Grant agreement no: FP7-600877

SPENCER:

Social situation-aware perception and action for cognitive robots

Project start: April 1, 2013

Duration: 3 years

DELIVERABLE 4.2

Behaviour evaluation through user studies report

Due date: month 34 (January 2016)

Lead contractor organization: UT

Dissemination Level: PUBLIC

Main Authors:

Michiel Joosse (UT)
Manja Lohse (UT)
Vanessa Evers (UT)

Version History:

0.1, initial version, mj, June 2015
0.2, updated abstract, sections A, D.1-D.4 and conclusion, mj, October 2015
0.3, updated sections D.1 and D.5, mj, November 2015
0.4, updated conclusion, mj, January 2016

Contents

Abstract

The EU FP7 project SPENCER is concerned with the development and the deployment of an autonomous guiding robot. This robot is envisioned to guide passengers in a crowded (airport) environment from A to B. The SPENCER project will extend the state-of-the-art in socially normative human-robot interaction (HRI) by determining socially normative motion behaviours (in terms of spatial motion and head motion) that will significantly impact the users' acceptance of the robot in crowded environments and the ease of use in group interaction [1].

This deliverable follows up on D4.1, and summarizes the user studies conducted within task 4.1 of the SPENCER project. The basis of these user studies is a literature overview, of which we summarize the most relevant findings. We provide an overview of the concept of social norms. As a part of this, we discuss social psychological and HRI literature on norms such as personal space. Given that the SPENCER robot will encounter people with different cultural backgrounds, it is not unlikely these people have different understandings of what is appropriate robot behaviour. Therefore, we provide an overview of research into (human-robot) cross-cultural research. The capstone of the literature review is an overview of the state-of-the-art in HRI social norms, and the identification of gaps that we intended to address based upon this overview.

Following the literature overview, we introduce the seven user studies we have conducted as part of the SPENCER project.

A Introduction

The SPENCER robot will guide transfer passengers at an airport from their arrival gate to their gate of departure. Given that flying is not an everyday activity, such a transfer can be stressful. An ethnographic study by the SPENCER industrial partner KLM showed that even experienced flyers mostly associate negative emotions with the process of "transferring". In order to help transfer passengers during their transfer, a robot will guide passengers to their next gate.

Task 4.1 is concerned with the selection and evaluation of the SPENCER robot's spatio-temporal behaviors and ensuring that they adhere to social norms for the scenarios as specified for the project.

This deliverable follows up on D4.1 which was delivered in M12. In D4.1 we presented a literature overview, a summary of two studies conducted and research questions guiding our research during the next two years.

In this deliverable we present the same literature overview (Section B). The core of this report consists of the extended abstracts of studies conducted in year 1, (Section C, previously reported in D4.1) and studies conducted in years 2 and 3 (Section D). Full papers of the seven user studies are included at the end of this deliverable. After presenting the extended abstracts, we present a conclusion of the deliverable in Section E.

The *Functional noise study* (Section C.1) summarizes a lab experiment in which we added artificial noise to two robots that differed in height. In this section we show the importance of functional noise [25, 38] when approaching people. The *Culture spacing survey* (Section C.2) deals with cross-cultural human-robot interaction, and summarizes the results of a survey distributed to three countries (China, Argentina and the United States) [26, 24].

The *Contextual analysis* (Section D.1) describes a systematic observation we conducted at Schiphol airport to further refine the use case scenarios and identifies possible normative behaviors for the SPENCER robot. In Section D.2 we describe the *Telepresence murder study*, which we conducted in collaboration with the EU project TERESA[1]. In this study we measured a groups subjective and objective behavior when a robot partner joined and left the group [61]. In the *Robot appearance study* (Section D.3) we investigated different appearances of the robot in terms of head direction while driving [22]. We concluded that during driving the head of the robot should face forward, somewhat in contrast to previous research which indicated that a backwards driving robot is more effective at attraction people's attention [54]. The *Speed study* (Section D.4) contains two studies in which we investigated how fast the SPENCER robot should drive, both under general conditions and under the specific constraints of the SPENCER MCT use case. Finally, the *Social situations survey* (D.5) contains a study in which we asked people from China and the United States how they believe the SPENCER robot should respond to various social situations. From the results we could see that especially the group size- and time pressure affected the appropriateness of various robot responses.

[1]Telepresence Reinforcement-Learning Social Agent, Grant number EU-FP7-611153

B Related work

Parts of this section have been published as:
Joosse, M.P., Lohse, M., & Evers, V. (2014) Lost in Proxemics: Spatial Behavior for Cross-Cultural HRI. Proceedings of the 2014 ACM/IEEE Conference on Human-Robot Interaction Workshop in Culture Aware Robotics, 3 March 2014.

When human guides guide passengers around an airport, questions like "How fast should I walk?", "Am I walking on the right side of the corridor?", and "Should I overtake these people in front of me?" usually do not arise. People unconsciously negotiate their way even in crowded public space. However, these spatial behaviours are not as straightforward for a robot [23]. In order to program a robot to navigate as successfully as a human, we have to define how a robot should behave given certain situations [57]. Our main question is:

"Which social normative (motion) behaviours does a robot require to guide multicultural passengers from A to B at an international airport?"

To address this question, we are particularly interested in finding out what specific patterns, or rules, people apply when navigating semi-public spaces. Therefore, in the following we discuss important spatial behaviours like distancing (proxemics), speed, and normative pedestrian behaviours in social situations - for instance overtaking and walking direction.

A more fundamental, underlying question of our work is if social robots require different normative behaviours compared to humans. In Section B we will provide a summary of our findings, which includes an overview of (social) norms and examples of research into those norms. We will answer the following questions:

1. How do social norms affect our lives?
2. Do people treat robots as they treat humans?
3. Are there reasons to assume cultural differences with relation to existent norms?
4. Are there examples of social norms for robots?

B.1 How do social norms affect our lives?

Webster [40] defines a norm as *"a principle of right action binding upon the members of a group and serving to guide, control, or regulate proper and acceptable behavior"*. A social norm is a specific type of norm. While scholars from different fields of research employ different definitions for social norms (see for example [13], [39], [52], [50]), one widely-used definition by Cialdini & Trost [10] seems appropriate for this research project: *"rules and standards that are understood by members of a group and that guide and/or constrain social behavior without the force of laws"*. Prentice [47] provides a similar definition, but focuses somewhat more on the situational-dependent nature of norms: *"[...] defined as socially shared and enforced attitudes specifying what to do and what not to do in a given situation"*. Related to social norms are personal and legal norms, habits and customs. Both personal norms and habits are norms held by individuals, with the difference between the two being that the violation of a personal norm leads to some sort of personal sanction, the latter does not. A legal norm is a norm that is codified, thus called a law. As a working definition, influenced by [10] and [47] we define a social norm as *a rule, standard, or convention, understood by members of a group which guides and/or constrains social behavior without the force of laws in a certain context*. From our working definition, we can conclude that a norm has to meet three prerequisites in order to be considered to be a social norm:

1. The norm is unwritten
2. The norm exists in a specific social context
3. The norm is culturally dependent, therefore it can vary between cultures.

Examples of research into human adherence to social norms include series of experiments by Cialdini et al. [9], Aarts & Dijksterhuis [2] and Keizer et al. [27]. The norms researched were social norms like littering, being silent in a library and adherence to prohibition signs. While above research provides insightful results, these are not automatically relevant or applicable for the SPENCER project. A robot does not litter, and the speech volume of the robot can be programmed. We are looking into human spatial behaviour: the norms and rituals people adhere to and expect from a robot in comparable situations. There are two norms that we believe are especially relevant for the SPENCER robot: adherence to others' personal space and positioning with respect to others in space. In the next sections we describe how these two norms affect HRI.

B.1.1 Adherence to others' personal space

Adherence to others' personal space is one aspect of the research on "proxemics". Edward T. Hall [17] introduced the term proxemics and defined it as *"the interrelated observations and theories of humans' use of space as a specialized elaboration of culture"*. The proxemics theory postulates that ones' body is surrounded by four eclipse-shaped bubbles: the intimate, personal, social and public spaces. Interaction partners mostly position themselves in the second- and third zone. For each of these zones, Hall defined approximate distances, as can be seen in Figure 1 and Table 1. In line with previous HRI work we will primarily focus on the identification of the personal space zone, in order to find out when a robot approaches too close for comfort.

There are several effects of personal space invasion. When one's personal space zone is invaded by "intruders" in semi-public spaces this is considered as a disturbance, an invasion of one's personal territory. Reactions to the invasion of personal space include avoidance behaviours to compensate for this invasion. The distance between two people is increased, for instance by avoiding eye contact

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

Table 1: Proxemics zones as defined by Hall [17]

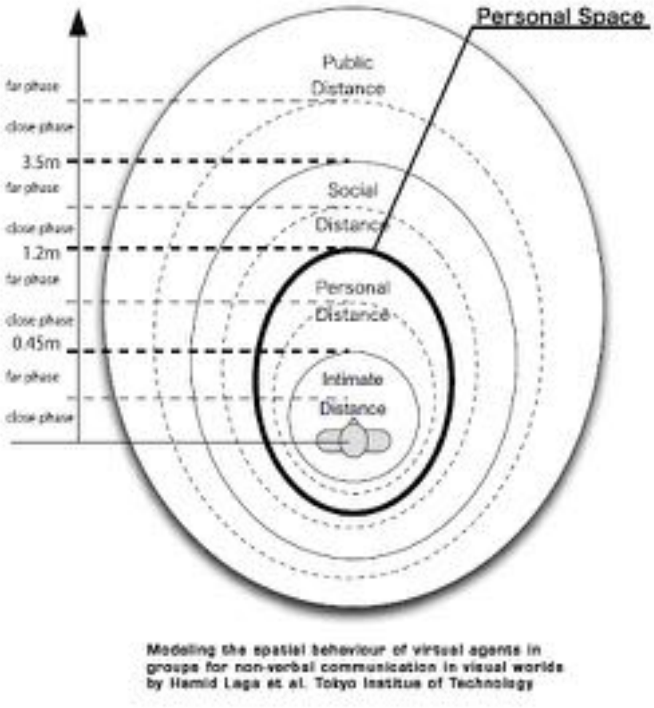


Figure 1: Proxemics zones as defined by Hall [17]. Image source [33]

(gaze), or physical flight behaviours, like leaning away or walking away entirely. Mediating factors influencing the size of one's personal space bubble include (but are not limited to) gender, age, personality, socioeconomic status, sociability, interpersonal likeability / attraction and gaze [19].

A detailed overview can be found in Hayduk [19]. The size of the proxemics zones is culturally dependent, as is explained in Section C.2. This is relevant for the SPENCER project as the SPENCER robot is envisioned to interact with a culturally diverse audience. Therefore, HRI research is needed into cross-cultural HRI proxemics.

The norm of adherence to personal space has received a lot of attention with respect to other factors though. Different scholars in HRI, most notably Michael Walters et al., researched whether or not the size of the personal space bubble would be equal when approached by a robot. Additional studies looked into which factors influence successful robot-human approaches. The work on HRI-related social norms relevant for SPENCER is summarized in Table 2. This research found that the baseline personal space distance for humans is around 57 centimeters and varies slightly depending on factors such as the appearance of the robot and the preference of the participant [65]. It has also been found that a full-frontal approach direction is not always considered to be the most appropriate [11]. However, most of the previous work focused on approaching single persons. For the SPENCER project this provides a basis upon which we can design our future user studies when we are going to approach small groups of people.

What?	Reported in
Approaching	
Personal space (zone)	
Identification of approach distance	[21, 64]
Influence of age factor	[63, 44]
Influence of gaze behaviour	[42, 58]
Influence of robot voice	[65]
Influence of approach angle	[31]
Negative effects of violation of personal space zone	[51, 23]
Approach angle while seated	[57, 11] [11]
Effect of robot height	[8, 25, 62]
Temporal stability	[31]
Spatial (F) Formations	
Formation assumed around a robot	[21]
Formation influenced by a robot	[32]
Social robot conventions while driving	[46, 48]

Table 2: Summary of HRI experiments related to physical robot behaviour

B.1.2 Positioning with respect to others (*F*-Formations)

People organize themselves not only in terms of interpersonal distance but also in terms of spatial arrangements. To capture this phenomenon, Kendon [28] introduced the concept of *F*-Formations. The space in which people direct their attention and manipulate objects can be called a transactional segment [28]. When two or more people interact, these segments overlap, thus, creating a joint transactional space: the O-space. Around the O-space, people arrange themselves in the P-space, in a certain *F*-Formation. Behind the P-space is the R-space which is everything not in the O- or P-space. These spaces have also been referred to as transactional region, agent region and buffer region [35].

Figure 2 shows an example formation with the O-, P-, and R-spaces. These three people are standing in a circular formation, however, also other formations are frequently observed; for instance, L-shape, vis-a-vis and circular formations. This is highly relevant for SPENCER because the robot has to approach groups of people. The formation and composition (size, male/female distribution, cultural background) of the group has an influence on which direction and distance of approach is considered to be the most appropriate. In Section C.2 we describe a survey that we conducted to gain insights into what would be considered being the most appropriate approach direction and distance.

B.1.3 Guiding people

Guiding people is believed to be a major application area for (social) robots. While the context may differ (e.g. an airport or a museum [53]) there are people being guided, and therefore there are similarities. One of these similarities might lie in the speed of the robot. Research by Garrell & Sanfeliu [14] showed that people's interest in following a robot decreased when the speed of the robot decreased to 0.4 m/s. A speed of 0.8 m/s seemed to be more appropriate. This is still not as fast as the average walking speed observed by researchers as Bohannon [6], so it could be the case

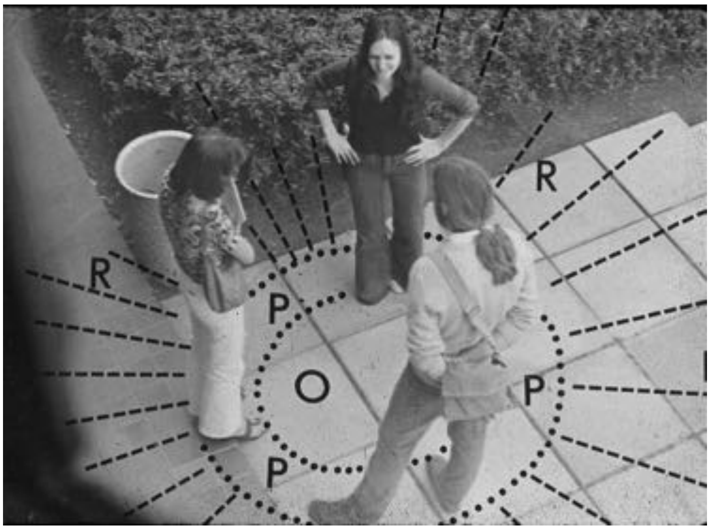


Figure 2: Typical circular *F*-formation [28].

the SPENCER robot could still comfortably guide passengers at a higher speed. Thus, incoherences in previous findings need to be addressed in order to determine the optimal speed for the SPENCER robot (see Section B.4).

Next to speed, a guiding robot should communicate where it is going in one way or the other [30], not only for the passengers being guided but also for people being present around the robot. Humans use head movement and non-verbal cues (facial expressions) to communicate such intentions. Modalities such as gaze and sound could also be used in HRI as our previous work has indicated [25, 38]. However, these aspects will need to be researched more in depth with respect to guiding behaviours (see Section B.4).

B.2 Do people treat robots as they treat humans?

When interacting with robot technology, or computer technology in general, people attribute certain human characteristics to both its appearance and behaviour. The term *anthropomorphism* describes the tendency to imbue the real or imagined behaviour of non-human agents with humanlike characteristics, motivations, intentions, or emotions [12]. In robot design, for instance, people have been found to automatically perceive anthropomorphic cues, for instance in that they prefer a more sociable robot head for a more social task [16].

Just as people anthropomorphize a robot's appearance, people sometimes e.g. scold at their computer in an attempt to make it work. This phenomenon can be explained by means of the Media Equation theory, also known as the CASA - Computers As Social Actors - paradigm. The Media Equation holds that people tend to treat computers and other media as if they were either real people or real places [49]. Supporting evidence has been found in the field of Human-Computer Interaction (HCI), e.g. participants attributed introverted or extroverted personality to a computer that read aloud book descriptions by ways of a TTS engine [43]. Similar evidence has been found in HRI literature: Lee et al. [34] programmed an AIBO robot to behave either introvert or extrovert and found that participants were able to distinguish between both personalities. This implies that the CASA paradigm might be equally valid for HRI as it is in HCI. However, HRI research has also revealed limitations of the Media Equation. Bartneck et al. [4] replicated Stanley Milgram's famous prison experiment [41] with a robot in place of the human. Unlike Milgram's experiment where 40% of the participants went up to the highest voltage setting to punish a person in another room, 100% of participants did so when asked to punish a robot [4]. While this could be considered to be an extreme example, it nicely points out that despite the fact that we might treat robots similar to humans to a certain extend, it is still necessary to study how norms in HRI differ.

Numerous HRI publications refer - either implicitly or explicitly - to the importance of conforming to social norms, the most popular of all norms appearing to be personal space (see Section B.1.1). Qian et al. [48] conducted a survey in order to identify important, socially acceptable, safety criteria for robots. Out of eight possible social conventions, four were reported as being the most important, according to the participants of the survey. These conventions are (1) adhere to personal space, (2) maintain visibility when approaching, (3) drive at the correct side of a hallway, and (4) give priority to humans should a robot and human appear to be on a collision course. All of these are included in the research questions that we will present in the conclusion of this section.

B.3 Are there reasons to assume that cultural differences with relation to norms (for robots) exist?

Human interaction is not governed by one specific set of social norms but the norms differ across cultures. Unfamiliarity with cultural differences can lead to misinterpretation, misunderstanding and even unintentional insult [59]. Culture is an ambiguous concept, therefore, we will first look at different definitions of culture. Originally, the term culture stems from the Latin word colore, and it *"[...] usually referred to something that is derived from, or created by the intervention of humans - culture is cultivated"* [18]. Triandis [60] divided culture into a subjective and material culture. Material culture consists of elements, for instance food, houses and tools. Subjective culture, on the other hand refers to the characteristic way in which a specific group perceives its environment [60]. When referring to culture, this review refers to subjective culture.

Brauer & Chaurand [7] compared 46 uncivil behaviours across eight countries, which varied along Hofstede's individualism-collectivism axis. For each of the behaviours, participants were asked to rate how uncivil they thought the behaviour was, how common it was in their country and how likely they would be to react negatively to that behaviour. Results indicated that if the behaviour was perceived as more deviant, participants would be more likely to react to it. In a similar study by Gelfand et al. [15], participants (N=6823) from over 33 countries were asked to rate the appropriateness of twelve behaviours in fifteen everyday situations, and, whether or not there were clear rules for appropriate behaviours in these situations. It was found that there was a high within-nation agreement about the level of constraint in everyday situations, and a high level of variability between-nations. The nation as unit of analysis appears to have proven to be a useful unit of analysis.

In the field of HRI, cross-cultural research has not yet focused on social norms. Cross-cultural HRI research up to date has primarily focused on general attitudes toward robots [5], and whether or not the mental model people have of robots is culturally-dependent [29]. Wang et al. [66] found people to be more willing to follow a robot's advice when the robot would provide advice in a culturally appropriate way. However, for the SPENCER project it is important that we investigate whether proxemic expectations of users toward the robot depend on culture. Research by social psychologists like Sussman & Rosenfeld [56] and Little [36] provides support for this hypothesis in human interaction. In Section C.2 we provide some empirical evidence for this hypothesis with respect to HRI.

B.4 Identifying research questions

The SPENCER robot has to execute four primary movement tasks: Approach Me, Walk With Me, Talk To Me, and Leave Me. The “me” in this is the participant, or in case of the SPENCER project the group of participants. Based upon our literature review we have gained an overview of what has and what has not yet been tested in relation to these tasks. We sum this up in the following with the goal to identify open questions that we have already answered in the studies presented below and/or will address in our future research.

HRI trials have focused on approaching and engaging single persons in the lab (e.g., [11, 21, 25, 31, 42, 58]). The SPENCER robot will have to interact with multiple passengers. Even though it might be possible to direct the robot’s focus of attention on a spokesperson, approaching a small group could lead to different proxemics preferences (and thus expectations) since members of the group influence each other. A gap in the state of the art is therefore the approach (initial contact) and engagement of small groups of people.

We have briefly discussed F-formations and associated different interaction spaces. From this, we know that a robot can influence the formation of the group by adjusting its position when part of the group [32]. What we do not know is how a robot should approach a group of people and get their attention when they are standing in a specific F-formation. This gap is not only relevant for a guiding robot but basically for any attention-seeking robot.

Also little is known about an appropriate robot speed when guiding people. As guiding passengers is the *raison d’etre* of the SPENCER robot, we intend to conduct an experiment to understand how quickly the robot should move in certain situations. One specific situation that is important in the SPENCER context is how the robot can cause groups of passengers to walk quicker when being guided.

Approaching a group has been discussed in the literature (see Table 2). While there are preferences for approach direction and -distance, it has not yet been investigated how a robot should leave a group.

From this we identified several research questions related to the four primary movement tasks:

Approach me / Talk to me

1. How close and from what angle should a robot approach a group of passengers in order to engage them in interaction?
   See Section C.2.

2. How does normative robot behaviour in terms of how close and from what angle to approach differ between situations with individuals compared to groups?
   See Section D.2.

Walk with me

3. How should a robot's motion behaviour be designed to compensate for its lack in (facial) expressiveness of non-verbal cues?
   See Sections C.1 and D.3.

4. a) What is the most appropriate robot speed when guiding a group of passengers?
   b) How can a robot make a group of passengers speed up / hurry?
   See Section D.4.

5. How should a robot keep passengers engaged while guiding them over a longer distance (>100 meter)?
   See Section D.3.

6. Do people expect a robot to give priority of way to elderly people at an airport?
   See Section D.5.

Leave me

7. How should a robot leave a small group of transfer passengers in a culturally appropriate way?

C Summary of experiments conducted in year 1

This section summarizes two studies we have conducted in year 1. An overview of symbols used to report the various statistical tests can be found in Section G.

C.1 Functional noise study

This section has been published as:
Joosse, M.P., Lohse, M., & Evers, V. (2014) Sound over matter: the effects of functional noise, robot size and approach velocity in human-robot encounters. Proceedings of the 2014 ACM/IEEE Conference on Human-Robot Interaction, pp. 184-185

We conducted a 2x2x2 between-groups experiment (N=80), manipulating the robot height, the approach velocity of the robot and the use of functional noise (increasing in volume when the robot accelerated and decreasing in volume when the robot decelerated). See also Figure 3. Our hypothesis was that *"a robot using functional noise to convey its intention to the user will be more positively perceived than a robot which does not use intentional functional noise"*.

When we combined this dataset with the one in [38], we found a significant main effect of functional noise on helpfulness. Participants found an intentional noise pattern (M=3.35, sd=1.122) significantly more helpful than a constant functional noise pattern (M=2.73, sd=.987), U=546.00, Z=-2.546, $p<0.05$. Furthermore, we found significant (2-tailed) main effects for functional noise on all Godspeed scales [3]: anthropomorphism (F(1,73)=7.685, $p<0.01$), animacy (F(1,75)=7.474, $p<0.01$), likeability (F(1,75)=9.336, $p<0.01$), perceived intelligence (U=520.00, Z=0.10, $p<0.01$) and perceived safety (U=607.50, Z=0.059, $p<0.05$). For the above scales the intentional noise conditions

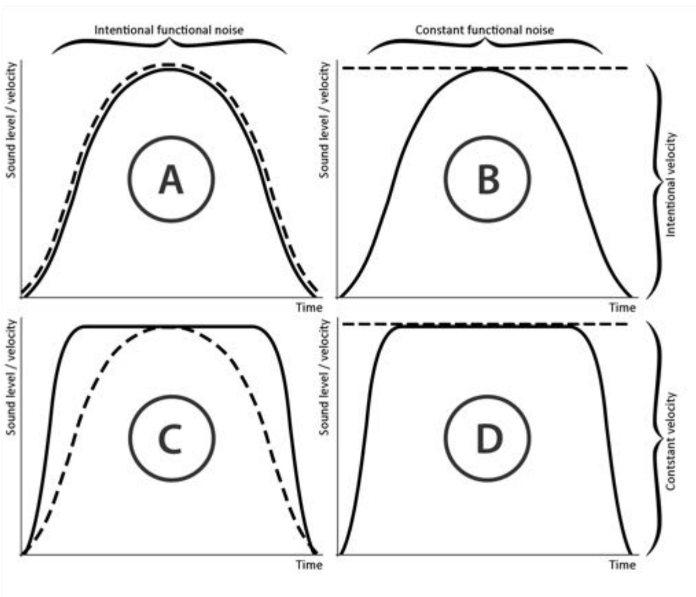


Figure 3: A 2x2 between-groups experiment was conducted, manipulating functional noise and velocity.

