and limitations of those findings (Section 3.3.5).

3.3.2 Theoretical background

Much of the work in social positioning for robots is based on two theories from sociology on social positioning in humans: proxemics and F-formations. Proxemics, first introduced by Hall [1] focuses on the distances people keep from each other. F-formations, as introduced by Kendon [7], describe the different spatial arrangements people can use in social interactions. As proxemics and F-formations would predict, many different social situations can be distinguished based on only position and orientation information (e.g. [17,18]).

We will give an overview of the existing work in social positioning for robots, including work which demonstrates how the behaviors displayed by participants controlling a telepresence robot can be used to investigate the suitability of different (telepresence) robot behaviors (Section 3.3.2.1). We will further discuss work that focuses on the different interaction dynamics (Section 3.3.2.2).

3.3.2.1 Measures of social positioning robot behavior

Previous work has applied and investigated proxemics and F-formations in the context of robotics. Significant effects have been found in various settings; in different contexts [19], with different properties of the robot [20], with relation to the background of the participants [3], and for different cultures [21]. These findings show that taking proxemics and F-formations into account can have a positive effect on the perceived appropriateness of the displayed robot behavior.

Similar questions have also been approached by having participants control the robot. By definition, this approach involves some form of telepresence, and often uses robots that are equipped with a video connection as well. The approach can be used to have participants experience the possibilities and limitations of the robot [8] or to inform design decisions [22]. Of particular relevance in the context of this study is the research of Kristoffersson et al. [23] and that of van Oosterhout & Visser [4], since they actively observed the displayed behaviors. Based on manual annotations of visual data (video/photo), they were able to observe relevant patterns in those behaviors. Van Oosterhout & Visser [4] for instance found that people generally position themselves within Hall's personal space zone. Kristoffersson et al. [23] found that when talking through a telepresence robot about a disembodied topic (here

a remote control) participants tend to assume a L-shape arrangement, as Kendon's F-formations would predict [7].
Actively observing the behaviors used by participants controlling a robot thus seems a fruitful approach to investigate suitable social positioning (telepresence) robot behavior.

3.3.2.2 Dynamics of social interaction

One factor that makes it hard to study human robot interaction is that it is a dynamic process. Or, as Hüttenrauch et al. [16] put it when investigating the applicability of proxemics and F-formations to the field of robotics, "The dynamic changes and transitions from one interaction episode state into the another one are difficult to express in terms of Hall's interpersonal distances and Kendon's F-formations arrangements when tried in a HRI scenario" ([16], p.5058).
There are two sides to these interaction dynamics. First, there are the temporal dynamics; since interactions all take place over time, requirements can change and movements rather than static positions are relevant. There is a limited set of papers that explicitly look into the temporal dynamics of people interacting with a robot [16,23,24]. Secondly, participants in an interaction all influence each other's behavior. Obviously, these dynamics become more complex as more entities become involved in the interaction (e.g. a robot interacting with a group). Complex as they are, these dynamics allow for many interesting applications. For example, by relying on people to get out of the way for a navigating robot [25], to signal approachability with a group of virtual agents [26], or to influence the formation of people interacting with a robot [15]. As evidenced by these papers, these interaction dynamics are relevant and can have a strong influence on what happens in the interaction.
However, as far as we are aware, the work presented in this study is the first to look into these interaction dynamics concurrently.

3.3.3 Method

The aim of this study is to generate hypotheses on (dynamic) features that could be taken into account when designing social positioning robot behavior for conversation with a group. However, it is still too much of a challenge to automatically generate robot behaviors that are sufficiently dynamic and appropriate for such a study. A solution discussed in the theoretical background is having participants control a (telepresence) robot and observing the behaviors they use. Therefore we created a setting in which small groups of four people were having a conversation. One of the participants was present trough a telepresence robot (the Visitor), and used the robot to interact with the rest of the group (the Interaction Targets).

In this section, we will discuss the task (3.3.3.1), the procedure (3.3.3.2), the data collection (3.3.3.3), the participants (3.3.3.4), and the data synchronization/segmentation (3.3.3.5) in more detail.

3.3.3.1 Task

The task had to motivate the participants to have a conversation in which the Visitor had to go through several cycles of approach/converse/retreat behavior. We thus asked our participants to solve a murder mystery, where the Visitor had to go and collect eight clues, and return to the group in order to share the clues. To eliminate effects of the specifics of the murder mystery, groups were randomly assigned to one of three murder mysteries. Preliminary analysis did not indicate any effect of the different murder mysteries, so this variable has been excluded from the analysis.

Each of the clues had to be picked up at different markers positioned around the interaction area (see Figure 2). The location of the marker for the next clue was provided to the Visitor 75 seconds after the previous clue was presented, which gave ample time for both approach and conversation (we confirmed this in a pilot study).

Each group of participants was thus part of a total of eight approach/converse/retreat cycles, separated by the Visitor having to go to a marker to collect the next clue. After these, rather than a ninth clue, the Visitor was given the instruction at the marker to decide as a group on a primary suspect. This resulted in one last approach, and a discussion that was ended by the experimenter when consensus was reached.

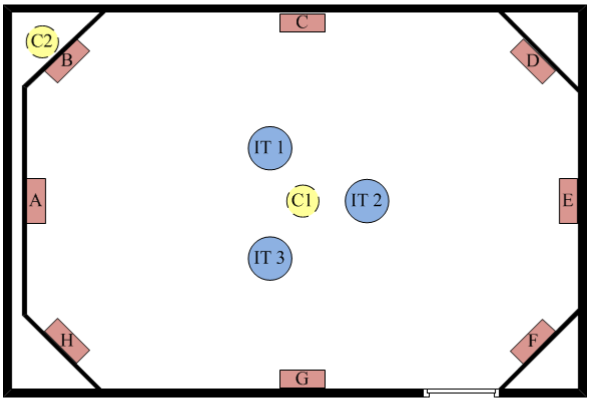

Figure 2 - Overview of the interaction area. On the circle in the middle the positions of the Interaction Targets are indicated (IT1, IT2, IT3), these were projected using a projector mounted to the ceiling, but only in between the approach/converse/retreat cycles. The rectangles near the border of the interaction area indicate the positions of the markers A-H. C1 and C2 indicate the positions of the cameras.

3.3.3.2 Procedure

The study took place in a controlled laboratory setting. For the study, we used a Teresa robot equipped with the hardware required for the data collection (Section 3.3.3.3). The robot was located in a room with the Interaction Targets (**interaction area**). The Visitor controlled the robot from a separate room using the standard Giraff software (Figure 1).

After the briefing, participants were randomly assigned to either be the Visitor (1 participant) or be an Interaction Target (3 participants). This was followed by task-specific instructions from the experimenter. The Interaction Targets were equipped with everything required for the data collection (Section 3.3.3.3) while the Visitor was given a brief training on controlling the Giraff (changing position, orientation and head tilt).

The Visitor approached the Interaction Targets for a total of 9 times. The first eight times the Visitor approached the Interaction Target from one of the eight markers shown in Figure 2. The final approach was from the same marker as the first approach. To eliminate possible ordering effects, the Visitor had to go to the different markers in one of eight randomly assigned pre-defined counterbalanced orders[3].

At the end of each cycle, before being given the next clue, we asked participants individually to fill in a brief questionnaire on the robot behavior during that cycle (Section 3.3.3.3). The next clue was presented after all participants had finished filling in the questionnaire.

While filling in the questionnaire at the end of each cycle, the Interaction Targets were asked to stand in a fixed formation which was temporarily projected on the floor. These projections had the effect of resetting the position of the participants before each cycle, which allowed for a fair comparison between cycles. The projections were not shown during the cycles and we explicitly told our participants that they were allowed to move around during the cycles. We used two formations to cover some of the variations that might naturally occur: a circular formation, with every participant occupying an equal amount of space, and a semi-circular formation featuring an open space [26]. Groups were randomly assigned to one of the formations. For the findings presented in this section, we did not introduce formation as a factor.

At the end of the interaction part of the study, after the group had reached a consensus on their primary suspect, we asked all participants individually to fill in a post-experiment questionnaire (Section 3.3.3.3).

3.3.3.3 Data collection

During the study, a variety of data has been collected. Here we will describe the methods we used for collecting objective data with various sensors and subjective data with questionnaires.