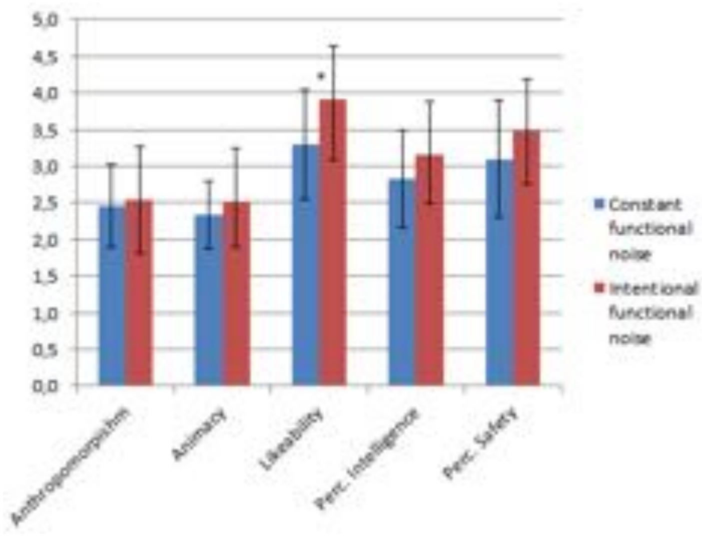


Figure 4: Mean ratings for the combined sample Godspeed scales.

were rated more positively than the constant noise conditions (Figure 4). No main or interaction effects for height were found for any of the Godspeed scales; except for perceived safety. Participants perceived the short Magabot (M=3.69, sd=.94) as safer than the taller Giraff (M=3.29, sd=.77), U=611.00, Z=-1.851, $p<0.05$ (1-tailed).

In conclusion, we found that a robot approaching with intentional noise (increasing in volume when the robot accelerated and decreasing in volume when the robot decelerated) was regarded more positively. Our study shows that functional noise could be a powerful tool to convey a robot's intentions when approaching a user.

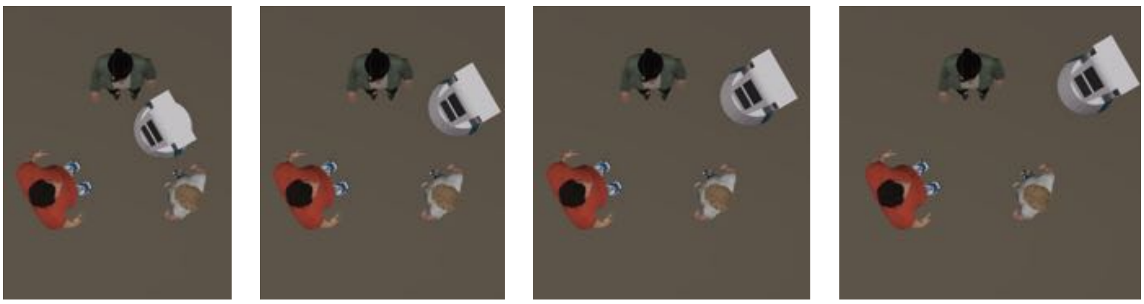
C.2 Culture spacing survey

This section has been published as:
Joosse, M.P., Poppe, R.W., Lohse, M., & Evers, V. (2014) Cultural Differences in how an Engagement-Seeking Robot should Approach a Group of People. Proceedings of the 5th ACM Conference on Collaboration Across Boundaries: Culture, Distance & Technology (CABS), pp. 121-130

The sociologist Hall [17] coined the term proxemics to indicate the studies of human's use of space. Social psychological research found these interpersonal distances to be culturally dependent [20, 36, 56] as previously explained in Section B.1.1. We set out to extend the state of the art on human-robot proxemics by investigating whether preferences for how a robot should approach a small group is culturally dependent. We present our first study in this, a set of measures and preliminary results of an online survey (N=181) distributed to people in China, the U.S. and Argentina. We chose these countries because related research indicated these three national countries would have significant different practices and values and that there would therefore have different proxemic expectations. A more detailed explanation is provided in the paper, which is included in the Appendix.

We conducted a 3 (nationality) x 3 (position in the group) x 6 (distance from the group) online study. A survey-based questionnaire was distributed through a crowdsourcing platform (crowdflower.com) to a targeted population. Participants were shown images of small families of 3D people and a robot (See Figure 5). These groups were composed of three people: a man, a woman and a child. Participants were asked to indicate on a 7-point Likert scale how appropriate they believed the position of the robot was after the robot had approached. The position and the distance of the robot were manipulated within-subjects, the nationality of the participants was a between-subjects variable.

Our results show that participants prefer a robot which stays out of people's intimate space zone just like a human would be expected to do. The cultural differences found were partly in line with previous socio-psychological research: Chinese participants believed that closer approaches were appropriate compared to the participants from the U.S. and Argentina. For the SPENCER project, this implies that we actually have to take the culture of the passengers into account.



(a) (b) (c) (d)

Figure 5: Example top-down stills as shown to participants.

D Summary of experiments conducted in years 2-3

This section summarizes two studies we have conducted in year 2 and 3. An overview of symbols used to report the various statistical tests can be found in Section G.

D.1 Contextual Analysis

This section has been published as:
Joosse, M.P., Lohse, M., & Evers, V. (2015) How a Guide Robot Should Behave at an Airport - Insights Based on Observing Passengers. Technical Report TR-CTIT-15-01, Centre for Telematics and Information Technology, University of Twente, Enschede. ISSN 1381-3625

As part of the input for future robot development, we conducted a contextual analysis at Schiphol Airport. The primary goal of the contextual analysis is to understand the behaviour of passengers (specifically transfer passengers) at Schiphol. This includes (observable) rituals and habits that guide their walking behaviour, and their needs when transferring.

Based upon the SPENCER use case[2] six situations and locations of interest were specified. Data was collected on 11 and 12 June 2014. This data consisted mainly of recorded video, specifically at various locations in the Non-Schengen part of the terminal - in particular lounges 2 and 3, and gates D, E, and F. Data analysis has been conducted in three phases, which were based upon an inductive data analysis approach by [37]. Phase 1 consisted of preparing the raw video data for analysis, which included merging, splitting and renaming video files. During phase 2, the open coding phase, a first categorization was made using open coding [55, p.28]. In the third phase, this coding was refined and all video material was cut and placed into one or multiple of the subcategories. This resulted in a total of 406 short video clips in 48 subcategories.

For each subcategory we described the behaviour which was observed, where possible supplemented with descriptive statistics. Each section contains conclusions, which are implications for the behaviour or for the perception capabilities of the SPENCER robot. These implications, or guidelines, are included in Appendix G.

[2]Deliverable 1.1

(a) People mostly walk on the right side of the corridor, even when not divided by a moving walkway

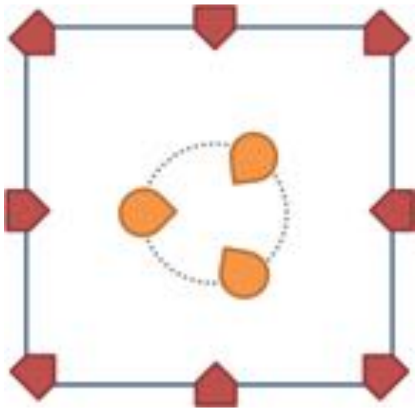
(b) Groups organize them in semi-circular formations around information monitors

Figure 6: Examples of data gathered during the contextual analysis

D.2 Telepresence murder study

The study reported in this section has been in collaboration with the EU FP7-project TERESA[3]*, and has been published as:*
Vroon, J.H., Joosse, M.P., Lohse, M., Kolkmeier, J., Kim, J., Truong, K.P., Englebienne, G., Heylen, D.K.J. & Evers, V. (2015) Dynamics of Social Positioning Patterns in Group-Robot Interactions. Proceedings of the 24th IEEE International Symposium on Robot and Human Interactive Communication (ROMAN)

(a) Groups interacted with a telepresence robot



(b) The robot approached the group (yellow markers) from eight different locations (red markers)

Figure 7: Groups of four participants solved a murder mystery, while one of the members approached, engaged and left the group from eight directions.

In this study groups of four people solved a collaborative task together (solve a fictional murder). One of the members of the group was present through a Giraff telepresence robot (called the Visitor), while the other three members (Interaction Targets) were asked to stand in one of two F-formations (Figure 7a). The Visitor was the only member who had access to vital information required to solve the task. Throughout the experiment the Visitor had to approach, conserve and leave the group 8 times, from various directions (Figure 7b). From the subjective and objective data we recorded we gathered patterns of spatio-temporal motion behaviour, both from the Visitor and the other group members.

We found the following patterns:

- 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.
- 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.
- In 33.9% of the retreats we observed (38 out of 112), to our surprise, that Visitors passed

[3]Telepresence Reinforcement-Learning Social Agent, Grant number EU-FP7-611153

straight through the group. In 16 of these 38 retreats the Visitor communicated this beforehand with the group.

- From the Interaction Target's subjective evaluation we learned that approaching with a smooth and steady path seems to be important for the average normalized ratings. Additionally we observed that the Visitor stopped at on average 1.25 meter from and aimed at the center of the group.
- In nine out of the ten highest rated retreats we saw that the Visitors explicitly communicated their goals (verbally) before driving.

D.3 Robot appearance study

This section has been published as:
Joosse, M.P., Knuppe, R.A., Pingen, G.L.J., Varkevisser, R.A., Vukoja, J., Lohse, M. & Evers, V. (2015) Robots Guiding Small Groups: The Effect of Appearance Change on the User Experience. Proceedings of the 4th International Symposium on New Frontiers in Human-Robot Interaction

We conducted 2 exploratory user studies in which we investigated subtle changes of a guide robot's appearance on the subjective evaluation by users. For this study a shell was attached on top of a remote-controlled Robotino robot platform[4] (Figure 8a). The height of the robot was 170cm and it drove at a speed of approximately 0.7 m/s. In both studies participants were provided a small guided tour (Figure 8b).

In the first study we manipulated the appearance of the robot in terms of the position of a tablet providing information (facing the group that was guided or the walking direction) and the type of information displayed (eyes or route information). 25 participants divided over 9 groups participated in this study. Our results indicate that the location of the screen can be either forward or backward, depending on the information displayed. In the case of eyes facing participants, our results showed that this was considered to be very unnatural and intimidating. On the other hand, when the tablet faced participants and route information was provided this was again evaluated as more useful.

(a) Modified robot

(b) A group being guided

Figure 8: For the robot appearance studies a shell was mounted on a Robotino platform.

Based upon comments received from participants a second study was conducted. In this study we focused specifically on the effect of different input modalities on user satisfaction. A total of 3 conditions were designed containing different input modalities (vocal or touch) and as with the first study either route information or eyes were displayed. 19 participants divided over 9 groups participated in this study. Participants showed a dislike for vocal input, however, it should be noted that this system had technical difficulties which could be the cause of this dislike.

In both studies participants commented on the low speed of the robot. Partially based upon this comment a series of studies has been conducted into robot speed while guiding (Section D.4).

[4]http://www.festo-didactic.com/int-en/learning-systems/education-and-research-robots-robotino/

D.4 Speed study

We conducted a study to investigate how people value different speeds of the SPENCER platform, within a fictional airport context, both under time pressure and without time pressure.

In this study we aimed at getting a benchmark of a triad's walking speed, with and without luggage, under various level of time pressure. Related work into human walking speed provides various studies studying the speed of both individual and groups of people, however, as we could not find this specific situation we conducted this study. The experiment was a modified version of a 6-Minute Walking Distance test [45]. Following the instructions the group walked laps around a 30-meter line, after 3 minutes we measured how far they had walked. The use of luggage was manipulated within-subjects, using counterbalancing to avoid learning effects. Participants were instructed using three different sets of instructions regarding time pressure (manipulated between-subjects):

1. "Your previous flight has been delayed due to bad weather, however, you have more than enough time to walk to your connecting flight." (**No pressure**)
2. "Your previous flight has been delayed due to bad weather, therefore, you will have to hurry a bit to catch your next flight." (**Mild pressure**)
3. "previous flight has been delayed due to bad weather, therefore, you have to hurry up in order to catch your next flight." (**Heavy pressure**[5])

39 participants, divided over 13 groups, together walked 15.451 meters. Table 3 shows the mean distance groups traversed based upon time pressure condition, and whether or not they carried luggage.

Table 3: Distance traversed by groups, based upon time pressure condition

	Without luggage		**With luggage**	
	(meters)	m/s	(meters)	m/s
A	155.48	0.86 m/s	148.36	0.82 m/s
B	240.60	1.34 m/s	224.00	1.24 m/s
C	237.43	1.32 m/s	228.04	1.27 m/s

In order to analyze the effect of luggage on the distance walked we conducted separate Mann-Whitney tests for each of the three time pressure condition. Participants walked less when carrying luggage, though this effect was only significant in condition B (U=27.5, Z=-2.570, $p<0.01$). The distance groups traversed different significantly by condition ($H(2)$=52.57, $p<0.001$). Post-hoc Mann-Whitney tests indicated the effect being condition A and B being significant different (U=.000, Z=-6.267, $p<0.001$), and condition A and C as well (U=18.000, Z=-5.954, $p<0.001$). This indicates that when provided some form of pressure the groups walked further. The specific type of condition did not seem to influence the results in this study.

Based upon the results of the benchmark study, a follow-up study will be designed with the SPENCER platform.

[5] SPENCER MCT scenario

D.5 Social situations study

This study concerned "everyday people's ideas of how the SPENCER robot should respond to various situations at the airport". We conducted a survey in which we give people various situations, and asked them to write down in a text box how they believed the SPENCER robot should respond to that situation. We defined 10 social situations which are likely to occur when a robot guides a group at an airport.

Furthermore, we manipulated four variables, which resulted in 8 between-groups conditions [6]

1. Time-pressure (between-subjects): participants are primed beforehand that they *have to hurry up* or that they *have more than enough time to walk to their connecting flight*.
2. The situation concerns either the minority or majority of the group (between-subjects).
3. Participants are either involved or not involved in the situation (within-subjects), e.g. you are part of the people who have go to the bathroom, or it's just other people who have to go to the bathroom.
4. Participants are part of a small or larger group of people (between-subjects). This was also enforced by a specific image participants were shown when completing the survey.

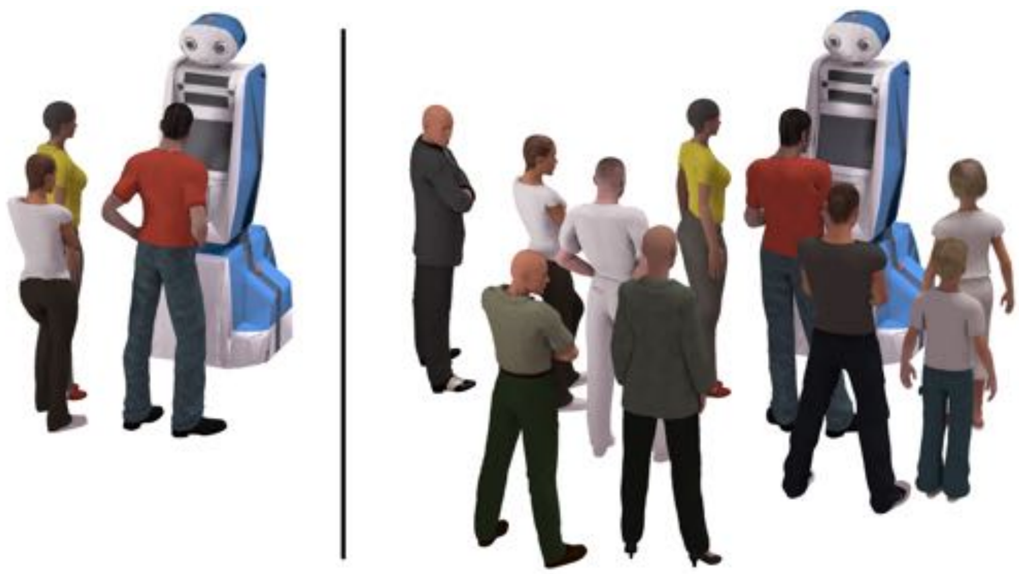


Figure 9: Participants in the survey were part of either a small or larger group.

Based upon the answers of 118 people, we drew conclusion regarding the behavior the robot should display. The general conclusions are provided in Table 4, and the full report can be found in the appendix. As can be seen in the table, the size of the group and the time pressure factors were especially relevant when determining the appropriate response by the robot.

[6]The specific instructions and social situations are available at https://svncvpr.informatik.tu-muenchen.de/redmine/attachments/download/243/situations_for_web.pdf

Table 4: Robot action to be conducted in various social situations to be encountered at an airport.

Situation	Applies to	Robot action
1. Fall behind	(all)	Slow down
	Large group	Slow down, or wait
2. Forgot passport	(all)	Continue
	Low pressure	Provide information
3. Enter a store	(all)	Wait
4. Go to the bathroom	(all)	Wait
	Small group	Provide information
5. Elderly crosses road	(all)	Move around the people
	Low pressure	Wait
	Large group	Wait
6. Tie shoelace	(all)	Wait
7. Someone is lost	China	Assist person
	United States	Continue
8. Take a selfie		Wait
9. Take a smoke	(all)	Continue
10. Meet familiar person	Low time pressure	Wait
	High time pressure	Continue
	Majority	Wait
	Minority	Continue

E Conclusion

In this deliverable, we have summarized our research activities over the past three years. We have conducted a literature review from which we defined seven research questions which we believed were especially relevant for the SPENCER project:

1. How close and from what angle should a robot approach a group of passengers in order to engage them in interaction?
2. How does normative robot behaviour in terms of how close and from what angle to approach differ between situations with individuals compared to groups?
3. How should a robot's motion behaviour be designed to compensate for its lack in (facial) expressiveness of non-verbal cues?
4. a) What is the most appropriate robot speed when guiding a group of passengers?
   b) How can a robot make a group of passengers speed up / hurry?
5. How should a robot keep passengers engaged while guiding them over a longer distance (>100 meter)?
6. Do people expect a robot to give priority of way to elderly people at an airport?
7. How should a robot leave a small group of transfer passengers in a culturally appropriate way?

We have addressed most of these questions through 7 user studies, of which 4 lab experiments, 1 observation study and 2 online surveys investigating cultural differences regarding attitudes towards robot behavior in China, the United States and Argentina. Findings have been incorporated in the SPENCER demonstrator robot, for example the results head direction study. While not all findings have been incorporated, studies such as the social situations survey and the telepresence murder study provide input for future projects featuring social robots in semi-public spaces.

One question we have not addressed is the last question, concerning the “leaving of the group”, and ending the interaction. Due to time constraints this was not possible, though we intend to include this question in the final user studies (D6.5) in March 2016 in order to draw at least some rudimentary conclusions and recommendations.

In general we believe that with these studies we have made a valuable contribution both to the SPENCER project and the HRI community. We look forward to the deployment and final evaluation of the SPENCER project.

F Acknowledgements

The authors wish to thank Robin Knuppe, Geert Pingen, Rutger Varkevisser, Josip Vukoja and Sanne van Waveren, all master students enrolled in the master Human Media Interaction of the University of Twente, for their contribution to this deliverable. The authors would furthermore like to thank Jered Vroon, Jan Kolkmeier, Khiet Truong, Gwenn Englebienne and Dirk Heylen (EU TERESA project) for their assistance in collection data for the *Telepresence Murder study*.

G Symbols used in this deliverable

The following symbol are used in this deliverable:

Symbol	**Denotes**
N	Sample size
M	Mean, or average
sd	The standard deviation of a sample of data
T	Test statistic for the T-test
H	Test statistic for the Kruskal-Wallis test
U	Test statistic for the Mann-Whitney test
F	F-ratio: the test statistic used in an analysis or variance (ANOVA)
p	Probability
Z	A data point expressed in standard deviation units

References

[1] Spencer description of work, 2011.

[2] H. Aarts and A. Dijksterhuis. The silence of the library: Environment, situational norm, and social behavior. *Journal of Personality and Social Psychology*, 84(1):18–28, 2003.

[3] C. Bartneck, D. Kulić, E. Croft, and S. Zoghbi. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International journal of social robotics*, 1(1):71–81, 2009.

[4] C. Bartneck, C. Rosalia, R. Menges, and I. Deckers. Robot abusea limitation of the media equation. In *Proceedings of the interact 2005 workshop on agent abuse, Rome*, 2005.

[5] C. Bartneck, T. Suzuki, T. Kanda, and T. Nomura. The influence of peoples culture and prior experiences with aibo on their attitude towards robots. *AI & SOCIETY*, 21(1-2):217–230, 2007.

[6] R. W. Bohannon. Comfortable and maximum walking speed of adults aged 2079 years: reference values and determinants. *Age and ageing*, 26(1):15–19, 1997.

[7] M. Brauer and N. Chaurand. Descriptive norms, prescriptive norms, and social control: An intercultural comparison of people's reactions to uncivil behaviors. *European Journal of Social Psychology*, 499(June 2009):490–499, 2010.

[8] J. Butler and A. Agah. Psychological effects of behavior patterns of a mobile personal robot. *Autonomous Robots*, 10(2):185–202, 2001.

[9] R. B. Cialdini, R. R. Reno, and C. A. Kallgren. A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology*, 58(6):1015–1026, 1990.

[10] R. B. Cialdini and M. R. Trost. Social influence: Social norms, conformity and compliance. 1998.

[11] K. Dautenhahn, M. Walters, S. Woods, K. L. Koay, C. L. Nehaniv, A. Sisbot, R. Alami, and T. Siméon. How may i serve you?: a robot companion approaching a seated person in a helping context. In *Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-Robot Interaction*, pages 172–179, doi:10.1145/1121241.1121272. ACM, 2006.

[12] N. Epley, A. Waytz, and J. T. Cacioppo. On seeing human: a three-factor theory of anthropomorphism. *Psychological review*, 114(4):864–86, 2007.

[13] E. Fehr and U. Fischbacher. Social norms and human cooperation. *Trends in cognitive sciences*, 8(4):185–90, 2004.

[14] A. Garrell and A. Sanfeliu. Cooperative social robots to accompany groups of people. *The International Journal of Robotics Research*, 31(13):1675–1701, 2012.

[15] M. J. Gelfand, J. L. Raver, L. Nishii, L. M. Leslie, J. Lun, B. C. Lim, L. Duan, A. Almaliach, S. Ang, J. Arnadottir, Z. Aycan, K. Boehnke, P. Boski, R. Cabecinhas, D. Chan, J. Chhokar, A. D'Amato, M. Ferrer, I. C. Fischlmayr, R. Fischer, M. Fülöp, J. Georgas, E. S. Kashima, Y. Kashima, K. Kim, A. Lempereur, P. Marquez, R. Othman, B. Overlaet, P. Panagiotopoulou, K. Peltzer, L. R. Perez-Florizno, L. Ponomarenko, A. Realo, V. Schei, M. Schmitt, P. B. Smith, N. Soomro, E. Szabo, N. Taveesin, M. Toyama, E. Van de Vliert, N. Vohra, C. Ward, and S. Yamaguchi. Differences between tight and loose cultures: a 33-nation study. *Science (New York, N.Y.)*, 332(6033):1100–4, 2011.

[16] J. Goetz, S. Kiesler, and A. Powers. Matching robot appearance and behavior to tasks to improve human-robot cooperation. In *Proceedings of the 12th IEEE International Workshop on Robot and Human Interactive Communication. ROMAN 2003*, pages 55–60, doi:10.1109/ROMAN.2003.1251796. IEEE, 2003.

[17] E. T. Hall. *The Hidden Dimension*. 1966.

[18] T. Hamamura. Are cultures becoming individualistic? a cross-temporal comparison of individualism-collectivism in the united states and japan. *Personality and Social Psychology Review*, 16(1):3–24, 2012.

[19] L. A. Hayduk. Personal space: Where we now stand. *Psychological Bulletin*, 94(2):293–335, 1983.

[20] H. Hø gh Olesen. Human spatial behaviour: The spacing of people, objects and animals in six cross-cultural samples. *Journal of Cognition and Culture*, 8(3):245–280, 2008.

[21] H. Hüttenrauch, K. Severinson Eklundh, A. Green, and E. A. Topp. Investigating spatial relationships in human-robot interaction. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 5052–5059, doi:10.1109/IROS.2006.282535. IEEE, 2006.

[22] M. Joosse, R. Knuppe, G. Pingen, R. Varkevisser, J. Vukoja, M. Lohse, and V. Evers. Robots guiding small groups: The effect of appearance change on the user experience. In M. Salem, A. Weiss, P. Baxter, and K. Dautenhahn, editors, *Proceedings of the 4th International Symposium on New Frontiers in Human-Robot Interaction*, pages 6–8, Canterbury, UK, 2015. University of Kent.

[23] M. Joosse, M. Lohse, and V. Evers. Short duration robot interaction at an airport: challenges from a socio-psychological point of view. 2013.