Objective measures: All three Interaction Targets were equipped with uniquely identifiable markers (one on the back of the chest, one on a cap), which were tracked by an OptiTrack

[3] We used a balanced latin square design for this, controlling for regularities in the order in which positions close-by and further to the previous position would be chosen.

(www.naturalpoint.com/optitrack/) motion capture system using 8 infrared cameras. The robot was similarly equipped. The system used allows sub-centimeter level precision tracking of both position and orientation of each marker. We optimized tracking for the center of the interaction area, to make sure we could properly capture the interaction. Markers near the edges of the interaction area could often not be tracked. To ensure proper tracking of the actual interaction, we informed the Interaction Targets about this and asked them not to get too close to the edges of the interaction area. In the analysis here presented, we will represent the position of the Interaction Targets by the position of the marker on their cap.
Speech of the Interaction Targets was recorded by equipping them with microphones for close talk recordings. The robot was equipped with a microphone array to record audio and a Kinect depth camera.
Two cameras were positioned in the interaction area. One camera provided a side view, the other a (fish eye) top down view. All interactions of the Visitor with the interface were recorded with screen capture software.

Subjective measures: After each approach/converse/retreat cycle (i.e. 9 times), all participants were given an in-between questionnaire. After the interaction part of the study a post-experiment questionnaire was administered.

a) In-between questionnaire: The in-between questionnaire consisted of five questions; two related to the usefulness of the clue and task progress. The remaining questions measured comfortability with the robot operators' driving behavior during approach and retreat, and the distance to the robot during conversation. For the robot operator, we instead used three questions assessing work load (based on [27]).
b) Post-experiment questionnaire: The post-experiment questionnaire consisted of 49 items. Among others we measured co-presence and attentional engagement [28]. Furthermore we measured the participants' attitudes towards robots (based on [29]) and workload [27].

3.3.3.4 Participants

A total of 56 participants participated, divided into 14 groups of 4 persons. Of these, 13 (23.2%) were female, 43 (76.8%) male. All were students, aged between 18 and 32 years with a mean of 20 (SD=2.2). Most participants had the Dutch nationality (85.7%).

3.3.3.5 Data synchronization and segmentation

After the experiment, we synchronized the data from the various sources in Elan[4] using points that were visually/auditory/motion-wise salient. We used the tracking data to determine when the robot was moving or not[5] and then used that information to segment the collected data.

Approaches are defined as the set of movements (and enclosed non-movements) between the Visitor being given a clue and the Visitor starting to (verbally) share that clue with the Interaction Targets. Likewise, **Retreats** are the set of movements (and enclosed non-movements) between the buzzer indicating that the next clue could be collected and the end of the recorded movement to the marker. The segment in between Approach and Retreat is a **Converse**.

In the segment between each Retreat and the next Approach the participants were filling in the questionnaires, we did not use this segment in our analysis. After the ninth Approach, the task of the participants changed, so we excluded that data from our analysis as well.

3.3.4 Findings

We will present findings from the (quantified) observations (3.3.4.1) and the investigation of the relations between features of the dynamics of the motion patterns and the ratings of the Interaction Targets (3.3.4.2).

3.3.4.1 Observed patterns of behavior

Under the assumption that the participants all try to behave as such, suitable interaction behavior would likely be more common. Thus, using inductive reasoning, patterns that are commonly observed in the interactions can be used as hypotheses for suitable behavior. We will introduce such patterns, organized by the phase of the interaction (Approach/Converse/Retreat) in which they occurred. Where applicable, we will quantify these patterns and use the tracking data to calculate how common they are.

Approach: During the Approach, most Visitors drove the robot towards the Interaction Targets (Table 2-1,4). Only in one of the groups we observed that the Visitor only turned the robot to face the Interaction Targets without driving to them.

---

[4] Annotation tool developed by the Max Planck Institute for Psycholinguistics (The Language Archive, Nijmegen, The Netherlands), available from tla.mpi.nl/tools/tla-tools/elan/

[5] We defined the robot to be moving if the position of the marker placed on its base, smoothed over 50 frames, changed more than 0.02cm between frames (2.4cm/s). This yielded some false positives.

	Quantified pattern	MIN	Q25	Q50	Q75	MAX
1	Distance between robot and center of the group at end of Approach	7cm	91cm	113cm	134cm	315cm
2	Angle (in degrees) between robot viewing direction and center of the group at the end of the Approach	0deg	5deg	10deg	18deg	133deg
3	Angle (in degrees) between the actual position of the robot at the end of the Approach and the position it would have had if it had moved in a straight line from the marker to the center of the group.	0deg	9deg	18deg	34deg	135deg
4	Distance between first and last detected position of robot during Approach	0cm	111cm	176cm	211cm	293cm
5	Distance between first and last detected position of Interaction Targets during Approach (averaged)	1cm	9cm	13cm	21cm	84cm
6	Distance between first and last detected position of robot during Converse	0cm	0cm	0cm	1cm	233cm
7	Distance between first and last detected position of Interaction Targets during Converse (averaged)	5cm	13cm	20cm	37cm	122cm

Table 2 - Quantified patterns of behavior with a five-number summary (minimum (MIN), lower quartile (Q25), median (Q50), upper quartile (Q75), and maximum (MAX)) of their distribution in the collected data.

When approaching, Visitors commonly aimed for the closest-by opening between the Interaction Targets they could see, rather than taking a larger detour to approach the group from another angle (Table 2-3). We only observed one Visitor taking multiple such detours; for this Visitor, the Interaction Targets were in the semi-circular formation and the detours seemed aimed at the large opening in that formation.
In some cases we saw that the Interaction Targets actively changed their position to accommodate for the approaching Visitor – e.g. by making the opening the Visitor was aiming at larger and/or by moving a little towards the Visitor. However, this pattern was only moderately common (Table 2-5).

Converse: During conversation, many Interaction Targets changed their position between the beginning and the end of the Converse segment, while movement of the Visitor was very rare (Table 2-6,7). When the Visitor did move, these movements were rotations that increased the visibility of the Interaction Targets.

Retreat: In 38 out of the 112 Retreats (33.9%) we observed, to our surprise, that Visitors passed straight through the group. This was always done to reach a marker located directly behind the group. In 42% of these situations the Visitors communicated this beforehand. Only in rare cases (9 cases, 8% of total Retreats) we saw that the Visitor backed up from the group and took a detour instead. The Interaction Targets actively assisted the Visitor, by pointing out the position of markers, by moving out of the way and even by actively inviting the Visitor to pass through the group.

3.3.4.2 Relating motion patterns with ratings

The ratings given by the Interaction Targets during the in-between questionnaire give additional information on whether the displayed behavior was actually perceived as more or less comfortable. Patterns in the relation between this information and (dynamical) aspects of the recorded behavior can be used as further hypotheses for suitable behaviors.
There are large individual differences between how the different Interaction Targets answered the in-between questionnaires, which makes it harder to reliably extract this information. To compensate for this, we used Gaussian normalization (normalizing the scores of an Interaction Target by subtracting the mean of those scores and dividing by their standard deviation), averaged over the three Interaction Targets in a group.
We will first describe some informal findings acquired by looking for patterns in the Approaches/Converses/Retreats that had the ten highest and ten lowest average normalized Ratings. Then we will discuss more quantified ways for looking at these findings.

Motion patterns with the highest/lowest ratings: Driving the robot with a smooth and steady path seems to be important for the average normalized ratings, since we observed this in most of the ten Approaches and Retreats that scored highest, while observing more ‘wobbly’ robot motion in many that scored lowest.
In most of the best Approaches we additionally saw that the Visitor ended up at on average 1.25 meter from and aimed at the center of the group, and changed the head tilt of the robot to face the group even better (see Figure 3a). In some of the worst Approaches the Visitor did not approach at all, or got so close to the Interaction Targets that they stepped away (see Figure 3b).
In nine out of the ten best Retreats we saw that the Visitors explicitly communicated their goals (verbally) before driving. The pattern we observed before, in which the Visitor passed straight through the group while retreating, was observed in both the best and the worst ten Retreats and thus seems to have had no strong influence of itself on the given ratings.
We did not observe any particularly salient patterns in the ten best Converses, but in the ten worst the robot was usually far away from the group center or relatively close to at least one of the Interaction Targets.