[24] M. Joosse, M. Lohse, and V. Evers. Lost in proxemics : Spatial behavior for cross-cultural hri. In *Proceedings of the 2014 HRI Workshop on Culture Aware Robotics*, 2014.

[25] M. Joosse, M. Lohse, and V. Evers. Sound over matter: the effects of functional noise, robot size and approach velocity in human-robot encounters. In *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*, pages 184–185. ACM, 2014.

[26] M. P. Joosse, R. W. Poppe, M. Lohse, and V. Evers. 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, pages 121–130. ACM, 2014.

[27] K. Keizer, S. Lindenberg, and L. Steg. The spreading of disorder. *Science (New York, N.Y.)*, 322(5908):1681–5, 2008.

[28] A. Kendon. *Conducting interaction: Patterns of behavior in focused encounters*, volume 7. CUP Archive, 1990.

[29] S. Kiesler and J. Goetz. Mental models of robotic assistants. In *CHI'02 extended abstracts on Human Factors in Computing Systems*, pages 576–577. ACM, 2002.

[30] N. Kirchner and A. Alempijevic. A robot centric perspective on the hri paradigm. *Journal of Human-Robot Interaction*, 1(2):135–157, 2013.

[31] K. L. Koay, D. S. Syrdal, M. L. Walters, and K. Dautenhahn. Living with robots: Investigating the habituation effect in participants' preferences during a longitudinal human-robot interaction study. In *Proceedings of the 16th IEEE International Workshop on Robot and Human Interactive Communication.*, pages 564–569, doi:10.1109/ROMAN.2007.4415149. IEEE, 2007.

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

[33] H. Laga and T. Amaoka. Modeling the spatial behavior of virtual agents in groups for non-verbal communication in virtual worlds. In *Proceedings of the 3rd International Universal Communication Symposium*, pages 154–159. ACM, 2009.

[34] K. M. Lee, W. Peng, S.-A. Jin, and C. Yan. Can robots manifest personality?: An empirical test of personality recognition, social responses, and social presence in human–robot interaction. *Journal of communication*, 56(4):754–772, 2006.

[35] F. Lindner and C. Eschenbach. Towards a formalization of social spaces for socially aware robots. pages 283–303, 2011.

[36] K. B. Little. Cultural variations in social schemata. *Journal of Personality and Social Psychology*, 10(1):1–7, 1968.

[37] J. Lofland, D. Snow, L. Anderson, and L. H. Lofland. Analyzing social settings: a guide to qualitative observation and analysis. *Belmont, CA: Wadsworth*, 2006.

[38] M. Lohse, N. van Berkel, E. M. van Dijk, M. P. Joosse, D. E. Karreman, V. Evers, et al. The influence of approach speed and functional noise on users' perception of a robot. In *Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on*, pages 1670–1675. IEEE, 2013.

[39] V. Melnyk. *What is normal to do? : social norms as determinants of consumer decision making*. Phd thesis, University of Wageningen, 2011.

[40] Merrian-Webster.com. ”norm”, 2013.

[41] S. Milgram and E. Van den Haag. Obedience to authority, 1978.

[42] J. Mumm and B. Mutlu. Human-robot proxemics: physical and psychological distancing in human-robot interaction. In *Proceedings of the 6th ACM/IEEE International Conference on Human-Robot Interaction*, pages 331–338. ACM, 2011.

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

[44] S. Y. Okita, V. Ng-Thow-Hing, and R. K. Sarvadevabhatla. Captain may i?: proxemics study examining factors that influence distance between humanoid robots, children, and adults, during human-robot interaction. In *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*, pages 203–204. ACM, 2012.

[45] A. C. on Proficiency Standards for Clinical Pulmonary Function Laboratories et al. Ats statement: guidelines for the six-minute walk test. *American Journal of Respiratory and Critical Care Medicine*, 166(1):111, 2002.

[46] A. K. Pandey and R. Alami. A framework for adapting social conventions in a mobile robot motion in human-centered environment. In *Advanced Robotics, 2009. ICAR 2009. International Conference on*, pages 1–8. IEEE, 2009.

[47] D. A. Prentice. *The psychology of social norms and the promotion of human rights*. Oxford University Press, USA, 2012.

[48] K. Qian, X. Ma, X. Dai, and F. Fang. Socially acceptable pre-collision safety strategies for human-compliant navigation of service robots. *Advanced Robotics*, 24(13):1813–1840, 2010.

[49] B. Reeves and C. Nass. *The Media Equation: How people treat computers, television, and new media like real people and places*. CSLI Publications and Cambridge university press, 1996.

[50] R. N. Rimal and K. Real. Understanding the influence of perceived norms on behaviors. *Communication Theory*, 13(2):184–203, 2003.

[51] A. Sardar, M. P. Joosse, A. Weiss, and V. Evers. Don't stand so close to me: users' attitudinal and behavioral responses to personal space invasion by robots. In *Proceedings of the 7th ACM/IEEE International Conference on Human-Robot Interaction*, pages 229–230, doi:10.1145/2157689.2157769. ACM, 2012.

[52] M. Sherif. *The psychology of social norms.* Harper. Oxford, England, 1936.

[53] M. Shiomi, T. Kanda, H. Ishiguro, and N. Hagita. Interactive humanoid robots for a science museum. In *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*, pages 305–312. ACM, 2006.

[54] M. Shiomi, T. Kanda, H. Ishiguro, and N. Hagita. A larger audience, please!: encouraging people to listen to a guide robot. In *Proceedings of the 5th ACM/IEEE international conference on Human-robot interaction*, pages 31–38. IEEE Press, 2010.

[55] A. L. Strauss. *Qualitative analysis for social scientists*. Cambridge University Press, 1987.

[56] N. M. Sussman and H. M. Rosenfeld. Influence of culture, language, and sex on conversational distance. *Journal of Personality and Social Psychology*, 42(1):66–74, 1982.

[57] D. S. Syrdal, K. Dautenhahn, S. Woods, M. L. Walters, and K. L. Koay. 'doing the right thing wrong'-personality and tolerance to uncomfortable robot approaches. In *Proceedings of the 15th IEEE International Symposium on Robot and Human Interactive Communication.*, pages 183–188, doi:10.1109/ROMAN.2006.314415. IEEE, 2006.

[58] L. Takayama and C. Pantofaru. Influences on proxemic behaviors in human-robot interaction. In *Proceedings of the 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 5495–5502, doi:10.1109/IROS.2009.5354145. IEEE, 2009.

[59] O. L. Taylor. Cross-cultural communication: An essential dimension of effective education. 1990.

[60] H. C. Triandis. *The analysis of subjective culture*. Wiley, New York, 1972.

[61] J. Vroon, M. Joosse, M. Lohse, J. Kolkmeier, J. Kim, K. Truong, G. Englebienne, D. Heylen, and V. Evers. Dynamics of social positioning patterns in group-robot interactions. 2015.

[62] M. L. Walters. *The design space for robot appearance and behaviour for social robot companions. School of Computer Science, Faculty of Engineering and Information Sciences, University of Hertfordshire*. PhD thesis, 2008.

[63] M. L. Walters, K. Dautenhahn, K. L. Koay, C. Kaouri, R. t. Boekhorst, C. Nehaniv, I. Werry, and D. Lee. Close encounters: Spatial distances between people and a robot of mechanistic appearance. In *Humanoid Robots, 2005 5th IEEE-RAS International Conference on*, pages 450–455. IEEE, 2005.

[64] M. L. Walters, K. Dautenhahn, R. Te Boekhorst, K. L. Koay, C. Kaouri, S. Woods, C. Nehaniv, D. Lee, and I. Werry. The influence of subjects' personality traits on personal spatial zones in a human-robot interaction experiment. In *Proceedings of the 14th IEEE International Workshop on Robot and Human Interactive Communication.*, pages 347–352, doi:10.1080/01463373.2011.563444. IEEE, 2005.

[65] M. L. Walters, D. S. Syrdal, K. L. Koay, K. Dautenhahn, and R. Te Boekhorst. Human approach distances to a mechanical-looking robot with different robot voice styles. In *Proceedings of the 17th IEEE International Workshop on Robot and Human Interactive Communication.*, pages 707–712, doi:10.1109/ROMAN.2008.4600750. IEEE, 2008.

[66] L. Wang, P.-L. P. Rau, V. Evers, B. K. Robinson, and P. Hinds. When in rome: the role of culture & context in adherence to robot recommendations. In *Proceedings of the 5th ACM/IEEE international conference on Human-robot interaction*, pages 359–366. IEEE Press, 2010.

Appendix: list of recommendations from Contextual Analysis

This appendix provides a summary of the recommendations we have formulated in the Contextual Analysis report. These recommendations are either implications for the behavior or implications for the perception capabilities of the SPENCER robot.

Section 3.1.2: People searching for the way

- Behavior: a robot which collects passengers should place itself, if possible, near the gate exit but opposite the walking direction (Figure 10) when collecting passengers.
- Perception: ”above average” head turning, especially in combination with a lower walking speed, could be an indication that someone is searching for the way.

Section 3.1.3: People re-packing luggage

- Behavior: when the robot detects passengers being guided are ”repacking luggage”, it should slow down or stop
- Perception: passengers holding their bags in front of them, passengers stopping, kneeling down and searching in their bags

Section 3.1.4: Groups waiting in the hallway

- Behavior: the robot should wait in appropriate places such as the ones shown in Figure 12.
- Perception: In general, locations outside the passenger flow are appropriate to wait.

Section 3.1.5: Individual passengers waiting in the hallway

- Behavior: the robot should wait in appropriate places such as the ones shown in Figure 12.
- Perception: not every individual who is standing still is lost, they might just be waiting for someone/something; people might roam around without goal while waiting

Section 3.1.7: Groups organize themselves in pairs and suddenly start walking behind each other

- Perception: groups of passengers will most likely organize themselves in pairs when following the robot, elderly people might walk one behind the other

Section 3.1.8: Passengers explain the way to others

- Perception: people from different groups can be observed explaining things to each other by gesturing (see also ”encounters with staff”)

Section 3.1.9: People are running

- Behavior: if the robot walks fast (runs) it hast to slow down when encountering crowded areas

- Perception: passengers belonging to one group while running might be one behind the other, the distance between these passengers might increase considerably

Section 3.1.10: Encounters with staff

- Behavior: having finished the conversation about where to go, the robot should indicate the direction with a gesture
- Perception: the robot should be enabled to read boarding cards as they seem to be a common way to share information

Section 3.1.11: Passengers overtaking

- Behavior: if the robot needs to overtake people, the left side in many situations might be the better side to do so

Section 3.1.12: Children running around

- Behavior: the robot might have to slow down or stop when detecting a running child
- Perception: it might be beneficial to be able to detect running children

Section 3.2: Self-service transfer machines

- Behavior: the robot should indicate the direction at the end of the interaction
- Perception: when the robot acts as SSTM, it might be useful to recognize from the positions of the users if a staff member is present who takes over some of the tasks for the robot

Section 3.3: Information monitors

- Behavior: it appears that it might be inappropriate for the robot to drive between information monitors and people looking at them
- Perception: the robot might have to be able to recognize people looking at information screens (groups facing the same direction, standing in half-circles or multiple of these behind each other)

Section 3.4: Encounters with vehicles

- Behavior: signal to other passengers in a polite way that the robot is approaching; give priority to passengers, e.g. by slowing down or adapting walking direction
- Perception: the robot should detect if its path will collide with other passengers' paths

Section 3.5: Moving walkway

- Behavior: the robot might have to adapt its speed to people walking on the mobile walkways
- Perception: the robot might have to track passengers while they are walking on the moving walkway, are overtaking, being overtaken, or blocked by other passengers

Appendix: papers

The remainder of this deliverable contains the papers written as part of this deliverable.

- **Functional noise study** (Section C.1)
  Joosse, M.P., Lohse, M., & Evers, V. (2014) Sound over matter: the effects of functional noise, robot size and approach velocity in human-robot encounters. *Proceedings of the 2014 ACM/IEEE Conference on Human-Robot Interaction*, pp. 184-185

- **Culture spacing survey** (Section C.2)
  Joosse, M.P., Poppe, R.W., Lohse, M., & Evers, V. (2014) Cultural Differences in how an Engagement-Seeking Robot should Approach a Group of People. *Proceedings of the 5th ACM Conference on Collaboration Across Boundaries: Culture, Distance & Technology (CABS)*, pp. 121-130

- **Contextual analysis** (Section D.1)
  Joosse, M.P., Lohse, M., & Evers, V. (2015) How a Guide Robot Should Behave at an Airport - Insights Based on Observing Passengers. *Technical Report TR-CTIT-15-01, Centre for Telematics and Information Technology, University of Twente, Enschede*. ISSN 1381-3625.

- **Telepresence murder study** (Section D.2)
  Vroon, J.H., Joosse, M.P., Lohse, M., Kolkmeier, J., Kim, J., Truong, K.P., Englebienne, G., Heylen, D.K.J. & Evers, V. (2015) Dynamics of Social Positioning Patterns in Group-Robot Interactions. *Proceedings of the 24th IEEE International Symposium on Robot and Human Interactive Communication (ROMAN)*

- **Robot appearance study** (Section D.3)
  Joosse, M.P., Knuppe, R.A., Pingen, G.L.J., Varkevisser, R.A., Vukoja, J., Lohse, M. & Evers, V. (2015) Robots Guiding Small Groups: The Effect of Appearance Change on the User Experience. *Proceedings of the 4th International Symposium on New Frontiers in Human-Robot Interaction*, 21 April 2015.

- **Miscellaneous**
  Joosse, M.P., Lohse, M., & Evers, V. (2013) Short Duration Robot Interaction at an Airport: Challenges from a Socio-Psychological Point of View. *Proceedings of the ICSR 2013 Workshop Robots in public spaces: towards multi-party, short-term, dynamic human-robot interaction*, University of Edinburgh.

  Joosse, M.P., Lohse, M., & Evers, V. (2014) Lost in Proxemics: Spatial Behavior for Cross-Cultural HRI. *Proceedings of the 2014 ACM/IEEE Conference on Human-Robot Interaction Workshop in Culture Aware Robotics*, 3 March 2014.

Sound over Matter: The Effects of Functional Noise, Robot Size and Approach Velocity in Human-Robot Encounters

Michiel Joosse
Human Media Interaction
University of Twente
Enschede, The Netherlands
m.p.joosse@utwente.nl

Manja Lohse
Human Media Interaction
University of Twente
Enschede, The Netherlands
m.lohse@utwente.nl

Vanessa Evers
Human Media Interaction
University of Twente
Enschede, The Netherlands
v.evers@utwente.nl

ABSTRACT

In our previous work we introduced functional noise as a modality for robots to communicate intent [6]. In this follow-up experiment, we replicated the first study with a robot which was taller in order to find out if the same results would apply to a tall vs. a short robot. Our results show a similar trend: a robot using functional noise is perceived more positively compared with a robot that does not.

Categories and Subject Descriptors

H.m [**Information systems**]: Miscellaneous

General Terms

Experimentation

Keywords

Social robot, functional noise, robot height, approach experiment, artificial noise

1. INTRODUCTION AND BACKGROUND

The first impression counts [1], and has already formed when approaching someone. Research has shown that when approaching each other, people exchange social signals using non-verbal communication [7]. Also their appearance is a signal that provides information to the other person [5]. While robot designers can control some of these latter signals by ways of morphological design, robots fall short in employing subtle (non-verbal) signals, such as short glances or gestures, due to technical limitations. Thus, we - as interaction designers - have to find ways to compensate for this lack to ensure that users understand and can predict the robot's behaviors.

Therefore we propose to add functional noise to robots to convey their intentions. Functional noise is added artificial noise to inform people. For instance to an electric car some engine noise may be artificially added so that people can hear it coming. We carried out a first study [6] in which we investigated the effect of fictional noise that communicates how fast the robot is going. This experiment provided us with results of how functional noise and approach velocity influence people's perception of robots.

In the current study we aim to address the biggest limitation in the first study: the height of the robot was only 78 cm [6]. This may explain the lack of effect of approach patterns on users' attitudes and behavior. Previous work on height in HRI found no significant result (120 cm vs. 140 cm) [8]. Or, when differences were found, both height and appearance were manipulated [3]. In order to address this limitation, we conducted an experiment in which we replicated the previous experiment [6] with a taller robot, having a height of 163 cm instead of 78 cm.

2. METHOD

We conducted a 2x2 between-groups experiment, manipulating two independent variables: robot (acceleration and deceleration) velocity and functional noise, see also Figure 1. Our hypothesis is that "*a robot using functional noise to convey its intention to the user will be more positively perceived than a robot which does not use intentional functional noise"*. A 163cm Giraff robot was used. On the screen of the robot, we displayed a pair of eyes, made up from static colored dots. The robot was programmed to accelerate and decelerate either slowly over time (0.1 m/s^2) and to drive "smoothly" or to accelerate and decelerate as fast as possible (1.35 m/s^2) and to drive in an "abrupt" way. The maximum speed of the robot was set to 0.69 m/s, and the robot would approach the participant by driving 4.9 meters in a straight line.

We created two different functional noises; a noise with "*constant noise* level" and a noise that increased in volume at the beginning of the approach and decreased in volume at the end, the latter called "*intentional noise"*. The manipulations resulted in four different experimental conditions.

A 32-item post-experiment questionnaire was used as dependent variable, measuring among others helpfulness (see [6]) and the Godspeed scales [2]. All five Godspeed scales had medium to high internal reliability. The Godspeed scales anthropomorphism (α=.740), animacy (α=.656), likeability (α=.898), perceived intelligence (α=.804) and perceived safety (α=.778).

The sample consisted of 40 participants (25 males, 15 females) with a mean age of 21.25 years (sd=2.30). Participants were equally distributed over the experiment conditions. The participants, mainly students, were recruited from the premises of the University of Twente. After being provided with a short explanation about the experiment, participants filled out a consent form. The robot approached the participants once, after which they filled out the post-experiment questionnaire.

HRI'14, March 3–6 2014, Bielefeld, Germany
ACM 978-1-4503-2658-2/14/03.
http://dx.doi.org/10.1145/2559636.2559822

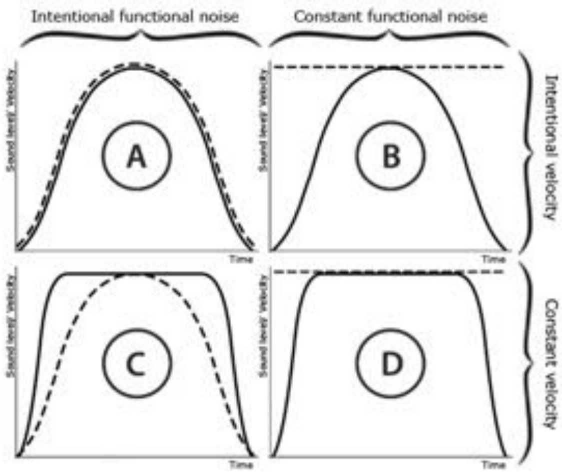


Figure 1. A 2x2 between-groups experiment was conducted, manipulating functional noise and velocity.

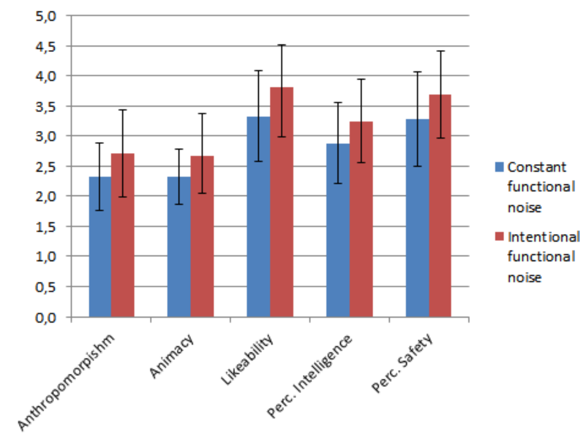


Figure 2. Mean ratings for the combined sample Godspeed scales.

3. RESULTS AND DISCUSSION

Participants liked the robot more in the functional noise conditions, instead of a constant noise conditions, $F(1,39)=3.844$, $p<0.05$. A main effect was found for functional noise on perceived helpfulness: participants rated the functional noise conditions (M=3.35, sd=1.089), as being significantly more helpful than the constant noise conditions (M=2.70, sd=1.081), $U=135.5$, $p<.05$.

When we combined this dataset with the one in [6], we found a significant main effect of functional noise on helpfulness. Participants found an intentional noise pattern (M=3.35, sd=1.122) significantly more helpful a constant functional noise pattern (M=2.73, sd=.987), $U=546.00$, $Z=-2.546$, $p<0.05$. Furthermore, we found significant (2-tailed) main effects for functional noise on all Godspeed scales: anthropomorphism ($F(1,73)=7.685$, $p<0.01$), animacy ($F(1,75)=7.474$, $p<0.01$), likeability ($F(1,75)=9.336$, $p<0.01$), perceived intelligence ($U=520.00$, $Z=0.10$, $p<0.01$) and perceived safety ($U=607.50$, $Z=0.059$, $p<0.05$). For the above scales the intentional noise conditions were rated more positively than the constant noise conditions as can be seen in Figure 2.

No significant effects were found between size of the robots. Both short and tall robots were simple-looking robotic devices without moveable arms. It could be that a robot with a more anthropomorphic, or sophisticated shape, yields different results.

We are aware that we have introduced limitations towards the validity of our work. Previous work in HRI has found that full-frontal robot approaches are not necessarily the most comfortable. The experiment procedure perhaps made participants unnaturally well aware of the approaching robot; participants were focused on the robot from start to finish.

In conclusion, we found that a robot approaching with intentional noise (increasing in volume when the robot accelerated and decreasing in volume when the robot decelerated) was perceived more helpful, and was regarded more positively. Our study shows that functional noise could be a powerful tool to convey a robot's intentions when approaching a user.

4. ACKNOWLEDGMENTS

This work has partly been supported by the European Commission under contract number FP7-ICT-600877 (SPENCER).

5. REFERENCES

[1] Ambady, N.E., and Skowronski, J.J. 2008. *First impressions.* Guilford Publications.

[2] Bartneck, C., Kulić, D., Croft, E. and Zoghbi S. 2009. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International Journal of Social Robotics* 1, 1 (Jan. 2009), 71-81. DOI= http://dx.doi.org/10.1007/s12369-008-0001-3

[3] Butler, J.T. and Agah, A. 2001. Psychological effects of behavior patterns of a mobile personal robot. *Autonomous Robots*, 10, 2, 185-202.

[4] Fong, T., Nourbakhsh I., and Dautenhahn, K. 2002. *A survey of socially interactive robots: Concepts, design and applications.* Technical report CMU-RI-TR-02-29.

[5] Hegel, F., Gieselmann, S., Peters, A., Holthaus, P., and Wrede, B. 2011. Towards a typology of meaningful signals and cues in social robotics. In *Proc. of the RO-MAN'11* (Atlanta, Georgia, July 31-August 3, 2011). RO-MAN 2011. IEEE, 72-78.DOI = http://dx.doi.org/10.1109/ROMAN.2011.6005246

[6] Lohse, M., Van Berkel, N., Van Dijk, E.M.A.G., Joosse, M. P., Karreman, D.E., and Evers, V. 2013. The influence of approach speed and functional noise on the perception of a robot. In *Proc. IROS'13* (Tokyo, Japan, Nov. 3-8, 2013). IROS2013. IEEE, 1670-1675.

[7] Picard, R.W. 2000. *Affective computing*. MIT press.

[8] Walters, M.L. 2008. *The design space for robot appearance and behaviour for social robot companions*. PhD thesis, School of Computer Science, Faculty of Engineering and Information Sciences, University of Hertfordshire.

Cultural Differences in how an Engagement-Seeking Robot should Approach a Group of People

Michiel Joosse
Human Media Interaction
University of Twente
Enschede, the Netherlands
m.p.joosse@utwente.nl

Ronald Poppe
Information and Computing Sciences
University of Utrecht
Utrecht, the Netherlands
r.w.poppe@uu.nl

Manja Lohse
Human Media Interaction
University of Twente
Enschede, the Netherlands
m.lohse@utwente.nl

Vanessa Evers
Human Media Interaction
University of Twente
Enschede, the Netherlands
v.evers@utwente.nl

ABSTRACT
In our daily life everything and everyone occupies an amount of space, simply by "being there". Edward Hall coined the term proxemics for the studies of man's use of this space. This paper presents a study on proxemics in Human-Robot Interaction and particularly on robot's approaching groups of people. As social psychology research found proxemics to be culturally dependent, we focus on the question of the appropriateness of the robot's approach behavior in different cultures. We present an online survey (N=181) that was distributed in three countries; China, the U.S. and Argentina. Our results show that participants prefer a robot that stays out of people's intimate space zone just like a human would be expected to do. With respect to cultural differences, Chinese participants showed high-contact responses and believed closer approaches were appropriate compared to their U.S. counterparts. Argentinian participants more closely resembled the ratings of the U.S. participants.

Author Keywords
human-robot interaction, cross-cultural survey, proximity, social robotics, social interaction, online survey.

ACM Classification Keywords
J.4 [**Computer Applications**]: Social and Behavioral Sciences

INTRODUCTION
In our daily life everything and everyone occupies an amount of space, simply by "being there". When moving through around, people keep a certain distance between each other, and this distance depends on factors like culture, familiarity and personality, as well as the context of the situation.

ACM 978-1-4503-2557-8/14/08…$15.00.
http://dx.doi.org/10.1145/2631488.2631499

In 1966 Hall coined the term proxemics to describe this phenomenon. According to Hall [5], one's body is surrounded by ellipse-shaped bubbles. Each of these bubbles is appropriate for different social interactions. One of these zones, the personal space zone, acts as a virtual buffer zone around our body. Hall describes it as *"a small protective sphere or bubble that an organism maintains between itself and others"*. When this buffer zone is invaded, people compensate for this intimate contact, by non-verbal or verbal compensation behaviors such as stepping away, or limiting eye contact [14]. While every human adheres to others' personal space, what individuals regard as appropriate distances in certain social situations depends on culture [e.g., [19], [7], [17]).

Individual people keep certain distances towards each other, but small groups of people also organize themselves spatially in patterns; such as circles or lines. When such a pattern is stable, it is called a formation. Kendon [10] introduced the term *F*-formation to refer to a specific formation which occurs "*whenever two or more people sustain a spatial and orientational relationship in which the space between them is one to which they have equal, direct and exclusive access"*.

Our work focuses on the spatial organization of small groups in Human-Robot Interaction (HRI). Previous research has provided support for the Media Equation theory, which holds that people treat computers and other media as if they were either real people or real places [15]. A most relevant example is a study by Hüttenrauch et al. [8], which found that most people place themselves in Hall's personal zone (between 0.45 and 1.2 meters distance) when interacting with a robot.

While research in HRI has focused to some extent on the concept of proxemics, this research has been limited in that it has mostly studied robots approaching single persons – usually from Western countries - in controlled lab settings. We intend to extend this state of the art by looking at small groups of people from different cultures. Specifically, we try to identify optimal approach and placement position for a robot which is seeking to gain the attention of a small group of people. As social robots are envisioned to operate in contexts in which they have to interact with people having different cultural backgrounds (such as airports and fairs), we are particularly interested in finding out if a robot requires different spatial behavior depending upon the cultural background of its users. To do so, we have conducted an online survey which we distributed to three different cultural regions in the world through a crowdsourcing platform. In this paper we report on the methodology we used and we provide first results.

RELATED WORK

This section reviews the two major themes of our work: cross-cultural proxemics and group formations. We will conclude this section with our hypotheses, which provide the basis for the experimental method.

Proxemics and culture

In his book, *The Hidden Dimension*, Hall [5] defined four interpersonal distance zones. These zones are called the *intimate, personal, social* and *public* space zones (Table 1).

As stated in the introduction, research has found that the proxemics zones depend on multiple factors, among which culture. Based upon observations, Hall noted that people from low-contact cultures maintain a larger personal space compared with their counterparts from high-contact cultures. Northern European cultures are considered being low-contact, whereas Southern European, Southern American and Arab [4, 5, 19] cultures on the other hand are considered high-contact cultures.

Little [12] used the placement of dolls to infer at which distance people from either the U.S., Sweden, Scotland, Italy and Greece would place people in 19 different social situations. He found that people from North European cultures placed dolls significantly further apart compared with their Mediterranean counterparts. This could be explained by Hall's explanation of high contact- and low contact cultures.

Sussman & Rosenfeld [19] conducted a study in which 105 students from three different countries (Japan, U.S. and Venezuela) had a five-minute conversation with a same sex, same-nationality confederate. They found that, when they were speaking English, participants from the low-contact culture (Japan) sat further apart from each other compared to participants from a high-contact culture (Venezuelan). Within their respective cultural groups, male participants sat further apart than female participants.

Zone	Range	Situation
Intimate	0-0.45m	Lover or close friend
Personal	0.45-1.2m	Conversation between friends
Social	1.2-3.6m	Conversation
Public	3.6m+	Public speech

Table 1: Proxemics zones as defined by Hall [5].

Furthermore, when speaking in their native language, participants from high-contact culture sat closer together than when speaking English. This research implies that human personal spaces zones are dependent on peoples' cultural background.

Also in the field of Human-Robot Interaction (HRI) some studies have been conducted in the area of proxemic zones. Research on proxemics found that people appear to "respect" a robot's personal space zone [8, 23] and maintain a distance from a robot that would be considered respectful when approaching a fellow human. When a robot approaches a person, the comfortable approach distance has been found to be roughly 57 cm [22], which is comparable with distances between people when they have a conversation (see Table 1). Furthermore, similar to human encounters behaviors such as a robot's gaze can influence the distance people chose [21]. If the robot is gazing at people, they tend to stay further away. Work on proxemics in HRI also found that people show similar compensating behavior as they would do when a person invades their personal space [18]. While these findings provide important insights for robot behavior design, HRI research has not yet taken the impact of users' culture into account for proxemics research. As culture is an important factor in human spatial interaction, our work centers around this factor.

F-formations

People organize themselves spatially not only by interpersonal distance, but also in terms of their spatial arrangement when being part of small groups. Kendon [10] introduced the concept of *F*-Formations to capture this phenomenon. According to Kendon, activity is always located in a space. This space can be called the 'transaction segment'.

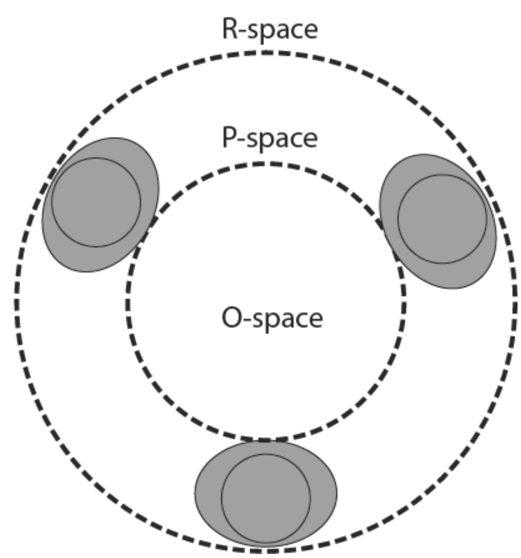


Figure 1 circular *F*-formation around an O-space.

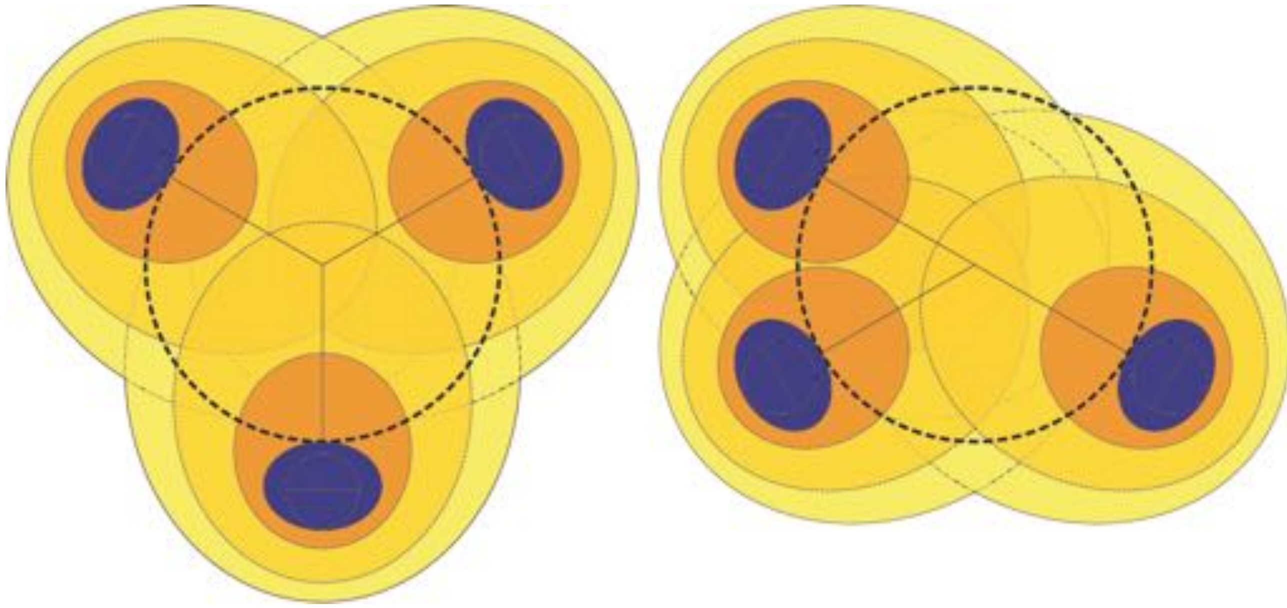


Figure 2: Circular *F*-Formation with congruent (left) and incongruent (right) angles.

When two or more people form a group, they arrange the spatial formation of the group in such a way that the individual transaction segments overlap; thus creating a joint transactional space, also called the o-space (Figure 1). Whenever two or more people establish an o-space, an *F*-formation exists. The o-space is enclosed by the p-space, in which the persons making up the formation are located. The r-space is the space located beyond the p-space. Kendon [10] describes this latter space as *"under the influence of the F-formation […]"*, and provides as example that when multiple *F*-formations occur in a space without physical barriers, these formations tend to be spread out over the space.

F-formations can assume different spatial arrangements. For instance, a circle such as in Figure 2 but also other formations such as a side-by-side or vis-à-vis arrangement. The type of arrangement depends on a number of factors, for instance the number of participants and the context in which the arrangement occurs [10].

Rehm et al. [16] report the "six most occurring formations", and divide these six in open and closed formations. People in open formations are said to allow others to join the conversation; while this is not the case with closed formations. In an experiment with virtual characters, Rehm et al. [16] found that participants were more likely to join an open formation (84% of the trials) than a closed formation. All participants positioned themselves at a social distance, half in the close-social, and half in the far-social distance. However, the authors found that two Arabic participants positioned themselves in the close-social space, which is consistent with findings in cross-cultural research in that Arabic people generally stand closer to each other.

The role different people take on in the *F*-formation could be related to their spatial position. For instance, Kendon [9] observed that speaking rights are reflected in the formation. In a circular formation, rights tend to be equal, in other formations such as a rectangular formation the one in the spatially differentiated position (i.e. the one person sitting opposite others) has the right to speak more compared with others [9].

The arrangement of an *F*-formation can change depending on numerous factors. According to Kendon [10, pg. 221] an L-shape arrangement can for instance become a side-by-side arrangement when the participants focus their attention at an event in the vicinity instead of each other. A participant joining or leaving the specific *F*-formation can also result in a change as the group maneuvers' to maintain the *F*-formation. Thus: *F*-formations can be highly dynamic.

In the field of HRI, research has been conducted investigating the use of *F*-formations in modelling a robot's position. Yamaoka et al. [25] developed a model in which the o-space was established between a robot, listener and an object. The position based upon the developed model was preferred over positions in which the robot was placed either close to the object or to the listener. Kuzuoka et al. [11] investigated the capability of an information-providing robot to change the *F*-formation of a group of listeners. The underlying premise is that a robot which can change the *F*-formation can thereby direct the attention of its listeners. It was found that a robot could achieve this most effectively by rotating it's whole body. While these results are really important for robot design, in HRI, the role of culture with respect to a robot's most optimal position within the *F*-formation has not yet been taken into account.

Personal space and *F*-formations in HRI - Hypotheses

Work on personal space zones has mostly focused on the personal space of single people, and while numerous works call these zones "elliptical", only one distance is reported, which is the distance to the front of the participant. The diameter of the different zones can be estimated, but has not been researched extensively up till now.

Figure 2 contains two different *F*-formations: a circular formation with congruent angles between participants, and a more open formation with incongruent angles. There are three figures along a circle with a diameter of 122 cm (or 4 feet). The circles around the participants represent our hypothesized proxemics zones, these being the intimate zone, close personal and far personal space zone, respectively. The initial position where an actor places him-/herself to join a group can be found more appropriate or inappropriate. We would like to introduce this optimal approach position as a combination of the position an actor chooses with respect to the group members in between which he/she approaches, and the distance he/she takes from those actors.

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

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

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

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

Given that different cultures exist, and that research by Rehm et al. [16] found that participants from high-contact cultures stand closer to a group of people compared with people from low-contact cultures, we hypothesize a similar cultural dependent preference will exist when a robot approaches.

> H3: Participants from a high-contact culture (China, Argentina) are more comfortable with a closer approach by a robot than participants from a low-contact culture (U.S.).

METHODOLOGY

We conducted a 3 (nationality) x 3 (position in the group) x 6 (distance from the group) online study. A survey-based questionnaire was distributed through a crowdsourcing platform (crowdflower.com) to a targeted population. Participants were shown images of small families of 3D people and a robot (see Figure 3). These groups were composed of three people: a man, a woman and a child. The survey consisted of an introduction that contained detailed instructions as well as a picture of the family (Figure 4). Participants were asked to indicate how appropriate they believed the position of the robot was, imagining that the position was the position after the robot had completed its approach. The position of the robot was manipulated two-fold within-subjects (see next section), the nationality of the participants was a between-subjects variable. A questionnaire was used to measure the dependent variables.

For the groups, a circular formation with congruent angles was chosen. We are aware of the fact that people will often stand in non-congruent angle formations, however, if we were to introduce a formation with non-congruent angles and/or people spaced differently we would introduce a multitude of factors that would be hard to control for and that would make the study overly complex.

The diameter of the o-space was set to 122 cm, which corresponds to Hall's social space. The height of the participants was based upon average international height[1]. The male was scaled to 178 cm, the female to 152 cm, and the child to 140 cm. The height of the robot was scaled to 140 cm, as can been seen in Figure 4.

Figure 3 Example top-down still as shown to participants

Figure 4 The fictional family was scaled to average international dimensions based on [1]

[1] http://dined.nl//ergonomics/

Independent variables

Two variables were manipulated within-subjects: approach position (the position between which family members the robot approached, Figure 5), and the approach distance of the robot. We refer to the combinations of position and angle as scenes.

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

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

Dependent variables

The dependent variables were measured using a 112-item online questionnaire, measuring a total of 6 constructs. The questionnaire was divided into three consecutive blocks: appropriateness rating of the robot-group scenes, questions regarding participants' cultural background and personality, and general demographic questions.

In the first block, participants were asked to rate the 36 'robot approaches a family' scenes that have been described in the previous section. To avoid order-effects, the order of all scenes was randomized. Participants were provided with the instruction: "*The robot approached the family and has come to a halt between particular family members at a particular distance. Now it will interact with them*", and asked to indicate on a 7-point Likert scale how appropriate the position of the robot was. Another four items were included in this block to measure how participants themselves would approach the family. Two items were included to check the approach position- and distance manipulation. Here participants were provided with statements such as *"the robot generally approached from the same direction"* and *"the robot generally approached from different directions"*. Participants were forced to choose which of the statements was true. A final item was included in which we asked participants if they could indicate where they thought the family they had seen in the situations originated from.

The second block of the questionnaire consisted of a series of validated scales measuring four dependent variables. An indication of whether participants were members of a high-contact or low-contact culture was assessed by measuring *closeness* as people from a high-contact culture have been found to sit significantly closer to each other compared with members from a low-contact culture [19]. Five items from the IPROX (iconic proximity) questionnaire were used [7]. Participants' general attitude towards robots was measured by the *Negative Attitude Towards Robots* scale, a 14-item 7-point Likert scale. Hofstede [6] identified five dimensions of culture, one of these being Individualism-Collectivism. One way to explain cultural differences is by measuring *individual* and *group self-representations*. Individual self-representations refer to whether the self is represented as "a separate, unique individual" [1] whereas group-self representation refers to one who is "an interchangeable part of a larger social entity" [1]. This was operationalized using 7 items, by Brewer & Chen [1].

The final construct in this block was personality as we figured this could influence people's preference for a robot position (e.g., more extrovert people preferring the robot to come closer or to approach at their side of the group). We measured the Big Five personality traits using the 20-item Mini-IPIP scale [2].

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

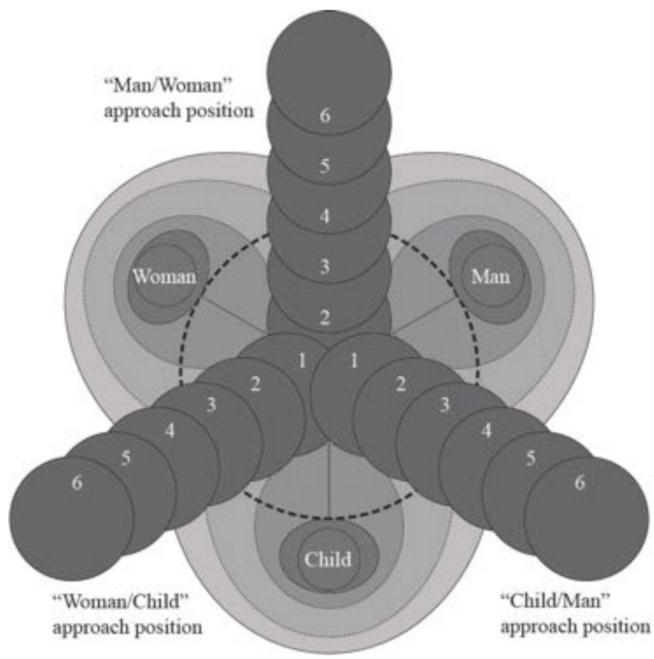


Figure 5: Participants standing in a circular F-Formation with a diameter of 122 cm. Dark grey indicates possible location of the robot. Grey: intimate zone, light grey: personal space zone.

	N	Mean age (sd)	Male / Female
U.S.	86	43.27 (12.25)	26 / 60
China	29	30.48 (8.93)	19 / 10
Argentina	66	33.06 (10.90)	48 / 18
Total	181	37.50 (12.54)	93 / 88

Table 2 Number, nationality, mean age, and gender of the participants

Participants

Participants were recruited from three different countries: China, Argentina and the United States. People from these three countries are generally considered culturally different; not only because they are geographically on different continents, but also because various studies have shown cultural differences [3, 4, 17] in for instance societal values.

For each country, participants were recruited through the Crowdflower platform, which allows for specification of the target country. 244 participants completed the questionnaire; each being paid $1 for completion of the survey. Responses were limited to one per IP address. Participants who failed to correctly answer the two manipulation checks were excluded from the sample. A second control method was the analysis of the robot-scene questions, which were 18 situations rated twice by each participant on a 7-point Likert scale. Participants who rated four or more situations with a difference of 3 or more points were also excluded from the survey. In total 63 participants (26%) were excluded. After applying the exclusion criteria, the total sample contained 181 participants, as specified in Table 2.

Data analysis

The results presented in this paper focus on the ratings of the scenes and on the closeness scale (five items from the IPROX questionnaire, see Dependent variables). Internal reliability of all scales was assessed by calculating Chronbach's α, and deemed acceptable for all scales.

As stated in the previous section, the participants rated all scenes twice as form of control method. After having excluded participants these ratings were averaged per participant and scene.

Approach in between	Mean	SD
Man-Woman	4.11	0.095
Woman-Child	4.16	0.100
Man-Child	3.93	0.093

Table 3 Mean appropriateness scores and standard deviation grouped by approach direction

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

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

RESULTS

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

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

In H1 we hypothesized that participants would rate it as more appropriate if the robot was positioned out of every group member's intimate space zone. A repeated mixed-model ANOVA revealed that participants rated the robot positions in the intimate space zone as significantly less appropriate (M=3.14, sd=1.25) compared with those positions where the robot was positioned outside the intimate space zone (M=4.61, sd=.99), $F(1, 100.658) = 109.567$, $p<0.01$. We therefore **accept H1**: a robot which stays out of the intimate space zone of each of the group members is considered to be more appropriate. These ratings were neither affected by gender (p=.87) nor by nationality (p=.60).

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

A factorial repeated-measures ANOVA with two independent variables (distance and position) and two between-subjects factors (gender and nationality) was conducted. Mauchly's test indicated that the assumption of sphericity had been violated for the main effects of distance, $\chi^2(14) = 613.9$, $p<0.001$, and angle, $\chi^2(2) = 76.37$, $p<0.001$. Sphericity had also been violated for the interaction effect (distance*direction), $\chi^2(54)=183.55$, $p<0.001$. The degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity (ε=.42 and ε=.74 for the main effects, and ε=.81 for the interaction effect).There was a significant main effect of the approach distance ($F(2.09,365.19)=54.37$, $p<0.001$), and a non-significant effect of approach angle on the appropriateness of the robot's position ($F(1.47, 258.25)=2.857$, $p=0.075$).

Post-hoc contrasts revealed a significant difference of appropriateness between the “Woman/Child” and “Man/Child” approaches: the appropriateness of the “Woman/Child” approach was significantly higher compared with the “Man/Child” approaches, $F(1, 175) = 11.71$, $p<0.001$ (See Table 3). The appropriateness of the “Man/Woman” approaches was equally appropriate as the “Woman/Child” approach, We therefore only **partially accept H2**, in which we hypothesized that participants would rate a robot approach as less appropriate if a robot approached in between a child and parent, as compared with approaching in between both parents. Instead, participants indeed found an approach between parent and child less appropriate but only for the position between father and child. The most appropriate approach position was generally thought to be in between the mother and the child (see Table 4).

Influence of cultural background on appropriateness

To check whether the countries that we chose actually differed in the low-high contact dimension, we analyzed the items from the closeness questionnaire. There was a significant difference between the ratings, $F(2) = 15.528$, $p<0.001$. As can be seen in Figure 6, participants from the United States gave significantly higher ratings on the closeness measure (M=4.96, sd=1.05), which indicates they put more distance between themselves and other people. This effect was vice-versa for Chinese people, as expected (M=3.88, sd=1.20). The Argentinian participants rated in between (M=4.11, sd=1.19).Therefore, we can assume that the national groups included in this sample can indeed be considered to have different cultural backgrounds concerning the low-high contact dimension.

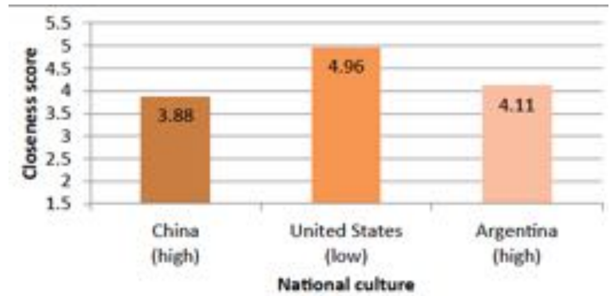


Figure 6 Participants from what are considered low-contact cultures scored indeed significantly higher on the “closeness” construct (scale: 1: high contact, 6: low contact. Mean scores provided in bars).

Our third hypothesis was that participants from high-contact cultures (such as China and Argentina) would rate a close approach as more comfortable than participants from a low-contact culture (United States). There was a significant three-way interaction effect between the nationality of the participant, distance, and position of the robot on appropriateness of the scene, $F(16.20, 1417.24)=1.912$, $p<0.05$. This effect can be seen in Table 4 and Figure 7.

The Table and Figure show that the U.S. and Argentinian participants gave similar appropriateness ratings for the

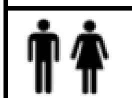
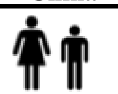
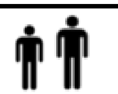
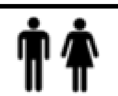
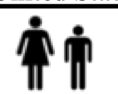
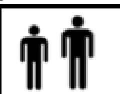
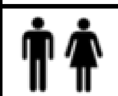
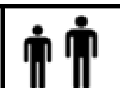
distance	China			United States			Argentina		
									
20	3.052 (1.555)	3.862 (1.870)	3.741 (1.751)	2.721 (1.560)	3.023 (1.627)	2.721 (1.516)	2.697 (1.544)	2.909 (1.446)	2.530 (1.364)
40	3.345 (1.748)	4.017 (1.740)	3.310 (1.785)	3.552 (1.690)	3.843 (1.676)	3.407 (1.474)	3.477 (1.515)	3.614 (1.230)	3.038 (1.178)
60	3.862 (1.737)	4.155 (1.895)	3.759 (1.740)	4.308 (1.478)	4.552 (1.596)	4.128 (1.468)	4.083 (1.583)	4.439 (1.383)	3.689 (1.202)
80	4.259 (1.766)	4.466 (1.732)	4.414 (1.547)	5.047 (1.490)	4.988 (1.538)	4.709 (1.523)	4.795 (1.48)	5.000 (1.547)	4.220 (1.356)
100	4.517 (1.825)	3.672 (1.649)	4.414 (1.753)	5.337 (1.428)	4.977 (1.439)	4.994 (1.610)	5.136 (1.423)	5.220 (1.465)	4.644 (1.315)
120	3.724 (1.893)	3.103 (1.749)	3.603 (1.655)	4.913 (1.722)	4.692 (1.736)	4.837 (1.591)	4.917 (1.690)	4.841 (1.813)	4.689 (1.583)

Table 4 Mean appropriateness ratings for the Chinese, U.S. and Argentinian sample. Distance indicates distance between the center of the circle and the robot (in cm). Mean appropriateness ratings on a scale from 1 to 7, standard deviations between brackets.