Quantified relations with ratings: We wanted to quantify the relation between the ratings and several aspects of the used motions. To do so, we used Spearman’s rank correlation

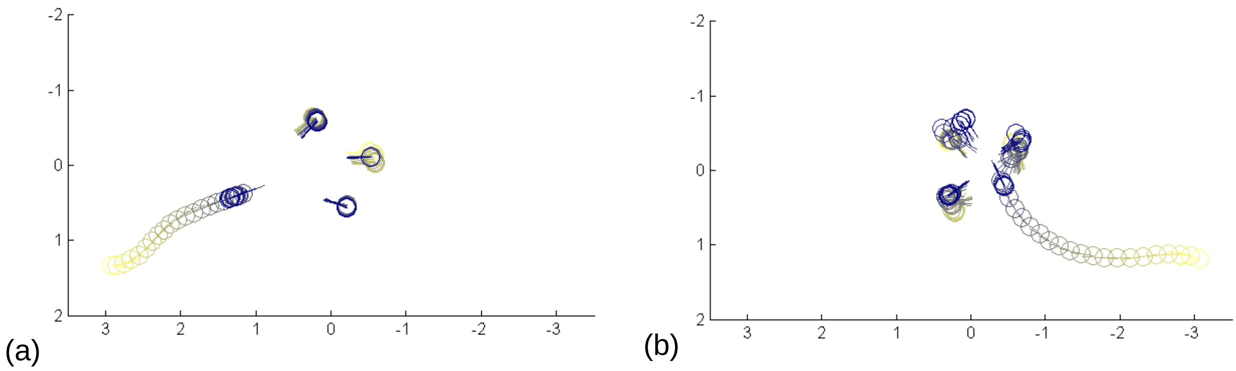

Figure 3 - Representation of head tracking data from two Approaches, one with a high average normalized rating (a) and one with a low average normalized rating (b). The circles with lines show the positions and orientations of the Visitor and Interaction targets in the interaction area. Indicators near the end of the Approach are darker. Axes indicate distance (in meter) from the center of the interaction area in the horizontal and vertical direction.

since it is robust against outliers and non-normally distributed data (the average normalized ratings were not normally distributed, p=0.0306 in a Kolmogorov-Smirnov test). We did not find a significant correlation for distance between the robot and the center of the group at the end of the Approach (ρ = 0.109, p = 0.220), nor for the speed used during the Approach (ρ = -0.008, p = 0.929). We did, however, find a significant correlation for angle between the direction of the robot and the center of the group at the end of the Approach (ρ = -0.218, p = 0.014). This indicates a positive relation between how well the robot faces the center of the group and the ratings.

Further analysis, for example based on mutual information, will likely reveal other measurable relations. Different aspects of the motion of the robot and its behavior in general are still open for investigation, and may well reveal subtle yet strong indicators of proper robot behavior.

3.3.5 Conclusions and Discussion

We have introduced a study in which a Visitor controlling a telepresence robot went through several approach/converse/retreat cycles with a group of three Interaction Targets. During these cycles, they together attempted to solve a murder mystery, with the Visitor leaving repeatedly to collect clues. We then identified various qualitative and quantitative patterns in the data we recorded in these interactions; common behaviors, regularities in the behaviors that were rated as most/least comfortable, and a correlation between these ratings and a particular positioning.

Using inductive reasoning, these patterns can be used as hypotheses for more general settings. These could be settings with a different task and different people. One of the limitations of inductive reasoning is that it is impossible to know beforehand if such a generalization is justified. For example, since our patterns were found in a setting with students, there is no guarantee they will translate to our target group. It is for this limitation of inductive reasoning that it is important to realize that our findings are hypotheses only. They could well generalize to various other settings, but this can only be proven through deductive reasoning.

We have used this inductive reasoning to generate a variety of hypotheses on social positioning. These include, in line with what proxemics would pose [1], the hypothesis that a (telepresence) robot should make an approach motion to get within approximately 1.25 meter of the individual interaction targets it wants to interact with. Based on our findings we can also hypothesize a relation between how well a robot faces the center of a group and how comfortable the group rates that positioning. In addition we found that dynamics indeed play a role in these interactions, since both the Visitor and the Interaction Target adapted their position and orientation to each other in various ways. This for example led to the hypothesis that a robot could pass through a group when retreating without this effecting how comfortable that retreat is.

Given the rich data that we collected, there are many opportunities for further analysis, in particular into the relation between aspects of the motion of the robot and how comfortable it is rated to be.

Overall, we have introduced an inductive methodology and used it to generate various hypotheses on social positioning. Our findings furthermore show that interaction dynamics can play a role in the interaction between a (telepresence) robot and a group.

4 Conclusions

In this deliverable we have discussed our investigation into socially normative behaviors that can be used by the Teresa robot. Based on the limited degrees of freedom of the robot and existing literature, we have chosen to focus on social positioning behaviors, for example proxemics.

With that in mind, we have conducted observations on the target group, both during social activities and during interactions with/through the Teresa robot. These have yielded insights into the contexts in which the robot will be applied. For example, the areas in which the Teresa robot will have to navigate will probably be relatively free of obstacles because those would also hinder the member of the target group. In addition, we also found that controlling the robot requires a lot of training and mental effort for members of the target group, which supports the aim of the Teresa project of making the robot more autonomous. We have also observed some social positioning behaviors common in the target group, the most salient of which was “leaning behavior” in elderly with hearing problems. The behavior of the Teresa robot will probably have to take this and/or the underlying needs into account.

After these observations, we have further performed an inductive user study, in which we collected a large data set of groups interacting with the Teresa robot. In this data set a variety of socially normative behaviors has been found. For example the hypotheses that when approaching a group, the Teresa robot should aim to get within 1.25 meter of the center of the individual group members and to be oriented towards the center of the group. We also found that it might be completely acceptable for the Teresa robot to pass straight through a group when retreating from it.

All these findings can guide the design of the behaviors for the Teresa robot to make them both more social and more suitable for the contexts in which the Teresa robot will be used. Due to the methodologies used (observations and an inductive user study) all our findings are merely suggestions for what could be socially normative robot behavior. Further work within the Teresa project, both in the joint consortium experiments/evaluations and for work package 3 can implement these suggestions and put them to the test. In addition, more potentially socially normative behavior patterns can be derived from the data set collected during the inductive user study.

We hope and expect that the socially normative behaviors we found can in these ways help the acceptance of the Teresa robot and similar robots.

5 References

[1] Hall, E. T. (1969). The hidden dimension. Anchor Books, New York.

[2] Adams, L., & Zuckerman, D. (1991). The effect of lighting conditions on personal space requirements. *The journal of general psychology*, *118*(4), 335-340.

[3] Takayama, L., & Pantofaru, C. (2009, October). Influences on proxemic behaviors in human-robot interaction. In *Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on* (pp. 5495-5502). IEEE.

[4] van Oosterhout, T., & Visser, A. (2008). A Visual Method for Robot Proxemics Measurements. 3rd ACM. In *IEEE International Conference on Human-Robot Interaction Workshop on Metrics for Human-Robot Interaction, Technical Report* (Vol. 471).

[5] Pacchierotti, E., Christensen, H. I., & Jensfelt, P. (2006, September). Evaluation of passing distance for social robots. In *Robot and Human Interactive Communication, 2006. ROMAN 2006. The 15th IEEE International Symposium on* (pp. 315-320). IEEE.

[6] Kristoffersson, A., Eklundh, K. S., & Loutfi, A. (2013). Measuring the quality of interaction in mobile robotic telepresence: a pilot's perspective. *International Journal of Social Robotics*, *5*(1), 89-101.

[7] Kendon, A. (1990). Conducting interaction: Patterns of behavior in focused encounters (Vol. 7). Cambridge University Press.

[8] Beer, J. M., & Takayama, L. (2011, March). Mobile remote presence systems for older adults: acceptance, benefits, and concerns. In *Proceedings of the 6th international conference on Human-robot interaction* (pp. 19-26). ACM.

[9] Boissy, P., Corriveau, H., Michaud, F., Labonté, D., & Royer, M. P. (2007). A qualitative study of in-home robotic telepresence for home care of community-living elderly subjects. *Journal of Telemedicine and Telecare*, *13*(2), 79-84.

[10] Tsai, T. C., Hsu, Y. L., Ma, A. I., King, T., & Wu, C. H. (2007). Developing a telepresence robot for interpersonal communication with the elderly in a home environment. *Telemedicine and e-Health*, *13*(4), 407-424.