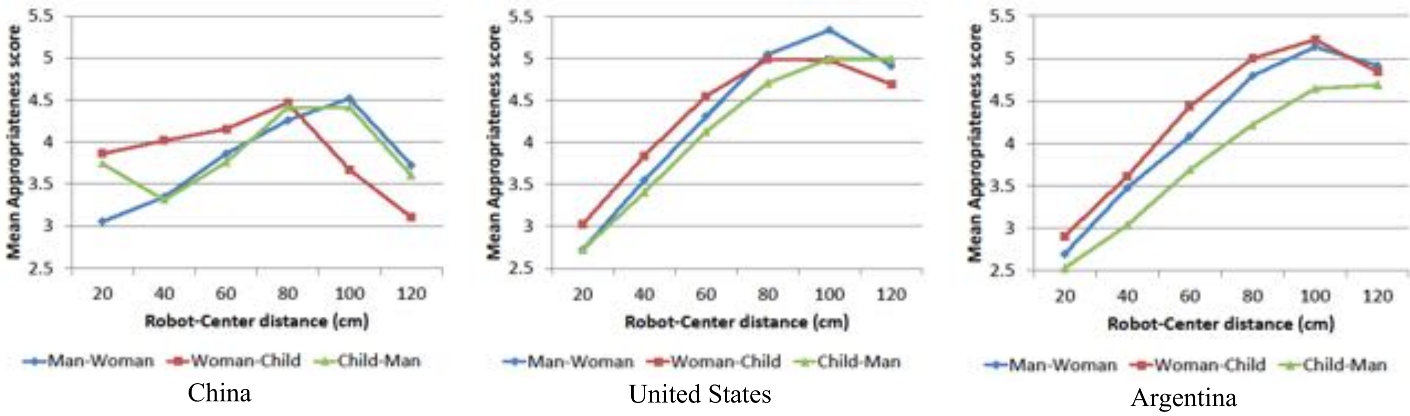


Figure 7 Mean appropriateness ratings for the Chinese, U.S. and Argentinian sample. Appropriateness on a scale from 1 to 7.

approach distances, but that one particular approach in between the "Man/Woman" was considered most appropriate by the U.S. participants, whereas the Argentinian believed the "Woman/Child" position was more appropriate.

The Chinese participants' ratings were generally lower and a notable difference was that the closer approaches (within the intimate zone) were actually considered to be quite appropriate. Like the U.S. and Argentinian participants, the Chinese also had a preference for a further stop distance (80-100 cm), though this difference was much less pronounced.

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

DISCUSSION

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

We hypothesized that participants would find approaches in between the parents more appropriate compared with the approaches where a child is cut off from one of the parents. The reason for the unexpected finding that approaches between mother and child were found quite appropriate could be a pragmatic one, which we had not considered. By approaching in between the mother and child the robot directly faced the father of the family. It could be that a robot's frontal approach to a male is seen as more appropriate. Even though previous work by Walters et al. [24] did not confirm this notion, this warrants further investigation into differences in gender preferences.

Figure 7 shows similarities in the appropriateness ratings of the U.S. and Argentinian samples respectively, despite the fact that Argentinian's closeness scores indicate a higher-contact culture. Therefore, we expected they would find it more appropriate if the robot approached closer. Thus, it could be that the high-low contact culture dimension is too simple and did not completely capture the subtleties of high-low contact cultural backgrounds and that there are more factors at play. One possible explanation can be found in Hofstede's work [6]. On the Individualism dimension, the U.S. scores are high (91 points), and Chines scores are relatively low (20). Argentinians scores are at 46 points. This is still closer to China than the U.S., however, if we look at other Latin American countries, such as Ecuador (8), Venezuela (12), Colombia (13) and Chili (23) it appears to be that Argentina is a rather individualistic country. This might partly explain why Argentinian participants showed a preference for a further positioning of the robot. However, this issue deserves further investigation.

Furthermore, we have not yet analyzed the relation between personality and the appropriateness of robot scenes. Previous work in HRI has shown that a high score on extraversion leads to more tolerance to uncomfortable robot approaches [20]. It could very well be that personality also influences ratings of appropriateness. In a similar way *attitude towards robots* and *individual* and *group self-representations* could influence the results in subtle ways, which we have not yet analyzed.

To analyze cross-cultural differences in proxemics, we used an online questionnaire as this allowed us to distribute the survey to geographically dispersed samples. The survey contained static images, and while the results do support most of our hypotheses, the ecological validity of our research is limited because groups are dynamic entities. The formation of the group changes when a new member joins the group, and our images might very well not have been able to capture these subtle dynamics. In future work, we will conduct a study where actual groups of people are approached by a robot– primarily to see if the results found with this survey are replicable when such an experiment is conducted in a lab or real world setting.

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

Another limitation of the experimental design concerns the chosen *F*-formation. As we explained in the methodology section, we chose for a closed circular formation with congruent angles (Figure 2). It could very well be that another formation, for instance with incongruent angles, yields different results; either because of the position (and status) of the members within the group, or simply because there is more room for a robot to approach when the angles are not congruent. This issue will also be addressed in future research.

Finally, the context of our stimuli could be debatable. The reason for not providing a specific context in which this group and the robot would interact (for instance a domestic environment, airport or shopping mall) was that we did not want our participants to have a predisposed opinion on for instance the feasibility or acceptability of a robot in a certain context. However, how different real-world contexts influence the ratings is a highly interesting question as future robots will be operating in such contexts.

As stated in the introduction, this is a first study. In order to improve ecological validity and generalizability of our results more research has to be conducted. Our future work will focus on replicating a similar experimental setup in either a physical lab or field setting in order to account for some of the limitations that arose in this experiment, as also pointed out in literature (f.e. [10]).

CONCLUSION

In this paper we have presented the first results of a survey that we distributed to three countries (China, the U.S. and Argentina). We were interested in finding out whether or not people from different nationalities have different proxemics expectations from a robot which approaches a small family.

The most appropriate approach distance appears to be somewhere near 80 cm from the center of the circle. Our results also show that while participants found a robot more appropriate when it stayed out of the intimate space zone, there are cultural differences which surface when comparing China with the other two countries:

- Argentinian and U.S. participants rated approaches in Hall's intimate space zone as clearly inappropriate whereas the Chinese participants rated approaches farther away (100-120 cm) as more inappropriate.
- Argentinian and U.S. participants rated an approach in between the child and man as less appropriate, Chinese participants did not have a clear preference.

Unexpectedly, the Argentinian ratings were closer to the U.S. ratings even though both Argentina and China were considered to be high-contact cultures, and both scored as such on our closeness measure. Hence, there seem to be many factors that contribute to the cultural identity of people that we will look into in the future, among others the interplay between personality and culture, as well as to the limitations caused by the methodological choices.

Overall, the influence of culture on HRI has turned out to be a promising research direction with respect to proxemics. Our first research shows that researchers need to take culture into account when building robots that operate in intercultural environments.

ACKNOWLEDGEMENTS

This work has partly been supported by the European Commission under contract number FP7-ICT-600877 (SPENCER).

REFERENCES

1. Brewer, M.B. and Chen, Y.-R. 2007. Where (who) are collectives in collectivism? Toward conceptual clarification of individualism and collectivism. *Psychological review*. 114, 1 (Jan. 2007), 133–51.
2. Donnellan, M.B., Oswald, F.L., Baird, B.M. and Lucas, R.E. 2006. The mini-IPIP scales: tiny-yet-effective measures of the Big Five factors of personality. *Psychological assessment*. 18, 2 (Jun. 2006), 192–203.
3. Gupta, V., Hanges, P. and Dorfman, P. 2002. Cultural clusters: Methodology and findings. *Journal of World Business*. 37, (2002), 11–15.
4. Hall, E.T. 1963. A system for the notation of proxemic behavior. *American anthropologist*. 65, 5 (1963), 1003–1026.
5. Hall, E.T. 1966. *The Hidden Dimension*. Anchor Books.
6. Hofstede, G. 2001. *Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations across Nations*. Sage Publications, Inc.
7. Høgh-Olesen, H. 2008. Human Spatial Behaviour: The Spacing of People, Objects and Animals in Six Cross-Cultural Samples. *Journal of Cognition and Culture*. 8, 3 (Aug. 2008), 245–280.
8. Hüttenrauch, H. and Eklundh, K. 2006. Investigating spatial relationships in human-robot interaction.

Proceedings of the 2006 IEEE Conference on Intelligent Robots and Systems. (2006), 5052-5059.

9. Kendon, A. 1973. The role of visible behaviour in the organization of face-to-face interaction. In M. von Cranach and I. Vine, eds., *Social Communication and Movement Studies of Interaction and Expression in Man ad Chimpanzee*. New York: Academic Press, p. 29-74.
10. Kendon, A. 1990. *Conducting interaction: Patterns of behavior in focused encounters*. Cambridge University Press.
11. Kuzuoka, H., Suzuki, Y., Yamashita, J. and Yamazaki, K. 2010. Reconfiguration Spatial Formation Arrangement by Robot Body Orientation. *Proceedings of the 5th ACM/IEEE international conference on Human-robot interaction,* 285-292
12. Little, K.B. 1968. Cultural variations in social schemata. *Journal of Personality and Social Psychology*. 10, 1 (Sep. 1968), 1–7.
13. Nomura, T., Kanda, T., Suzuki, T. and Kato, K. 2008. Prediction of Human Behavior in Human--Robot Interaction Using Psychological Scales for Anxiety and Negative Attitudes Toward Robots. *IEEE Transactions on Robotics*. 24, 2 (Apr. 2008), 442–451
14. Patterson, M.L., Mullens, S., and Romano, J. 1971. Compensatory Reactions to Spatial Intrusion. *Sociometry*. 34, 1 (1971), 114-121
15. Reeves, B. and Nass, C. 1996. *The Media Equation: How People Treat Computers, Television and New Media Like Real People and Places*. Cambridge University Press.
16. Rehm, M., André, E. and Nischt, M. 2005. Let ' s Come Together — Social Navigation Behaviors of Virtual and Real Humans. *INTETAIN 2005* (2005), 124–133.
17. Remland, M.S., Jones, T.S. and Brinkman, H. 1995. Interpersonal distance, body orientation, and touch: Effects of culture, gender, and age. *The Journal of Economic Perspectives*. 135, 3 (1995), 281–297.
18. Sardar, A.H., Joosse, M.P., Weiss, A. and Evers, V. 2012. Don't stand so close to me: users' attitudinal and behavioral responses to personal space invasion by robots. *Proceedings of the 2012 ACM/IEEE Conference on Human-Robot Interaction*, 229-230.
19. Sussman, N.M. and Rosenfeld, H.M. 1982. Influence of culture, language, and sex on conversational distance. *Journal of Personality and Social Psychology*. 42, 1 (1982), 66–74.
20. Syrdal, D.S., Dautenhahn, K., Woods, S.N., Walters, M.L. & Koay, K.L. 2006. 'Doing the right thing wrong'- Personality and tolerance to uncomfortable robot approaches. *Proc. of the 15th IEEE Symposium on Robot and Human Interactive Communication*, 183-188
21. Takayama, L. and Pantofaru, C. 2009. Influences on proxemic behaviors in human-robot interaction. *Proceedings of the IEEE/RSJ International Conference on Robots and Systems (IROS2009)* (2009), 5495–5502.
22. Walters, M.L. 2008. *The design space for robot appearance and behaviour for social robot companions*. University of Hertfordshire.
23. Walters, M.L. and Dautenhahn, K. 2005. Close encounters: Spatial distances between people and a robot of mechanistic appearance. *Proceedings of the 2005 IEEE-RAS Conference on Humanoid Robots* (2005), 450-455.
24. Walters, M.L., Dautenhahn, K., Woods, S.N. and Koay, K.L. 2007. Robotic Etiquette: Results from User Studies Involving a Fetch and Carry Task. *Proceedings of the 2007 ACM/IEEE Conference on Human-Robot Interaction,* 317-324.
25. Yamaoka, F., Kanda, T., Ishiguro, H. and Hagita, N. 2008. How close?: model of control for information-presenting robots. *Proceedings of the 3rd ACM/IEEE international conference on Human-robot interaction,* 137-144.

How a Guide Robot Should Behave at an Airport - Insights Based on Observing Passengers

Michiel Joosse, Manja Lohse, Vanessa Evers *

31 January 2015

1 Introduction

This report describes a contextual analysis conducted as part of the EU FP7 project SPENCER. This contextual analysis is based on the specifications in SPENCER Deliverable 1.1 (D1.1). In accordance with the use case specification, we specify aspects of the contextual analysis that we conducted at Schiphol airport.

As part of the SPENCER project a robot demonstrator will be developed which provides location based services (information, guiding) to passengers in the context of an international airport. Based upon the use case meeting held 7 April 2013, the following four main tasks for the robot[1] have been defined:

1. Give directions to passengers at gate
2. Guide MCT[2] passengers (non-runners) to the fast-track Schengen filter
3. Guide non-MCT passengers
4. Help passengers who miss their connection by use of a boarding card printer

Our specific work package (WP4) is concerned with the identification and evaluation of social normative motion behaviors for the robot. Given that the robot should act, and be perceived as, a companion, or intelligent agent, insight is needed into which norms, routines and rituals people adhere to when in transit. For the sake of simplicity, we refer to these as norms.

We use the term contextual analysis here to describe a scientific method to discover how people behave in a given context (here Schiphol airport) and in relevant situations within this context. Thus, the data for this contextual analysis has been collected at Schiphol Airport. Generally spoken, the main goal of the contextual analysis is to analyze human behavior at Schiphol airport in order to identify normative behaviors that the SPENCER robot should employ in the same context. The goal is explained in more depth below. The Method section will describe the research design in more detail, and provides information on the specific locations where we collected data. These data include observations from researchers, interviews with KLM staff and video recordings in representative environments. We will conclude with the data analysis in which we specify how we analyzed and structured the video data. In the results section we will describe the results gained from the observations, and finally in the conclusions section we will provide recommendations for the SPENCER robot's behavior (see 5).

* All authors are with the Human Media Interaction group, Department of Electrical Engineering, Mathematics and Computer Science, University of Twente, the Netherlands. Contact information: {m.p.joosse; m.lohse, v.evers }@utwente.nl

[1] Robot, or the robot refers to the demonstrator robot that will be developed.

[2] MCT: Minimum Connection Time; for a transfer between an intercontinental and domestic flight this is 50 minutes or less before next flight departs

Figure 1: The Schengen barrier, seen from the Non-Schengen area. The fast-track lane is at the left side (but not visible)

1.1 Goal

As specified in the description of work, the goal of this contextual analysis is *"to identify [] socially normative motion behaviors that humans adopt when navigating through crowded environments and interacting with groups in real world settings relevant to the project such as airports"*.

In order to attain this goal of the contextual analysis we have specified a set of contexts in which people navigate at the airside (behind customs) of an airport (see Method section). This data has been analyzed and transformed into a set of behavior recommendations for the SPENCER robot.

2 Method

To achieve our goal, we first specified situations and places that were of interest for the scenario (see Section 2.1). In these places we collected data as described in Section 2.2 and analysed it as explained in Section 2.3.

2.1 Situations and places of interest

In the current scenario, passengers disembark an aircraft, and walk to their next destination, which can either be the exit if Schiphol Airport is their destination or another aircraft if Schiphol serves as a connecting hub. Thus, the arrival gate and the ways taken to the next destination are of interest for us. As we mainly focus on passengers with connecting flights, we chose to observe people who walked to the Schengen barrier rather than the ones walking towards the exit.

MCT for international flights is 50 minutes. On average 80% of the passengers from a Chinese flight are transferring to a connecting continental flight, and have to pass the Schengen barrier (Figure 1) to enter the Non-Schengen area. The Non-Schengen area is a clean area, e.g. your hand luggage is checked before entering the area. Passengers first have to pass a passport control, followed by a security check. There is a fast-track lane, and a staff member is always present to direct passengers to which lane they should take. Only passengers whose flight will department within 20 minutes are allowed to use the fast-track lane.

Along the way, connecting passengers are provided with flight information by way of the information monitors positioned at various locations throughout the terminals. Passengers can also use the Self-Service Transfer Machines, which are placed throughout the terminal building. With these machines passengers can (among others) get travel information and rebook their flight without the assistance of staff. Given that the SPENCER robot is envisioned to assist passengers with these functions, these machines are of interest to us. The self-service transfer machines (SSTM's) are placed at various locations, some are placed in the middle of the gate (Figure 2), others at strategic points called "transfer areas" (Figure 3). These transfer areas are numbered, and positioned throughout the airport. Transfer 6 (Figure 3, Figure 4) is the largest transfer desk area in the Non-Schengen area, Transfer 2 in the Schengen area.

Figure 2: Four SSTM's located in the hallway of the E-gates; near gate E9

Figure 3: Transfer 6 self-service area

SSTM's support a variety of languages, including Chinese. It is possible for passengers to rebook flights, check flight information among others. At times the transaction cannot be completed at the SSTM, and the passenger will receive instructions to ”report to the transfer desk”. Nancy, a member of KLM's ground staff at Schiphol told us that older people in particular are more comfortable with a person helping them; they prefer a person to the machine. She told us that this could be because they are not used to work with computers.

The SPENCER robot will not be the first driving vehicle at Schiphol. For passengers with disabled or reduced mobility, Schiphol provides assistance in the form of staff-operated vehicles to bring passengers from the arrival gate to departure gate. We decided to observe how these passengers moved around. With the knowledge above we defined the six situations described in Table 1 of which we think information could benefit the development of the SPENCER robot.

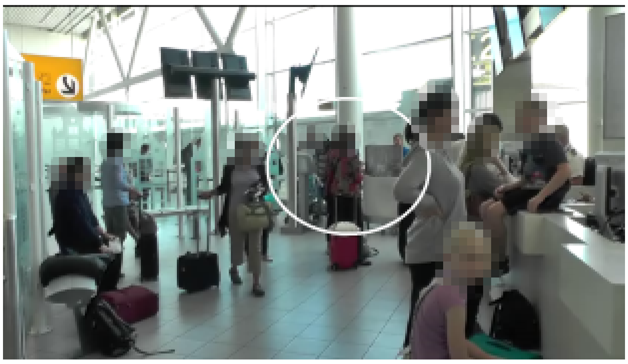


Figure 4: Transfer desk of transfer 6. Note that the ticket officer (seated in the back) first screens the needs of passengers

	Description of situations	Questions of interest	Observation method	Place(s) for data recording
1	passengers walking around without guidance	How does the group make decisions? Is there a leader? How do group members position themselves with respect to each other? (while walking or waiting)	non- participatory, hidden	on the way between gate and Schengen filter
2	groups approaching staff	How do people approach staff? Is there a leader in the group? Is the group leader in a specific position?	non-participatory, hidden	a) gate; b) on the way between gate and Schengen filter
3	staff approaching groups	How do staff approach groups? How are people instructed to use the fast-track?	participatory, hidden	a) gate; b) on the way between gate and Schengen filter
4	information screen	How do people position themselves around information screens? Do they stop to read the information?	non-participatory	a) gate; b) on the way between gate and Schengen filter
5	passengers at self-service transfer desks	How do people approach? How do they position themselves?	non-participatory, hidden	self-service transfer desk
6	airport vehicles	How do the airport vehicles interfere with other passengers?	non-participatory, hidden	b) on the way between gate and Schengen filter

Table 1: Situations to be observed at Schiphol Airport

The observations as described in Table 1 are non-participatory. In a participatory observation a person who knows about the study interacts with the participants (as opposed to a non-participatory observation). As the situations should be influenced as little as possible, non-participatory observations were preferred. Furthermore, observations can be open or hidden. In an open observation the participants know that and why they are observed. In a hidden observation, people do not know about it. Again, as the goal was to interfere as little as possible in the situation, hidden observations were preferred.

2.2 Data collection

Data was collected by the SPENCER consortium on 11 and 12 June 2014. Both days were average days in terms of traffic at Schiphol airport; it was not yet extremely crowded due to the holiday season. During our observations at the airside we were guided by two KLM staff members from the Translator Team (Dicky and Nancy). Their job is to assist passengers from China (a language members of this team speak fluently) when they arrive at Schiphol Airport. The UT group was guided around by Nancy, and before providing an overview of the data collected we will briefly describe their job.

Nancy has worked for seven years at Schiphol, before that she worked as a flight attendant for 4 years. The job of Nancy and Dicky is to be present at the arrival gate when a KLM flight from Japan, Korea or China arrives and assist in directing them to the right gate. Usually this does not entail physically walking with them, just giving directions, or sometimes helping them use the self-service transfer kiosk. The main reason is the language. More and more (young) people are starting to speak English, however, a lot of passengers do not. This is offered as a service by KLM. There is no staff at the airport which specifically roams around the airport looking for people in need of help. As we will also discuss

Figure 5: Layout of observerations around lounge 2

in Section 3.1.10 passengers ask help to staff who walk from A to B at the airport. Nancy explained that she rarely approaches people to assist them. They always come up to her (and her colleagues) because she's wearing a KLM uniform. What Nancy also mentioned was that even when she wasn't wearing the uniform, people did still tend to ask her for assistance.

2.3 Description of collected data

Data was recorded at various locations in the Non-Schengen part of the terminal - in particular lounges 2 and 3, and gates D, E, and F. Figure 5 and Table 2 provide further details of these observations. These data cover the situations as specified in Table 1. One situation we were also interested in was a situation in which participants would be approached by staff. However, as Nancy indicated there is no specific staff for this on Schiphol. During our time at Schiphol Airport we did notice some spontaneous staff-passenger interactions in the hallway. These interactions have been analyzed in the results section. However, for the purpose of our analysis these interactions were limited in that they were between staff and single passengers; not between staff and groups of participants as would be more relevant for the SPENCER project.

Video recordings were made from passengers exiting aircrafts, either assisted or not assisted by KLM ground staff. By following the general flow of passengers, we collected data on passengers using the SSTM's and information monitors.

The flights that we observed originated from destinations in Asia (Tokyo, Beijing), but also the United States (San Francisco) and Canada (Vancouver). One flight originated from the United Kingdom, which is a European country, though not part of the Schengen zone. Therefore passengers also disembark at the Non-Schengen gates. From the flight we followed two passengers from gate until they went to the lounges area. This proved to be difficult as it was obvious to the passengers that we were following them, and this also attracted the attention of other passengers, thus intruding in the context. Therefore, it was decided to not repeat this with more passengers. The data gathered elsewhere does provide sufficient information to arrive at conclusions concerning situation #2.

In two situations we collected data in the main terminal building for an extended amount of time. In the first situation, two cameras were placed in Lounge 2 for 30 minutes. The specific area is shown in Figure 5 and Figure 7, and as can be seen this area leads from one of the passengers' entrances to the Non-Schengen area, as well as the Schengen barrier, and two shops. The area also contains two series of information monitors. The second situation includes a camera placed for 15 minutes at the junction between the D-gates and lounge 2.

The total duration of the video data collected during this data collection session was 4.5 hours.

Figure 6: Data in lounge 2 was captured with 2 cameras

	Day	Time	Observation:	Gate	Type of observation	# of cameras
1	Wednesday	10.30	Arrival of flight from Vancouver (KL 0682)	E24	hidden observation of arriving passengers	1
2	Wednesday	15.15	Arrival of flight from Beijing (KL 0898)	E20	hidden observation of arriving passengers	2
3	Wednesday	15.00	Arrival of flight from Tokyo (KL 0862)	E19	hidden observation of arriving passengers	2
4	Thursday	9.50	Arrival of flight from San Francisco (KL 0606)	F8	hidden observation of arriving passengers	2
5	Thursday	10.10	Arrival of flight from Vancouver (KL 0682)	F4	hidden observation of arriving passengers	2
6	Thursday	13.35	Arrival of flight from Aberdeen (KL 1444)	D8	Followed two passengers	1
7	Thursday	12.00	Observed passengers at T-junction between Gate D and lounge 2		hidden observation of people walking	2
8	Thursday	14.15	Observed passenger behavior in lounge 2		hidden observation of people walking	2

Table 2: Specification of video observations at Schiphol airport terminal

2.4 Data analysis

Data analysis has been conducted in three phases, which were based upon an inductive data analysis approach by Lofland, Snow, Anderson, and Lofland (2006). Phase 1 consisted of preparing the raw video data for analysis, which included merging, splitting and renaming video files. During phase 2, the open coding phase, a first categorization was made using open coding (Strauss, 1987, p.28). Two researchers went through all video materials and noted down what they saw. For instance, ”passengers are running”. No quantitative statements were made about the data. Also, how the behavior was conducted was not described at this point of time (e.g. one passenger was running behind the other). Based on what we observed, we developed a coding scheme. An overview of these categories can be seen in Table 3. The behaviors identified in the open coding were grouped according to the situations that were observed. However, we had to adapt these based on the data. As a first adaptation we divided situation 1 (passengers walking around without guidance) into narrow hallways and wide hallways because we had the impression that the space made a difference on how people behaved. As we did not encounter situations where staff pro-actively approached groups, we eliminated that category. We added a situation called ”using moving walkways” as this turned out to be another interesting situation in the data collection.