[11] Kristoffersson, A., Coradeschi, S., & Loutfi, A. (2013). A review of mobile robotic telepresence. *Advances in Human-Computer Interaction*, *2013*, 3.

[12] Glass, T. A., de Leon, C. M., Marottoli, R. A., & Berkman, L. F. (1999). Population

based study of social and productive activities as predictors of survival among elderly Americans. *Bmj*, *319*(7208), 478-483.

[13] Webb, J. D., & Weber, M. J. (2003). Influence of sensory abilities on the interpersonal distance of the elderly. *Environment and behavior*, *35*(5), 695-711.

[14] Blevins, N., Deschler, D., & Park, L. (2014). Presbycusis. Retrieved September 15, 2014, from http://www.uptodate.com/contents/presbycusis

[15] Kuzuoka, H., Suzuki, Y., Yamashita, J., & Yamazaki, K. (2010, March). Reconfiguring spatial formation arrangement by robot body orientation. In *Proceedings of the 5th ACM/IEEE international conference on Human-robot interaction* (pp. 285-292). IEEE Press.

[16] Hüttenrauch, H., Severinson Eklundh, K., Green, A., & Topp, E. A. (2006, October). Investigating spatial relationships in human-robot interaction. In *Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on*(pp. 5052-5059). IEEE.

[17] Groh, G., Lehmann, A., Reimers, J., Frieß, M. R., & Schwarz, L. (2010, August). Detecting social situations from interaction geometry. In *Social Computing (SocialCom), 2010 IEEE Second International Conference on* (pp. 1-8). IEEE.

[18] Lau, B., Arras, K. O., & Burgard, W. (2010). Multi-model hypothesis group tracking and group size estimation. *International Journal of Social Robotics*,*2*(1), 19-30.

[19] Torta, E., Cuijpers, R. H., Juola, J. F., & van der Pol, D. (2011). Design of robust robotic proxemic behaviour. In *Social robotics* (pp. 21-30). Springer Berlin Heidelberg.

[20] Walters, M. L., Dautenhahn, K., Te Boekhorst, R., Koay, K. L., Syrdal, D. S., & Nehaniv, C. L. (2009). An empirical framework for human-robot proxemics. *Procs of New Frontiers in Human-Robot Interaction*.

[21] Joosse, M. P., Poppe, R. W., Lohse, M., & Evers, V. (2014, August). Cultural differences in how an engagement-seeking robot should approach a group of people. In *Proceedings of the 5th ACM international conference on Collaboration across boundaries: culture, distance & technology* (pp. 121-130). ACM.

[22] Jung, J., Kanda, T., & Kim, M. S. (2013). Guidelines for contextual motion design of a humanoid robot. *International Journal of Social Robotics*, *5*(2), 153-169.

[23] Kristoffersson, A., Eklundh, K. S., & Loutfi, A. (2013). Measuring the quality of interaction in mobile robotic telepresence: a pilot's perspective. *International Journal of Social Robotics*, *5*(1), 89-101.

[24] Mead, R., Atrash, A., & Matarić, M. J. (2013). Automated proxemic feature extraction and behavior recognition: Applications in human-robot interaction.*International Journal of Social Robotics*, *5*(3), 367-378.

[25] Kruse, T., Kirsch, A., Sisbot, E. A., & Alami, R. (2010, September). Exploiting human cooperation in human-centered robot navigation. In *RO-MAN, 2010 IEEE* (pp. 192-197). IEEE.

[26] Rehm, M., André, E., & Nischt, M. (2005). Let's come together—social navigation behaviors of virtual and real humans. In *Intelligent Technologies for Interactive Entertainment* (pp. 124-133). Springer Berlin Heidelberg.

[27] Hart, S. G. (2006, October). NASA-task load index (NASA-TLX); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting*(Vol. 50, No. 9, pp. 904-908). Sage Publications.

[28] Biocca, F., & Harms, C. (2003). Guide to the networked minds social presence inventory v. 1.2.

[29] Heerink, M., Kröse, B. J. A., Wielinga, B. J., & Evers, V. (2006). Studying the acceptance of a robotic agent by elderly users. *International Journal of Assistive Robotics and Mechatronics*, *7*(3), 33-43.