Based on the categories from the open coding, we watched all videos again looking for situations where we observed examples for each of the categories. We cut the scenes and put them into a folder

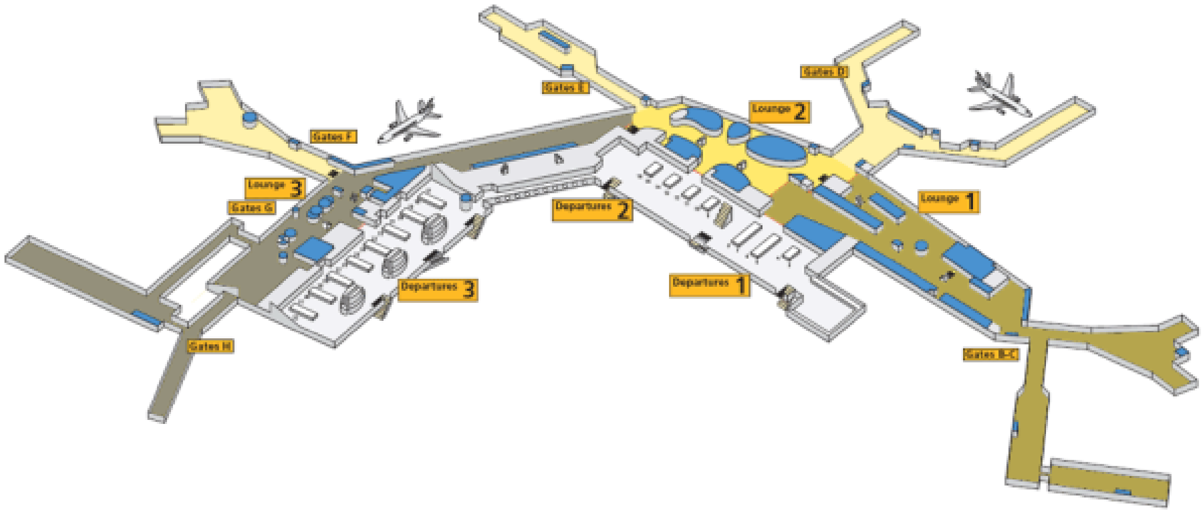


Figure 7: Map showing Schiphol filming locations. Legend: Narrow hallway, Wide hallway, Schengen area; no data recorded, Non-Schengen area; no data recorded

for the respective category. It was possible to have certain parts of the video in multiple folders. In some cases it was not immediately clear to which passenger(s) a video applied, for example in some of the scenes in the wide hallway, where numerous passengers and staff were present at a given moment. In those cases, the videos were edited with circles indicating where to watch so as to avoid confusion for the coders during the final phase of the analysis.

Interaction occurred in either a *narrow hallway* or *wide hallway*. The difference between those two was that we defined a wide hallway as the primary hallways to get from A to B, and the narrow hallways as the secondary hallways which in reality were the gate areas. This is shown in Figure 7. The five other categories (using SSTM, using moving walkway, using information monitors, encounters with staff and encounters with vehicles) can occur in either of those two hallways, and those situations have been categorized independently of the specific hallway (a situation having a passenger look at an information monitor situated in a narrow hallway will thus be categorized in the same category as a similar situation which takes place in a wide hallway).

All video material was cut and placed into one or multiple of the subcategories. This resulted in a total of 406 short video clips. As some material fitted multiple categories, these do not constitute 406 unique clips. In the third step, the focused coding phase, all video material was watched by subcategory, to get to an understanding of how people behaved in certain situations (e.g. how did they position themselves with respect to others and the environment). In the final phase, observations from the data were written down in tables, and from this conclusions were drawn as for the robot behavior. This phase is described in the results section and summarized in the discussion section in which we relate some of our findings to literature.

2.4.1 Exclusion criteria

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

Similarly, in situations where the mobile data collection stroller of our project partners was visible in the videos, these were not coded as it was observed this stroller attracted a lot of attention and thereby created a situation which was not normal - in contrast to the general hidden non-participatory observations.

category	subcategory	#
Narrow Hallway (NH)	1. Pair/Triad exceptions; deviations from normal behavior	2
	2. People walking together with ¿3 people next to each other	5
	3. People search for the way	27
	4. People re-packing luggage	8
	5. Group waiting in the hallway	18
	6. Individuals waiting in the hallway	2
	7. Passengers evading other passengers who move in the opposing direction	4
	8. Pairs walk in 1+1 behind each other	5
	9. Passengers explaining the way to other passengers	1
	10. People are running	3
	11. People overtaking others	2
	12. Passenger sitting on the floor	1
	13. People waving to others	1
	14. Groups organize themselves in pairs	3
Wide Hallway (WH)	1. Pair / Triad exceptions; deviations from normal behavior	n/a
	2. People search for the way	5
	3. People re-packing luggage	7
	4. Groups waiting in the hallway	11
	5. Individuals waiting in the hallway	7
	6. People pass through other groups	3
	7. Groups organize themselves into pairs	7
	9. People are running	20
	10. Pairs walk in 1+1 behind each other	13
	11. Passengers watch their luggage while standing still	3
Encounters with staff (ES)	1. People queuing for staff member	1
	2. Staff member looking around to assist passengers	2
	3. Staff member assisting passengers	17
	5. Passengers are getting information from staff at the SSTD	4
	6. Passenger showing documents to staff	7
Using information monitors (IM)	1. Passenger standing next to (1) other passenger (strangers)	4
	2. Passengers form a 2nd row	2
	3. Passenger looking at boarding pass	2
	Information monitors in wide hallway (Schengen)	
	1. Passenger standing next to (1) other passenger (strangers)	30
	2. Passengers form a 2nd row	21
	3. People take the place of others	2
Using moving walkway (CB)	1. Walking people overtake those who are standing still	12
	2. Passengers do not step on moving walkways currently not in use	1
	3. Passengers walking on the moving walkway	72
	4. Group stands still at the right; leaving left side open for passengers to overtake.	1
	5. Standing still on the moving walkway	11
Using SSTM (TD)	1. Passengers approaching machines	8
	2. Passengers using machines	9
	3. Passenger looking in bag to find documents	2
Encounters with vehicles (EV)	1. Airport vehicles driving around passengers	14
	2. People making way for airport vehicles	2
	3. No interference	13
	(undecided)	9

Table 3: Categories and subcategories after phase II data analysis. Underlined number indicates this NH subcategory is not present in WH category (and vice-versa)

3 Results

In this section we describe our findings and formulate recommendations for the SPENCER robot. Specifically for each subcategory we described the behavior which was observed, where possible supplemented with descriptive statistics. Each section ends with conclusions, which are implications for the behavior or for the perception capabilities of the SPENCER robot. Also we have provided references to illustrative video clips. These clips are referenced to using a path, consisting of a 2-character main category, followed by a 1-digit folder number, and finally the name of the file. The 2-character main category explanation can be found in Table 3 as well as the Section 5 at the end of this report.

3.1 People passing in the hallways

In the first subsection we will discuss all general behavior that we observed when people were walking through the hallway. We will present findings for the narrow and wide hallways together, and where possible we will compare and differentiate between both categories. In some cases subcategories only existed in either the ”narrow hallway” or ”wide hallway” category. Those subcategories will be discussed at the end of this subsection.

3.1.1 Passenger flows

The passenger flows in the wide and narrow hallways differed somewhat. When traversing a space people walking into the same direction organize themselves into flows in order to reduce time needed to traverse. Among others, Daamen and Hoogendoorn (2006) conducted a study where the walking speed of various passenger flows was recorded. Examples of situations include unidirectional, opposite- and crossing flows. Daamen and Hoogendoorn (2006) found, in support of evidence, that for instance speed decreases when encountering a situation with opposite- and crossing pedestrian flows.

In our observations the narrow hallway situations contained various unidirectional flows. To achieve this the flows were seperated by infrastructure, such as a moving walkway (Figure 8). In the wide hallway the flows were more complex, among others due to the multitude of destinations passengers could go to within that hallway (e.g. Schengen barrier, various gates, shops). The various passenger flows in our wide hallway observations have been visualized by ways of a schematic with arrows depicting the various flows over time (Figure 9).

3.1.2 People searching for the way

General search behavior consisted of strolling, sometimes stopping (WH2-Lounge1; NH3-SanFrancisco17; NH3-Vancouver8). We further observed that these people turned their head more than those passengers not searching. In case of a group searching for the way, hand- and arm gestures were observed (WH2-Lounge3; WH2-Lounge4; WH2-Lounge5). In most cases people were looking, and stopping, while walking in a general straight direction, however, at times the opposite was observed in that

Figure 8: The D-gates constitute a narrow hallway, with passenger traffic divided into two unidirectional flows by ways of the moving walkway

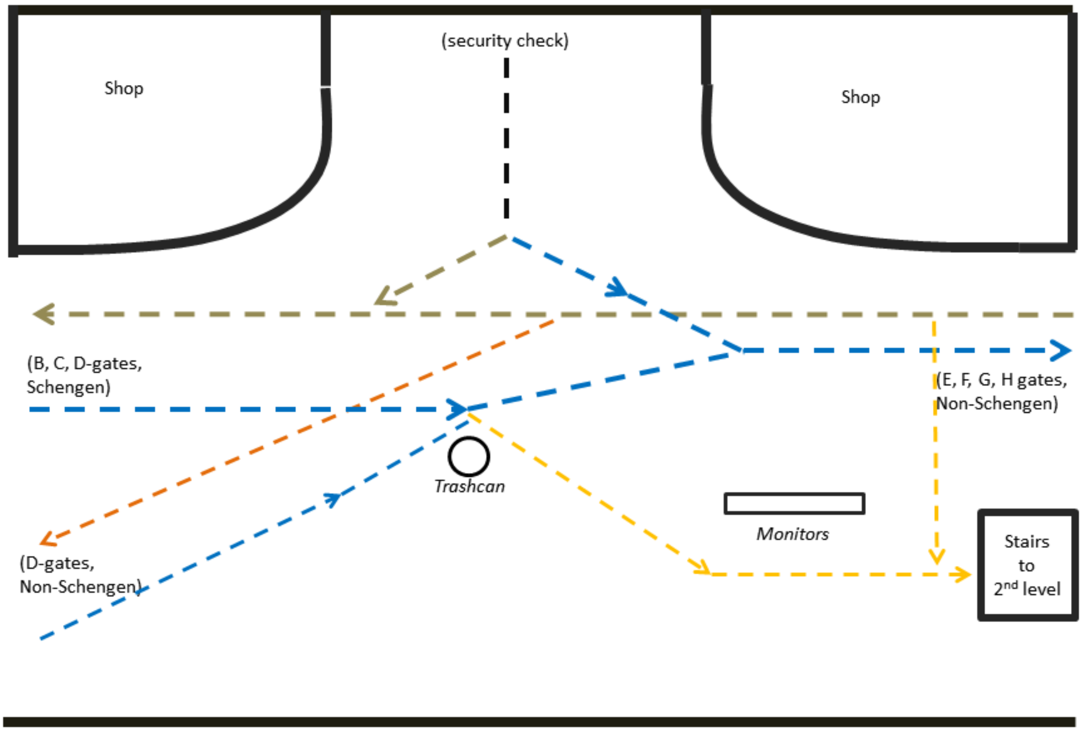


Figure 9: Schematic depicting the various passenger flows in the wide hallway

passengers continued walking, however, they did so while making a zigzag-like pattern through the hallway(NH3-SanFrancisco10; NH3-SanFrancisco13).

When people were exiting the aircraft, they were also observed looking for the way to go. In general, this behavior did not deviate that much from the other ”search” behavior (NH3-Beijing1; NH3-Beijing3). However, what we did see was that after observing the Tokyo flight, more people stopped at an ”inconvenient” place, causing (minor) problems for their fellow passengers (NH3-Tokyo5; NH3-Tokyo7) by standing in their way. We did not observe this for the disembarking process of the Beijing flight. One of the differences between these flights was that at the Beijing flight, KLM staff were visibly present. Furthermore, passengers from the Beijing flight could almost directly step onto a moving walkway, whereas the passengers from the Tokyo flight first had to walk a few meters to the left or right. This situation is illustrated in Figure **??**. Instead of standing in the way to figure out where to go, a better solution could have been the one observed by someone who seemed to be a tour guide in (NH3-Beijing5). She looked for a good place to collect her group, which was not in the walking direction of the majority of the other passengers. In a later video it was observed that she collected and guided a group of passengers. Two conclusions can be drawn:

- Behavior: a robot which collects passengers should place itself, if possible, near the gate exit but opposite the walking direction (Figure 10) when collecting passengers.
- Perception: ”above average” head turning, especially in combination with a lower walking speed, could be an indication that someone is searching for the way.

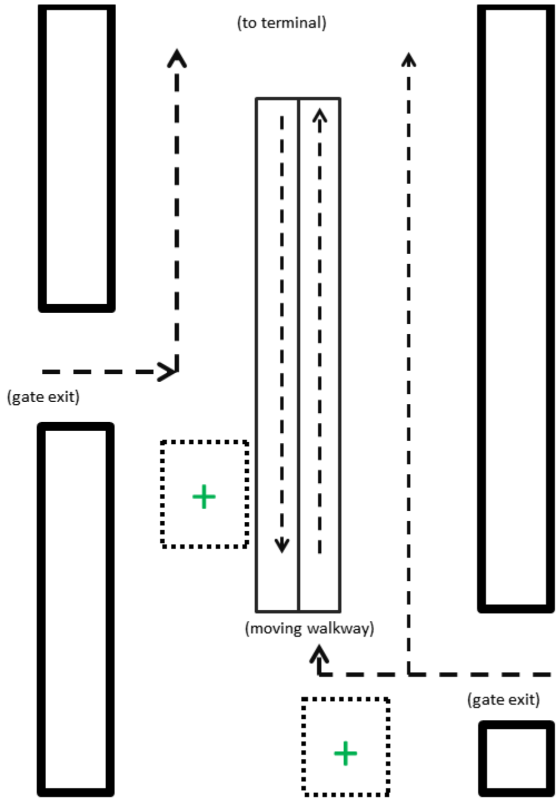


Figure 10: A robot collecting passengers at the gate exit best positions itself opposite the walking direction (marked with a green ”+”). Arrows indicate various passenger flows.

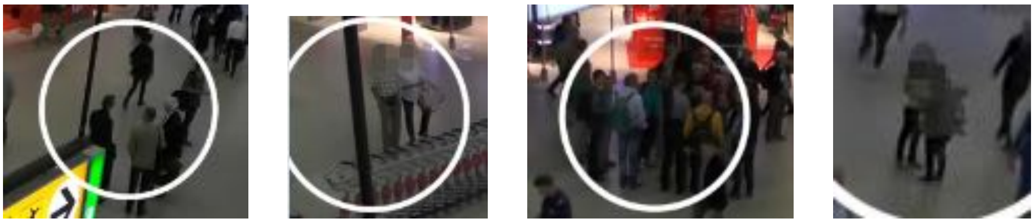
3.1.3 People re-packing luggage

At times we observed people who stopped briefly (NH4-SanFrancisco11), for instance to re-pack their luggage, or search for documents. We observed this kind of behavior 8 times in the narrow hallway, and 7 times in the wide hallway. Also here we observed that there are positions which might be considered inefficient (or inappropriate) by other passengers (WH8-Lounge5). A particularly good position to stand appeared to be the side of the hallway just before entering the T-junction (WH8-Lounge4), or the side in general (NH4-Vancouver5; NH4-Vancouver14). A schematic for appropriate locations in the wide hallway can be found in Figure 19. In general, the insights gained from this is that people could be expected to stop suddenly to re-pack, or re-arrange their luggage. There were two ways for people to re-pack their luggage: either they moved their bag to the front to search something, or they put it on the floor kneeling down. The first behavior resulted in shorter waiting times than the second. So in the first case the robot might just slow down while it might have to stop in the second case.

- Behavior: when the robot detects passengers being guided are ”repacking luggage”, it should slow down or stop
- Perception: passengers holding their bags in front of them, passengers stopping, kneeling down and searching in their bags

3.1.4 Groups waiting in the hallway

We observed several groups of people waiting in the hallway. In total 11 groups were observed in the narrow hallways, and 14 groups in the wide hallway. The size of these groups varied from small groups (2-4 people) to larger groups of 7 and 11. In particular these last two groups seemed to pose an obstacle for other passengers. As can be seen in (NH5-Beijing8, from about 00:10) passengers have to overtake. Of particular interest could be (NH5-Beijing7) where one can see how this group slowly forms, from 2



(a) Circular formation (b) Side-by-side formation (c) Circular formation (d) Vis-a-vis formation

Figure 11: Pairs and triads in the hallway were observed to stand in side-by-side formations, and vis-a-vis formations, while the larger group (c) formed a circular formation

people who do not pose an obstacle to a large group of 11 people. The group of 7 passengers also first stop in the middle of the hallway, however, afterwards reposition themselves to the side of the hallway (NH5-SanFrancisco14). We could not conclude that a larger group size necessarily leads to providing more hindrance, as also a small group appeared to be standing in a particularly bad position for a period of 6 minutes (NH5-SanFrancisco3), and got overtaken on several occasions (NH5-SanFrancisco1), which for the group could have provided an indication that they were standing in the way.

Of the eleven groups observed to be standing still in the wide hallway, only one group was a ”big” group (WH4-Lounge28). It appeared that they all came from the Non-Schengen passport control, and grouped before moving on. Similar behavior was observed for four other groups. The remaining six groups were observed to walk from the Non-Schengen area to the D-gates (either the Schengen barrier, or the Non-Schengen D-Gates), or vice-versa. One of the places groups were standing still more than once was before the exit of the shop (WH4-Lounge27; WH4-Lounge18), also sometimes to wait for a travel companion who was in the shop. Overall it did not appear as if there were big differences between the narrow- and wide hallway here. Given that we have more observations from the same area we made a map of better and worse locations where a group could stand still if necessary.

For the ”groups standing in the hallway”, we believe that there were different reasons for as why they would stand still. The San Francisco group of four people is standing there for conversation. On the other hand, in one instance (NH5-Vancouver3) numerous individuals form a group as they are all waiting to go to the bathroom. Also, sometimes people stand still to get their bearings, as can be seen in (NH5-Tokyo7). Pairs and triads were observed to be more likely to stand in side-by-side formations, the two large groups all stood in circular formations (Figure 11).

From this we conclude that there are spots in the wide hallway which are suited for waiting with passengers. These spots are mostly located at the sideways of the hallway, and in general away from junctions and information screens as shown in Figure 12.

- Behavior: the robot should wait in appropriate places such as the ones shown in Figure 12.
- Perception: In general, locations outside the passenger flow are appropriate to wait.

3.1.5 Individual passengers waiting in the hallway

We observed only one instance of individual people waiting in the narrow hallway and seven in the wide hallway. The person in the narrow hallway stood still and continued strolling on (NH6-SanFrancisco12).

Individual people seem to be waiting for different reasons. These could be to kill the time, but also waiting for travel companions who might be buying something in a shop or taking a bathroom break. From our observations we noted that individuals more often wait in the wide hallway compared with the narrow hallway, and we speculate that this could be because there are more facilities. A popular place appeared to be the trashcan, which could be as it is an obstacle anyway, they would not be standing ”in the middle of the hallway”, nor form an extra obstacle (NH5-Lounge22; NH5-Lounge13). In one instance we observed that a woman was waiting while her travel partner was buying something in a shop (NH5-Lounge17). We noticed that when people are texting (or otherwise engaged with their cellphone) they do not appear to pay that much attention to their surroundings (NH5-Lounge19).

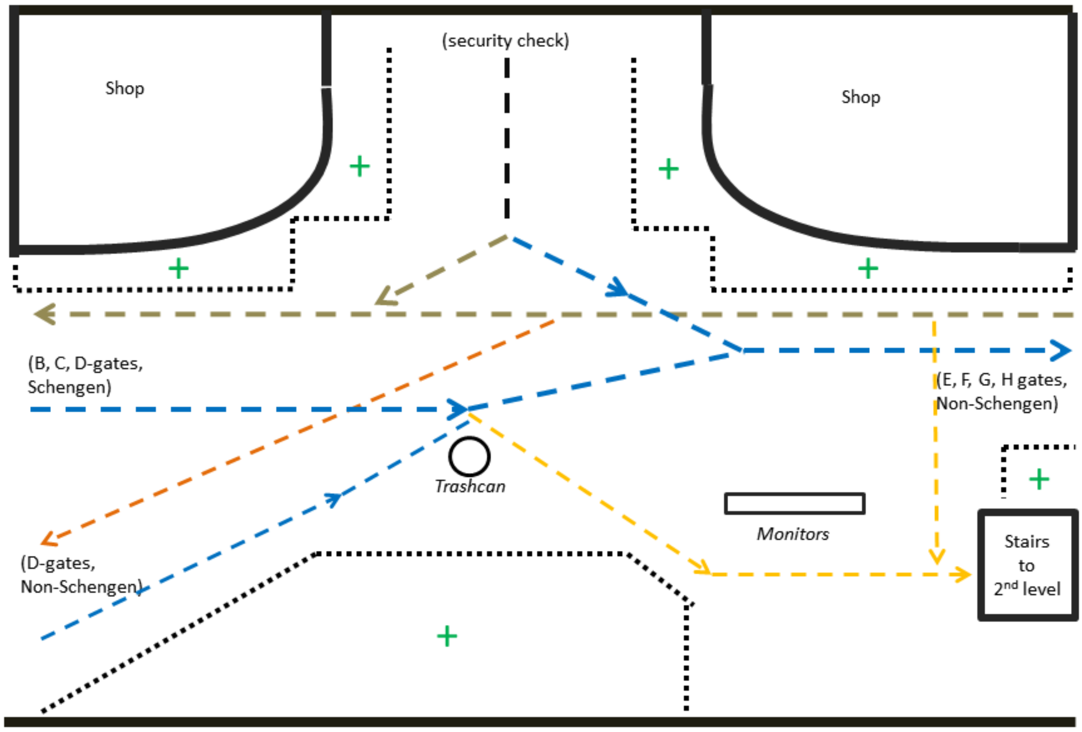


Figure 12: Appropriate positions for waiting with a group in the wide hallway are indicated with green "+"-signs within dotted lines. Colored arrows indicate various passenger flows (different colors represent different directions).

- Behavior: the robot should wait in appropriate places such as the ones shown in Figure 12.
- Perception: not every individual who is standing still is lost, they might just be waiting for someone/something; people might roam around without goal while waiting

3.1.6 Passengers avoiding collisions with other passengers

In the narrow hallway we observed people evading others in the opposing direction. Here the rule of thumb seemed to be that both parties adapt by (slightly) modifying their path. We once observed that a group with a luggage cart seemed to take priority over a group without one (NH7-Beijing7). However, given that we only observed one such a situation we cannot draw any conclusions based upon this.

Encounters with people walking in the opposite direction were more subtle in the wide hallway. At three times we noticed one group split, or forced to split, by passengers moving in the opposite direction. Once a group of three walking next together was passed through by a single person (WH6-Lounge31). Walking with more than 2 people next together seems to happen in a minority of situations, as we will discuss in the next section. If the group had walked in pairs (here 2 + 1) behind each other, this might not have occurred. Two situations with pairs occurred, in one situation two pairs (of equal composition) both split up when encountering each other. The reason for this could be "historical" as one of the pairs was already forced to split mere seconds earlier by an airport vehicle (NH6-Lounge32). In the third occurrence, one of two staff pairs split up when overtaking someone for unknown reasons (NH6-Lounge33).

All these groups were of different composition, and this situation was not observed enough to draw

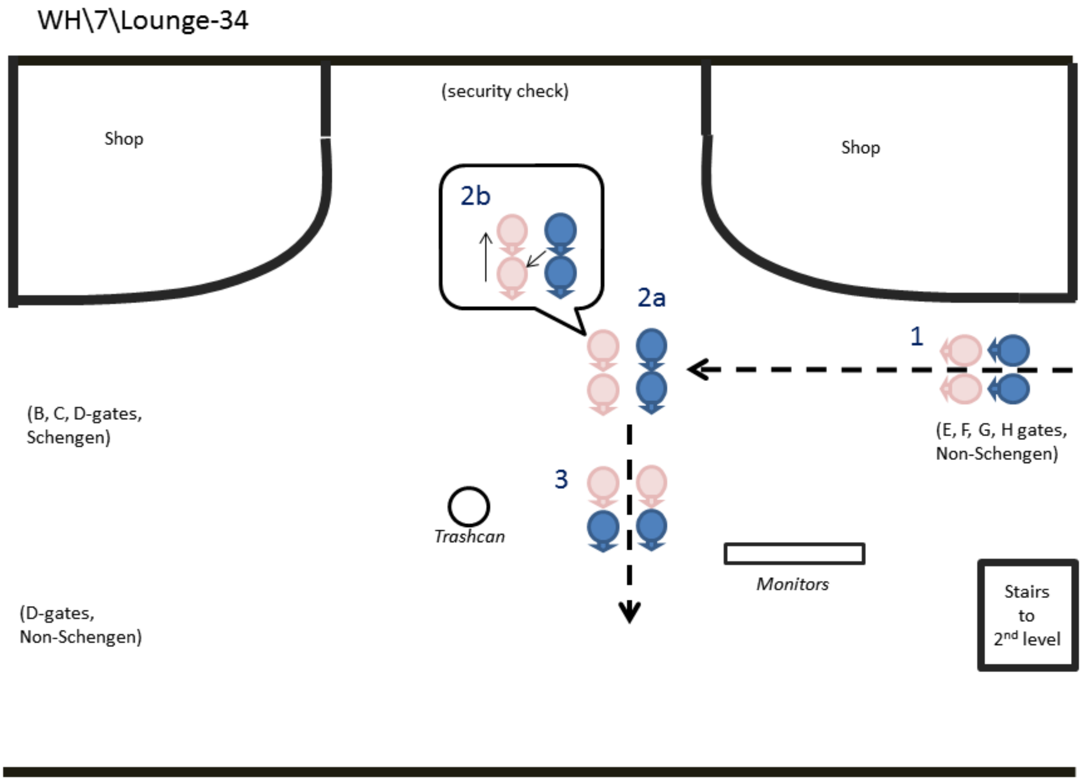


Figure 13: A group consisting of 2 pairs makes a 90 degree turn.

conclusions for how the SPENCER robot should deal with this situation. Thus, we believe that under normal circumstances the SPENCER robot will not encounter situations in which it will force other groups to split up.

3.1.7 Groups organize themselves in pairs and suddenly start walking behind each other

One of the most unexpected findings was that groups mostly organized themselves in pairs. And they always seemed to do so, even when the hallway was wide enough, and the area traversed not crowded thus providing ample space for more people walking next to each other. These pairs also seemed to walk behind each other most of the times. This was noted seven times in the narrow hallways, and also seven times in the wide hallway. For the narrow hallway it should be noted that we here grouped some big groups together, for instance the flight for Beijing (NH14-Beijing9), but also with a 4-person family group (NH14-Junction5) and two groups of staff members (WH7-Aberdeen1; WH7-Lounge36) with 13 and 9 persons, respectively.

Of the five times a 4-passenger group walked into pairs, all but one group organized them in a ”block” of 2x2 passengers. One interesting example (WH7-Lounge34) shows a 4-person group making a 90 degree turn; particular attention should be paid to the way of how the two male persons positioned themselves at the head of the group (Figure 13.

Sometimes the pairs split up and started to walk behind each other as we observed in both types of hallways; 5 times in the narrow hallway, and 13 times in the wide hallway. In most of the cases people did this because of an obstacle being in the way (twice in the narrow hallway, 9 times in the wide hallway). This obstacle was crowds of other people (WH10-Lounge70; WH10-Lounge71), other people looking at information screens (WH10-Lounge68) or people who decided to stop and talk (NH8-SanFrancisco26). We observed three elderly pairs who walked behind each other for no obvious reasons; there was hardly any crowd, or other obstacle (NH8-Vancouver7; WH10-Lounge64; WH10-Lounge69). The reason could be in different levels of individual fitness as will be explained in the discussion.

- Perception: groups of passengers will most likely organize themselves in pairs when following the robot, elderly people might walk one behind the other

3.1.8 Passengers explain the way to others

Once we observed that passengers explained each other the way. It should be noted that no audio recordings were made, so these two could also have been just friendly chats (NH9-Tokyo9). The reason that we believed this is a case of ”explain the way” was that the two passengers were not part of the same group of people (they were individuals), and both were observed to make sense of the information screen together, by use of various pointing gestures.

- Perception: people from different groups can be observed explaining things to each other by gesturing (see also ”encounters with staff”)

3.1.9 People are running

From our observations in the wide hallway we observed recurring behaviors when people tried to run, or otherwise hurry up to make their way through a (semi-)crowded area. When people encountered groups of people (small crowds), they slowed down, and started running again when the area was clear of other people (WH9-Lounge13; WH9-Lounge3). In the case of passengers travelling together, Kendon’s side-by-side formation arrangement (Kendon, 1990, p. 250) could be observed with a father and son who were trying to overtake others (WH9-Lounge6). This was ineffective as they could not pass. However, if they were to split up and run behind each other this would make overtaking more successful, as we observed multiple times (WH9-Lounge10; WH9-Lounge15; WH9-Lounge19; WH9-Lounge20). In some of these situations the distance between the members of the groups increased due to some of them being faster than others.

- Behavior: if the robot walks fast (runs) it hast to slow down when encountering crowded areas
- Perception: passengers belonging to one group while running might be one behind the other, the distance between these passengers might increase considerably

3.1.10 Encounters with staff

While KLM (or Schiphol) does not provide staff to actively search for passengers in need of help, staff do walk around, and from what we noticed they are always willing to direct passengers. This was of course not limited to KLM ground staff, but also applies to security- and cleaning staff. Assisting someone generally consisted of establishing contact, exchanging information (multiple times) and saying goodbye; leaving. In the event that there were multiple passengers waiting for one staff member, people waited, or queued. While we have initially looked at this queuing behavior as well we only found one example (ES1-Beijing6), and whether or not this was a queue was also arguable hence we did not go into this further. We categorized the videos from passenger encounters into two categories, these being ”staff member assisting passengers” and ”staff looking around to assist passengers”.

Staff member looking around to assist passengers Twice we observed that a staff member actively approached passengers (ES2-Vancouver2; ES2-Vancouver3). This occurred at the disembarkation from the Vancouver flight, and the staff members were part of the company which assists passengers requiring assistance. Therefore, the staff already knew they were looking for certain passengers. Apart from this we have not seen this behavior during the remainder of our observations; the initiation of contact seemed to occur at the side of the passenger, who would approach a staff member walking around.

Staff member assisting passengers We observed (partial) interactions with staff and passengers 16 times. Five of these were during the disembarking of the Beijing flight; these staff were specifically there to assist passengers who were either in need of directions or in need of a transfer. Four formations were observed after the interaction had started; side-by-side (3), vis--vis (6), L-shaped (5) and circular formations (2). It should be noted that there are three or more actors required for a circular formations; whereas in 9 interactions only two actors were present. This conforms with existing research findings that groups of size 2 tend to occur much more often than bigger sizes (Ciolek, 1976).

We observed that passengers showed travel documents to airport staff seven times. One of these encounters was identified as being "showing documents", however, due to the angle of the camera it was impossible to make out what document it was (ES4-Vancouver1). In four of the remaining six cases, the staff member physically took the boarding card and read it out loud (i.e. ES4-Beijing5), in the other two cases (ES4-SanFrancisco3), the staff member only read what was stated on the boarding card.

Out of the remaining 15 encounters, in 9 encounters staff were observed to gesture with their arms where the passenger had to go (i.e. ES3-SanFrancisco5). In four of these encounters the staff member also used his/her head to indicate direction (i.e. ES3-Beijing2; ES3-SanFrancisco7). In one encounter (ES3-Beijing1) we observed the staff member only using her head to indicate direction. Overall, indicating direction could be important as a way to end the guiding scenario for the SPENCER robot.

- Behavior: having finished the conversation about where to go, the robot should indicate the direction with a gesture
- Perception: the robot should be enabled to read boarding cards as they seem to be a common way to share information

3.1.11 Passengers overtaking

We also noted behaviors which did not occur frequently, but nonetheless could be important for the SPENCER consortium, at least to be aware of. In two instances we observed passengers actively overtaking passengers in front of them. With actively overtaking we refer to situations where people speed up, overtake, and then slow down. In both instances people overtook at their left side because the people in front of them walked right and there was no room (NH11-Vancouver15; NH11-SanFrancisco21). This could be culturally motivated as people in Western Europe mostly walk at the right and overtake at the left side.

- Behavior: if the robot needs to overtake people, the left side in many situations might be the better side to do so

3.1.12 Children running around

In two situations we observed children running away from their parents (WH9-Lounge54; ES3-SanFrancisco5). They moved rather quickly and not in line with passengers' general walking direction.

- Behavior: the robot might have to slow down or stop when detecting a running child
- Perception: it might be beneficial to be able to detect running children

3.2 Self-service transfer machines

It was, in general, observed that there were hardly any queues in front of the self-service transfer machines. We observed seven situations in which passengers approached the self-service transfer machine where it was clear to see what actually happened. In the other cases we observed passengers' usage of the machines without the actual approach to the machine. In most cases passengers just walked slowly up to the machine (i.e. TM1-SanFrancisco5; TM1-SanFransisco3). Two exceptions have been observed; one in which the passenger had to queue (TM1-SanFransisco10), and another example in which two passengers appeared to approach, however, it turned out that one passenger just had to find a place to tie his shoelaces while the other passenger was looking at the machine (TM1-SanFrancisco2).

After having made use of the SSTM it appeared that most people were still wondering where to go. We arrived at this conclusion as passengers who had help from staff as well as those who did not were

(a) Circular formation (b) Side-by-side formation (c) Circular formation

Figure 14: Passengers at the SSTM. In case of a group used a SSTM the number of passengers operating the machine was limited to 1 or 2

observed to look around after leaving the SSTM. Staff members who assisted in using the machines did provide directional cues to passengers (see below).

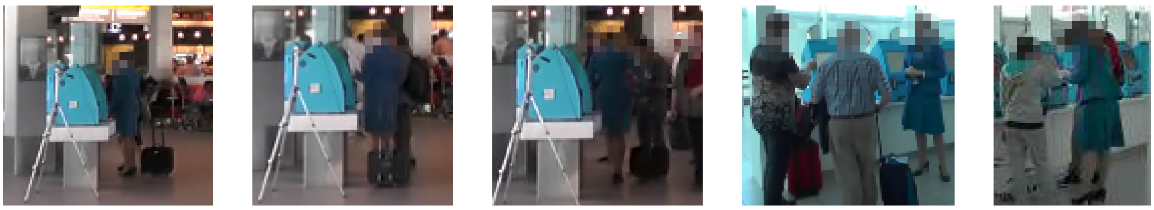
3.2.1 Passengers using machines without help of staff

10 times passengers were observed using a self-service transfer machine without help of staff. Of these nine videos, four were single passengers, the other small groups consisting of 2-4 persons each. Single passengers, were standing right in front of a machine (TM2-Vancouver1). Also in the groups at all times there was only one passenger operating the machine, however, usually one of the fellow passengers stood next to him/her (TM2-SanFrancisco4; see Figure 14), and the others all behind these two passengers (TM2-SanFrancisco9).

3.2.2 Staff member assisting passengers at SSTD

At the SSTM's we observed five interactions between staff, passengers, and machine. Three times these interactions took place at the F-gates. Here, the same staff member assisted three (single) passengers (Figure **??**: A, B and C). We observed that the staff member and individual passengers always formed a side-by-side formation facing the machine. At the conclusion of the transaction with the passenger, the staff member used a pointing gesture to indicate where the passengers had to go.

Two other interactions were (partially) observed at Transfer 6 (Figure 15: D and E). Here, a staff member assisted two couples. In both situations the staff member formed a rectangular formation; where the passengers were closest to the SSTD machine. No gestures by the staff member were observed; however this could be as we observed partial interactions.



(a) Single passenger (b) Single passenger (c) Single passenger (d) Pair (e) Pair

Figure 15: Staff assisting passenger at the SSTD; usually took the form of a circular or rectangular formation

Overall it can be noted that the formations of people in front of the machines seem to differ depending if staff is present or not. If staff is present, the formations are more circular or rectangular, whereas if no staff is present, the users rather position themselves in pairs directly facing the machines.

- Behavior: the robot should indicate the direction at the end of the interaction
- Perception: when the robot acts as SSTM, it might be useful to recognize from the positions of the users if a staff member is present who takes over some of the tasks for the robot

(a) (b)

Figure 16: Participants formed a formation around the information screen resembling a half circles

3.3 Information monitors

Within the ”information monitors” category we found four main spatial behaviors: *passenger standing next to a stranger*, *passengers forming a 2nd row*, *passengers forming a half circle around the information monitors* and *passengers looking at their boarding pass*. For the analysis, we looked separately at the data collected from the videos in the wide hallway and in the narrow hallway. The difference here was that the wide hallway appeared to allow for more people to look at the monitors, and that they were positioned more prominently. A second reason is that the fourth spatial behavior (looking at boarding pass) could not be observed in the wide hallway due to the distance between camera and passengers. Also, significantly more observations were made in the wide hallway (53 versus 7).

At two times we observed a passenger standing next to a stranger in the narrow hallway. In both cases we did not have measurements of the distances between both actors, and therefore did not know whether or not one would come too close. We estimated that in both cases there was about 30-40 centimeters space in between both actors. However, given that no physical compensatory behaviors such as stepping away were observed we are proceeding with the assumptions that this is not the case (IM2-Vancouver1; IM2-Aberdeen1). It could be that it is socially acceptable to stand next to a stranger provided that you keep at least some distance between one another, especially if there is not much space around. In the discussion we will briefly reflect on this.

For the information monitors in the wide hallway, similar behavior was observed: the passengers appeared to stand quite close together (IM2-Lounge11). One passenger displayed what might be a physical avoidance behavior to personal space invasion. In clips (IM52-Lounge15 and IM52-Lounge16) we observed that a passenger stepped to the left and right respectively when someone stood next to him (at the other side). One explanation could be a reaction to personal space invasion; e.g. the other one stood too close to him (see Section 4.1); another explanation is that coincidentally he was looking at another screen; therefore requiring him to make a small step.

In the wide hallway we have looked at people’s formations when standing next to each other, and facing the information monitors. What we observed in general was that the passengers were forming half-circles around the information monitors as can be seen in Figure 17. The crowdedness varied greatly. Even when it was not crowded around the monitors it could be observed that there are multiple layers (one very close, one halfway the hallway), and that the people were spread out more or less in a half circle. The boundaries of the area where you ”should” watch the information monitors appeared to be given by the surroundings; i.e. there were 6 monitors determining from where people could read the information. Also the hallway was at times crowded (IM51-Lounge34; IM51-Lounge42).

Similarly, at the temporary information monitors at the end of the E-gates (Figure 17a) 4 times passengers formed a second row instead of standing next to other passengers who were already there. From what we observed in the four videos, it appeared as if people started to form a second row as soon as they could either not see the monitors from the particular angle where they would have to stand, or that they would be in the way of other passengers by doing so (f.e. IM1-Vancouver4). An illustration of this problem is provided in Figure 17b. In the wide hallway this could be observed as well. Figure 19 enriches the figure of passenger flow and waiting areas with locations that appear to be appropriate for watching the information monitors in the wide hallway. As can be seen from the pas-

(a)

(b) Schematic showing socially appropriate area to watch monitors. Dotted lines indicate various passenger flows.

Figure 17: Three temporary information monitors located at the end of the E-gates

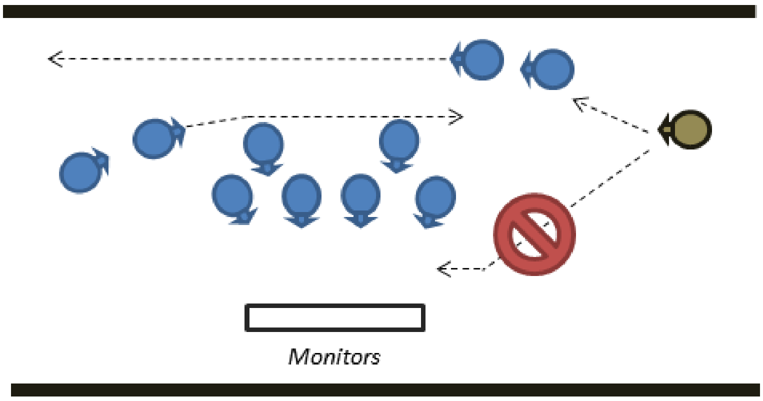


Figure 18: Scenario recommendation: do not drive in between information monitors and people looking at them. Dotted lines indicate various passenger flows.

senger flow, crossing between the people who look for information and the information screens might not be a good option. This particular robot behavior is also illustrated in Figure 18.

- Behavior: it appears that it might be inappropriate for the robot to drive between information monitors and people looking at them
- Perception: the robot might have to be able to recognize people looking at information screens (groups facing the same direction, standing in half-circles or multiple of these behind each other)

3.4 Encounters with vehicles

All powered vehicles driving at Schiphol are operated by airport staff. The size and purpose of these vehicles varies, as can be seen in Figure 20. From our observations we can make two general assumptions. The first assumption is that the vehicles drive relatively fast and are not equipped with a horn, or anything to signal. In order to make way the drivers would slow down, and usually raise their voice to indicate they would like to pass (see f.e. EV-Vancouver3). They do this in a polite way, as passengers are customers and should thus be treated as such. The second assumption is that people sometimes just do not hear the vehicle. There is a lot of ambient noise and the vehicles are battery-powered so the noise generated by the engines is low. For the SPENCER robot one recommendation could be to equip the robot with some sort of functional noise, as proposed in (Lohse et al., 2013).

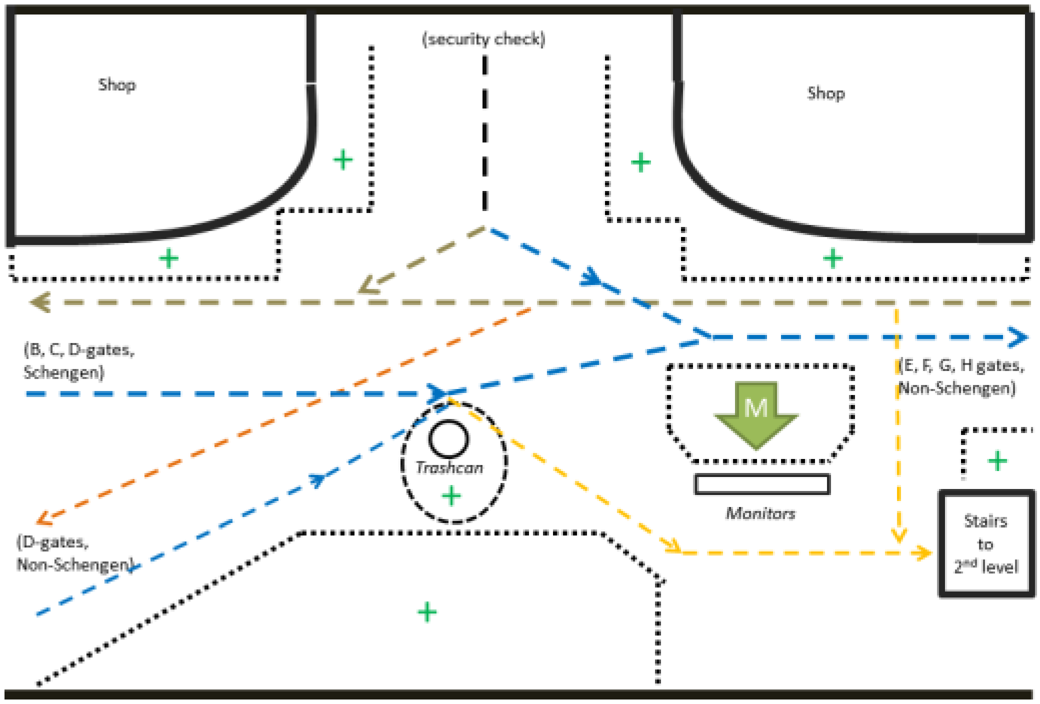


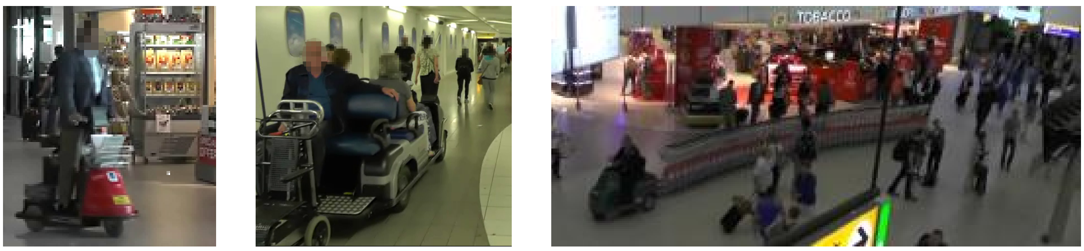
Figure 19: Wide hallway schematic showing appropriate locations to wait (or look at information monitors), indicated with green ”+”-signs within dotted lines. Colored arrows indicate various passenger flows (different colors represent different directions).

In total we observed 38 encounters with vehicles and passengers. Of these 38 encounters, a lot of times (13) there was no interference between vehicle and passengers; the driver of the vehicle would just drive around the passenger without having to slow down. 14 times we observed the driver of the vehicle going out of the way and making way for the passengers. Only 2 times we observed that the passengers went out of the way; which translated into going a little bit more to the right or left side of the hallway. Therefore it appears that passengers take priority at the airside. The remaining nine encounters did not show clearly whether it was the driver or passenger who adapted to the situation.

- Behavior: signal to other passengers in a polite way that the robot is approaching; give priority to passengers, e.g. by slowing down or adapting walking direction
- Perception: the robot should detect if its path will collide with other passengers' paths

3.5 Moving walkway

The narrow hallways are equipped with moving walkways for passengers, so as to reduce walking distance for passengers. Most moving walkways have a width that allows people to overtake others.



(a) Staff transportation (b) Passenger transportation (c) A number of luggage carts

Figure 20: Various of the powered vehicles at Schiphol

Walking speed	Single passengers				Dyads of passengers			Percentages (single/dyad)
	Left	Middle	Right	Undecided	Left	Middle	Right	
Standing still	2		1		1	2	1	8% / 13%
Slow	6	2	1	1	1	5	1	29% / 22%
Normal	9		5	1	4	10	2	42% / 55%
Fast	2	1	4	1	2		1	21% / 10%

Table 4: Comparison of single and dyad passengers motion behavior on a moving walkway

This can be necessary as we observed situations where all people are walking (such as CB3-Tokyo13 and CB3-Tokyo10) or where some passengers are just standing still (CB1-Tokyo3). In one situation we observed that a moving walkway was out of use. In this case the passenger did not walk on the moving walkway but next to it (CB2-Vancouver3).

In total we observed 83 situations on the moving walkway. The total duration of video material containing a moving walkway was 12 minutes. One independent coder coded these fragments to indicate whether people either were standing still (11), walking slowly (20), walking at a normal speed (39), or walking fast (13). 38 of these situations contained single passengers, 31 a pair of people, and the remaining situations featuring groups containing respectively three (4), four (1), five (3), six (3) and ten (2) passengers.

In Table **??** we compared the motion behaviors of single passengers and dyads of passengers. Of the 37 situations identified as containing single passengers, 19 were identified as walking at the left side, 3 in the middle, 11 at the right side and 3 undecided. Of the 31 situations with pairs of people, 8 pairs walked at the left side, 17 stood next together and thus occupied the whole width, and 5 passengers stood at the right side. From this it appears that passenger, if any, have a preference to stay at the right side of the moving walkway for other passengers to overtake. However, it is also possible that dyads block the whole walkway for other passengers behind them. If the people approaching from behind are the ones that the robot guides, this might pose a challenge for person tracking.

Overall there are no obvious differences between the individual passengers and the dyads. Compared to the dyads, more single passengers walk fast but then again also more walk quite slowly. A general trend is that passengers walk on the moving walkway instead of standing still. As the robot cannot get onto the walkways, this implies that it would have to drive quite fast to reach the other end of the walkway at the same time as the passengers.

- Behavior: the robot might have to adapt its speed to people walking on the mobile walkways
- Perception: the robot might have to track passengers while they are walking on the moving walkway, are overtaking, being overtaken, or blocked by other passengers

4 Literature related to our findings

In this section we will relate our findings to the existing related literature such as on pedestrian behavior and interpersonal distance. Where possible we relate to literature related to relevant contexts such as airports and train stations. We will discuss different topics, these being related to formations of groups (Section 4.1), speed (Section 4.2) and politeness-related behaviors (Section 4.3).

4.1 Formations & spatial organizations of small passenger groups

Based upon our general observations we found two important re-occurring patterns, which deal with how passengers navigate through a crowded area. The first general behavior was that a group would split itself into pairs (walking behind each other), even when there is enough space (width) in the hallway. The second observation deals with crowding, in that guided passengers will likely change formation in order to successfully traverse crowded areas.

A question would be how the robot should deal with this, from a technical point of view. For the first observation it would be necessary to know how many people walking behind each other the robot can track, and related to that if it would be possible for the robot to remember these persons even while they are not being tracked as still belonging to the group.

During our observations of passengers who were looking at information monitors, we observed that people frequently stood next to a stranger, at a distance that could be seen as socially inappropriate (too close), given the lack of crowdedness in the situations. In a normal situation, people maintain a minimum interpersonal distance between each other, not only when facing each other, but also when they stand next together. The reason why people stood seemingly close together could perhaps be explained by a combination of the theories of (Kendon, 1990) and (Hall, 1966).

The first thing to consider in that there is basically only one type of focused encounter. Due to the limited viewing angle people have to be able to see what is actually written on the monitors (see Figure 11 and Figure **??** for an illustration). This could be similar to a situation of a crowded shop, where people lean close to each other to grab products from a shelve. For the situation observed near the information monitors this implies: because people focus on the information monitors (and thus make no eye-contact with people standing next to them), they can stand closer together.

4.2 Walking speed

We observed that people tend to slow down when they encounter a crowd. In literature we could not find this exact finding, however, Young (1999) observed (adult) average walking speed at an airport. He observed that passengers tend to slow down when they are either:

1. Approaching a travel-path decision
2. Approaching / in the presence of directional signs and/or aircraft arrival-and-departure boards

We also observed people who slowed down in these situations, even though we did not actually measure speed. However, we can back this up by our observations. Young's results indicate that the average walking speed was 1.34 m/s (sd=0.27 m/s) under normal walking conditions (Young, 1999). Small differences were observed between man and woman, with man walking on average slightly faster. Furthermore, it was, perhaps unexpectedly, observed that people carrying bags walked significantly faster than people without bags. Our analysis was not detailed enough to either support or reject his final finding. However, it might be worthwhile to keep them in mind.

While age was not a main focus of our observations, it is not unlikely that the SPENCER robot will guide elderly passengers. As described above, we saw that sometimes couples of elderly walk behind each other, which could be due to one of the two being in a physically healthier state. When looking at the videos it seemed that elderly in general walked slower.

Bohannon (1997) compared both average and maximum gait speeds[3] of pedestrians from different age groups (N=230). Speed was measured by having participants walk in a lab over a distance of 7.62 meters (or 25 feet). He found that the maximum speed declined by age, whereas the comfortable walking speed was more or less stable over time (Figure 21). For the SPENCER robot this would imply that if people have to hurry and the robot drives faster than 1.6 m/s, it is imperative that the robot has some awareness of the age of the passengers. However, given that these speeds were measured over a short distance, other tests which yield reference values such as the 6-Minute Walking Test (6MWT) are potentially more useful in our context.

4.3 Politeness- and predictability-related recommendations

Oftentimes passengers were holding up, or strolling in the hallway. We have interpreted these situations as being one of two categories: searching for the way or waiting. For the SPENCER robot this would be an important feature to distinguish as the changes of a successful interaction would be greater when approaching those people who are likely in the need of help.

Based upon our observations from the airport vehicles, we noted that if one of the parties went out of the way to avoid a collision, this would be the vehicle driver. However, usually these vehicles drove at a high speed; not a speed a healthy human could easily keep up with. Therefore, this will not likely require the robot to manoeuver quickly through a crowd, and hoping the passengers who follow

[3]Gait speed is measured over a short distance, thus does not include endurance as a factor.

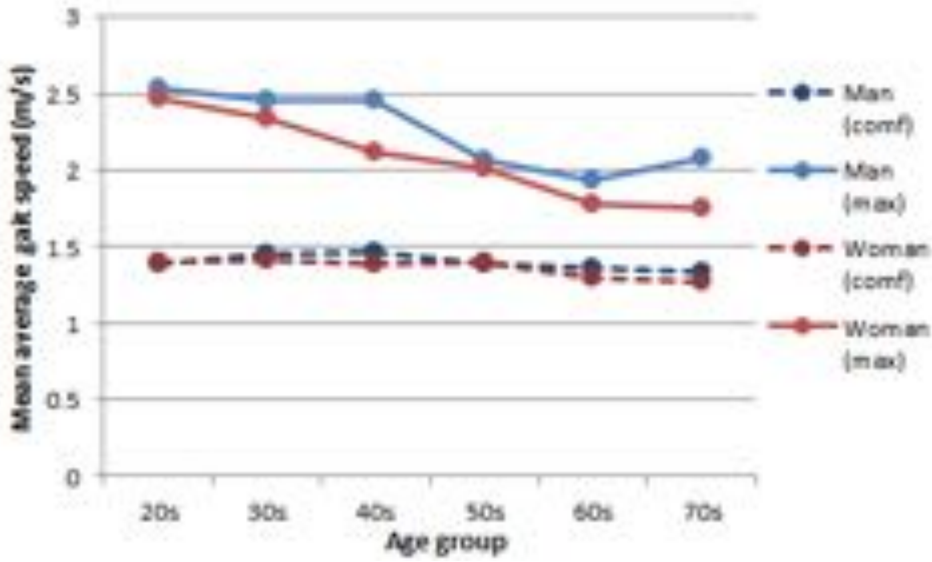


Figure 21: Average maximum gait speed (max) decreases significantly by age group, especially compared with comfortable gait speed (comf).

the robot keep up with the robot. We recommend that the robot displays politeness, and legibility (Lichtenthäler & Kirsch, 2013), for instance by slowing down and driving forward without making a lot of abrupt turns (Lichtenthäler & Kirsch, 2014).

In line with our observations of the closing of the interaction between staff and passengers, and with the KLM brand value ”positive concluding of the interaction”, we recommend that the robot provides the direction in a way which is as unambiguous as possible, by pointing where to go. We do not know yet whether or not the SPENCER robot should use its whole body to point or rather only its head.

Above recommendation is also related to research by Hicheur, Vieilledent, and Berthoz (2005), who found that head motion is a predictor of future walking direction. In an experiment where participants (n=10) were asked to walk along a 20-meter figure of 8, head motion was found to be a predictor of future walking direction. Similarly Hollands, Patla, and Vickers (2002) found that people use a combination of head and eye movement to indicate their heading. This unconscious nonverbal behavior could explain why people do not bump into each other when traversing a crowded public space, and might thus warrant implementation on the SPENCER robot.

A question which arose during the analysis was how to deal with the moving walkway. If the robot cannot track people who are walking on the walkway, we should design the instructions to the users in a way that they also will not do so; or are not encouraged to do so.

While a robot can be designed to have a humanoid appearance, and some behavior can indeed be matched, so as to copy human behavioral norm (such as slowing down, and adjusting speed), other behaviors are required to be robot-specific. For instance, we recommend, or observed, that in order to pass a crowded area effectively, the group has to break its formation. After having cleared a crowded area the robot could collect its passengers before moving on. This was not observed at Schiphol Airport, as the ”guiding” itself does not take place in this form.

One final recommendation perception-wise would be to consider that a robot in public space will be considered being a ”celebrity”, in that people will take pictures of the robot. How to deal with this during mission execution has yet to be decided, also, what to do when people block the way because they want to touch the robot, or take a picture with it.

5 Conclusion

In order to inform the design of the SPENCER guide robot’s motion behavior we have conducted a contextual analysis at the envisioned deployment location, Schiphol Airport. As this robot will guide passengers from their arrival gate to the Schengen barrier, we have collected video data at a variety of places which are relevant for the SPENCER robot, including the disembarkation of several flights, passengers making use of self-service transfer machines, encounters with staff members providing directions to passengers and current passenger movement through crowded areas.

We analyzed these data using an inductive data analysis approach by Lofland et al. (2006). This method consists of three phases, and after the third phase we made inferences about the data which were then structured into various categories. The results can roughly be divided into two categories: implications for the SPENCER robot's motion behavior and for the perception of the robot. For a brief list of all recommendations and observations we refer to Appendix A. In the discussion section these recommendations are related to existing literature.

Acknowledgements

This research has been supported by the European Commission under contract number FP7-ICT-600877 (SPENCER).

The authors would like to acknowledge Ruben Alblas, Manuel van Lijf and Dana Willemse from KLM for their support and help in acquiring the necessary permits from Schiphol Airport. We would also like to acknowledge Nancy Ho and Dicky Man from KLM Ground Staff who guided us around and provided us with invaluable insight information.

Hayley Hung, Marieke van Rooij and Lu Zhang provided insightful comments and feedback, both before and after the data collection event. Finally, thanks to Daphne Karreman for her help during data collection- and analysis.

References

Bohannon, R. W. (1997). Comfortable and maximum walking speed of adults aged 2079 years: reference values and determinants. *Age and ageing*, *26*(1), 15–19.

Ciolek, M. (1976). Location of static gatherings in pedestrian areas: an exploratory study. In *Ergonomics society of australia and new zealand conference, 13th, 1976, canberra.*

Daamen, W., & Hoogendoorn, S. P. (2006). Free speed distributions for pedestrian traffic. In *Trb-annual meeting, washington.*

Hall, E. T. (1966). *The hidden dimension* (Vol. 1990). Anchor Books New York.

Hicheur, H., Vieilledent, S., & Berthoz, A. (2005). Head motion in humans alternating between straight and curved walking path: combination of stabilizing and anticipatory orienting mechanisms. *Neuroscience letters*, *383*(1), 87–92.

Hollands, M. A., Patla, A., & Vickers, J. (2002). look where youre going!: gaze behaviour associated with maintaining and changing the direction of locomotion. *Experimental Brain Research*, *143*(2), 221–230.

Kendon, A. (1990). *Conducting interaction: Patterns of behavior in focused encounters* (Vol. 7). CUP Archive.

Lichtenthäler, C., & Kirsch, A. (2013). Towards legible robot navigation-how to increase the intend expressiveness of robot navigation behavior. In *International conference on social roboticsworkshop embodied communication of goals and intentions.*

Lichtenthäler, C., & Kirsch, A. (2014). Goal-predictability vs. trajectory-predictability: which legibility factor counts. In *Proceedings of the 2014 acm/ieee international conference on human-robot interaction* (pp. 228–229).

Lofland, J., Snow, D., Anderson, L., & Lofland, L. H. (2006). Analyzing social settings: a guide to qualitative observation and analysis. *Belmont, CA: Wadsworth*.

Lohse, M., van Berkel, N., van Dijk, E. M., Joosse, M. P., Karreman, D. E., & Evers, V. (2013). The influence of approach speed and functional noise on users' perception of a robot. In *Intelligent robots and systems (iros), 2013 ieee/rsj international conference on* (pp. 1670–1675).

Strauss, A. L. (1987). *Qualitative analysis for social scientists*. Cambridge University Press.

Young, S. B. (1999). Evaluation of pedestrian walking speeds in airport terminals. *Transportation Research Record: Journal of the Transportation Research Board*, *1674*(1), 20–26.

Abbrevations

KL	KLM flight number
KLM	Koninklijke Luchtvaart Maatschappij, or: Royal Dutch Airlines
MCT	Minimum Connection Time
SPENCER	Social situation-aware perception and action for cognitive robots. EU FP7 project FP7-600877
SSTM	Self-Service Transfer Machine
f.e.	For example
Abbreviations used for reference to video files in results section:	
WH	Wide hallway
NH	Narrow hallway
CB	Moving walkway
ES	Encounters with staff
EV	Encounters with vehicles
TD	Passengers using self-service transfer machines (either in the hallway or at the transfer area)
IM	Passengers looking at information monitors

A List of recommendations

This appendix provides a summary of the recommendations we have formulated in the Results section. These recommendations are either implications for the behavior or implications for the perception capabilities of the SPENCER robot.

Section 3.1.2

- Behavior: a robot which collects passengers should place itself, if possible, near the gate exit but opposite the walking direction (Figure 10) when collecting passengers.
- Perception: "above average" head turning, especially in combination with a lower walking speed, could be an indication that someone is searching for the way.

Section 3.1.3

- Behavior: when the robot detects passengers being guided are "repacking luggage", it should slow down or stop
- Perception: passengers holding their bags in front of them, passengers stopping, kneeling down and searching in their bags

Section 3.1.4

- Behavior: the robot should wait in appropriate places such as the ones shown in Figure 12.
- Perception: In general, locations outside the passenger flow are appropriate to wait.

Section 3.1.5

- Behavior: the robot should wait in appropriate places such as the ones shown in Figure 12.
- Perception: not every individual who is standing still is lost, they might just be waiting for someone/something; people might roam around without goal while waiting

Section 3.1.7

- Perception: groups of passengers will most likely organize themselves in pairs when following the robot, elderly people might walk one behind the other

Section 3.1.8

- Perception: people from different groups can be observed explaining things to each other by gesturing (see also "encounters with staff")

Section 3.1.9

- Behavior: if the robot walks fast (runs) it hast to slow down when encountering crowded areas
- Perception: passengers belonging to one group while running might be one behind the other, the distance between these passengers might increase considerably

Section 3.1.10

- Behavior: having finished the conversation about where to go, the robot should indicate the direction with a gesture
- Perception: the robot should be enabled to read boarding cards as they seem to be a common way to share information

Section 3.1.11

- Behavior: if the robot needs to overtake people, the left side in many situations might be the better side to do so

Section 3.1.12

- Behavior: the robot might have to slow down or stop when detecting a running child
- Perception: it might be beneficial to be able to detect running children

Section 3.2

- Behavior: the robot should indicate the direction at the end of the interaction
- Perception: when the robot acts as SSTM, it might be useful to recognize from the positions of the users if a staff member is present who takes over some of the tasks for the robot

Section 3.3

- Behavior: it appears that it might be inappropriate for the robot to drive between information monitors and people looking at them
- Perception: the robot might have to be able to recognize people looking at information screens (groups facing the same direction, standing in half-circles or multiple of these behind each other)

Section 3.4

- Behavior: signal to other passengers in a polite way that the robot is approaching; give priority to passengers, e.g. by slowing down or adapting walking direction
- Perception: the robot should detect if its path will collide with other passengers' paths

Section 3.5

- Behavior: the robot might have to adapt its speed to people walking on the mobile walkways
- Perception: the robot might have to track passengers while they are walking on the moving walkway, are overtaking, being overtaken, or blocked by other passengers

Dynamics of Social Positioning Patterns in Group-Robot Interactions*

Jered Vroon, Michiel Joosse, Manja Lohse, Jan Kolkmeier, Jaebok Kim,
Khiet Truong, Gwenn Englebienne, Dirk Heylen, and Vanessa Evers

***Abstract*— When a mobile robot interacts with a group of people, it has to consider its position and orientation. We introduce a novel study aimed at generating hypotheses on suitable behavior for such social positioning, explicitly focusing on interaction with small groups of users and allowing for the temporal and social dynamics inherent in most interactions. In particular, the interactions we look at are approach, converse and retreat. In this study, groups of three participants and a telepresence robot (controlled remotely by a fourth participant) solved a task together while we collected quantitative and qualitative data, including tracking of positioning/orientation and ratings of the behaviors used. In the data we observed a variety of patterns that can be extrapolated to hypotheses using inductive reasoning. One such pattern/hypothesis is that a (telepresence) robot could pass through a group when retreating, without this affecting how comfortable that retreat is for the group members. Another is that a group will rate the position/orientation of a (telepresence) robot as more comfortable when it is aimed more at the center of that group.**

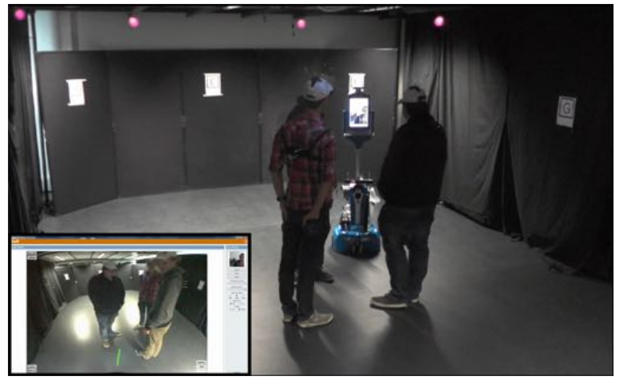


Fig. 1. Example of the interactions described in the paper. A group of four participants discuss a murder mystery. One of them is remotely present through a robot, and has to go through several approach/converse/retreat cycles. The inset shows the interface as seen by this participant.

I. INTRODUCTION

As robots are slowly making their way into situations where they interact with groups of people, it becomes more and more important that we understand which robot behaviors are suitable for such interactions. At the same time, interactions with people can involve a very wide range of different behaviors that may all be more or less appropriate. In other words; being social can pose a very complex, highly dynamic interactive challenge.

Since many robots have 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**). Social positioning plays a role during many of the different phases that could be entailed in an interaction, such as approach, converse, and retreat. As such, it has a very direct relevance to several settings, for example (semi-autonomous) telepresence robots[1] and robots giving people information at a mall, airport[2], or museum. As these examples show, many of these settings can involve (small) groups of users.

Another aspect of social positioning are two interaction dynamics that play an important role. First, there are the **temporal dynamics**; interactions take place over time, which implies that (social positioning) requirements can change and that movements rather than static positions should be taken into account [1]. Second, there are **social dynamics**; participants in an interaction respond and adapt to each other [2]. 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. Obviously, these dynamics become more complex as more entities become involved in the interaction (e.g., a robot interacting with a group).

This paper introduces a study aimed at collecting quantitative and qualitative data that can be used to inductively generate hypotheses on suitable social positioning behavior for a robot interacting with a group that can move and respond dynamically (see Figure 1). In particular, the study looks at the interactions involved in approaching a group, conversing with it, and retreating from it – since those elements are common in many of the previously mentioned settings in which mobile robots are used.

There are two reasons this aim is challenging. First, to allow for the dynamics, there should not be many constraints on the behavior of the participants, which reduces the control we have over the study. Second, it still is a challenge to implement temporally and the socially dynamic behaviors. We have here resolved this by using a telepresence robot controlled by one of the participants, though it should be noted that this may limit the generalizability of the found hypotheses to other kinds of robots.

* The work described in this paper has partly been supported by the European Commission under contract numbers FP7-ICT-611153 (TERESA), and FP7-ICT-600877 (SPENCER).

All authors are with the department of Human Media Interaction, Faculty of Electrical Engineering, Mathematics and Computer Science, University of Twente, 7500 AE Enschede, The Netherlands. {j.h.vroon, m.p.joosse, m.lohse, j.kim, k.p.truong, g.englebienne, d.k.j.heylen, v.evers}@utwente.nl and j.kolkmeier@student.utwente.nl

[1]www.teresaproject.eu

[2]www.spencer.eu

The contribution of this paper is twofold. First, we will introduce the method of the study in more detail as well as the data collected with it. Second, we will present social positioning patterns we found in the data, from which hypotheses can be derived. Note that though the data allows for a more extensive analysis, we consider this to be out of the scope (and size) of this paper.

II. THEORETICAL BACKGROUND

Much of the work in social positioning for robots is based on two theories from sociology on social positioning in humans. Proxemics, first introduced by Hall [3], focuses on the distances people keep to each other. F-formations, as introduced by Kendon [1], describe the different spatial arrangements people can use in social interactions. It has been shown that, as these theories would predict, many different social situations can be distinguished based on only position and orientation information (e.g. [4], [5]).

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 II-A). We will further discuss work that focuses on the different interaction dynamics (Section II-B).

A. Social positioning in robotics

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 [6], with different properties of the robot [7], with relation to the background of the participants [8], and for different cultures [9]. 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. To our knowledge, there is no work on social positioning for robots specific for interaction with small groups.

Social positioning has 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 [10] or to inform design decisions [11]. Of particular relevance in the context of this paper is the research of Kristoffersson et al. [12] and van Oosterhout & Visser [13], since they actively observed the displayed behaviors. Both used manual annotations of visual data (video/photo), to investigate relevant patterns in the behavior. Van Oosterhout & Visser [13] found that people generally position themselves within Hall's personal space zone. Kristoffersson et al. [12] 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 [1]. Actively observing the behaviors used by participants controlling a robot thus seems a fruitful approach to investigate suitable social positioning of (telepresence) robots.

B. Temporal and social dynamics of interaction

One factor that makes it hard to study human robot interaction is that it is a dynamic process. Or, as Hüttenrauch et al. [14] 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" *([14], p.5058)*.

There are two sides to these interaction dynamics. First, there are the temporal dynamics of movements and changing requirements. There is a limited set of papers that explicitly look into the temporal dynamics of social positioning for interactions between people and a robot [12], [14], [15], [16]. Second, there are the social dynamics of people adapting to a robot and other people, as well as expecting adaptation. 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 [17], to signal approachability with a group of virtual agents [18], or to influence the formation of people interacting with a robot [19]. As evidenced by these papers, the temporal and social dynamics are relevant and can have a strong influence on what happens in the interaction.

III. METHOD

The aim of this study was to collect data that can be used to generate hypotheses on (dynamic) features that could be taken into account when designing social positioning robot behavior for interaction with a small group. To achieve this, we created a setting in which groups of four people would go through several cycles of approach/converse/retreat behavior. One of the participants was present through a telepresence robot (the **Visitor**), and used the robot to interact with the rest of the group (the **Interaction Targets**).

One of the challenges to our aim was that to allow for the dynamics to arise we wanted to leave our participants as free as possible. At the same time, we wanted to keep the different cycles comparable, to make the comparison of the acquired quantitative data easier. Therefore, we created a somewhat controlled setting where we 'reset' the position of the Interaction Targets *between* the cycles, while allowing them to move *during* the cycles.

Another challenge was to automatically generate robot behaviors that are sufficiently dynamic and appropriate. As discussed in the introduction and theoretical background, we have here resolved this by having one participant control the telepresence robot used in the study.

A. 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 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.

B. Procedure

The study took place in a controlled laboratory setting. For the study, we used a Giraff (www.giraff.org) telepresence robot equipped with the hardware required for the data collection (Section III-C.1). 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 a 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 III-C.1) 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 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 III-C.2.a). 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. 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; a circular formation, with every participant occupying an equal amount of space, and a semi-circular formation featuring an open space [18]. Groups were randomly assigned to one of

[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.

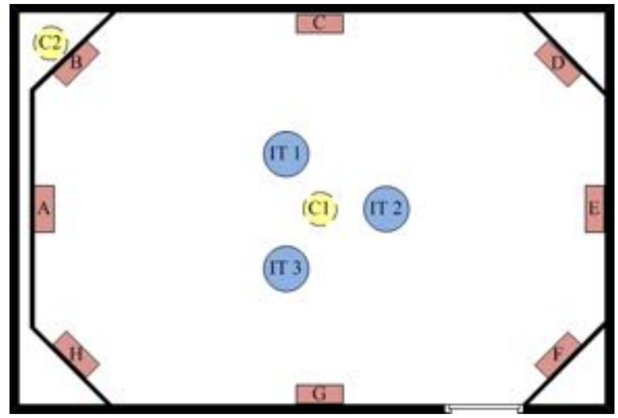


Fig. 2. Overview of the interaction area (approximately 6 by 4 meters). On the circle in the middle the positions of the Interaction Targets are indicated (IT1, IT2, IT3), these were projected using a projector mounted to the ceiling, but only in between the approach/converse/retreat cycles. The rectangles near the border of the interaction area indicate the positions of the markers A-H. C1 and C2 indicate the positions of the cameras.

the formations. This was not a condition, as it would have been in deductive research, but instead intended to cover some of the variations that might naturally occur.

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 (See Section III-C.2.b).

C. 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 (III-C.1) and subjective data with questionnaires (III-C.2). The tracking and questionnaire data is available from the first author upon request.

1) 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 (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 to not get too close to the edges of the interaction area. In the analysis here presented, we will take the marker on the cap worn by the Interaction Targets to represent their position.

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 sensor.

Two cameras recorded 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.

2) 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 [20]).

b) Post-experiment questionnaire: The post-experiment questionnaire consisted of 49 items. Among others we measured co-presence and attentional engagement [21]. Furthermore we measured the participants' attitude towards robots [22] and workload [20].

D. 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%).

E. 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.

IV. FINDINGS

We will present first findings from the (quantified) observations (IV-A) and the investigation of the relations between features of the dynamics of the motion patterns and the ratings of the Interaction Targets (IV-B). A more extensive analysis is out of the scope of this paper.

[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.

A. Observed patterns of behavior

Under the assumption that the participants all tried to display suitable social positioning, suitable behaviors would likely be more common. Thus, patterns that are commonly observed in the interactions can be generalized to hypotheses for suitable behavior with inductive reasoning. We will here introduce some of 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 were.

1) Approach: During the Approach, most Visitors drove the robot towards the Interaction Targets (Table I-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.

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 I-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 I-5).

2) 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 I-6,7). When the Visitor did move, these movements were rotations that increased the visibility of the Interaction Targets.

3) 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 observed 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.

B. Relating motion patterns with ratings

The ratings provided 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.

	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 I

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

There were large individual differences in 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 (IV-B.1). Then we will discuss more quantified ways for looking at these findings (IV-B.2).

1) 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 highest rated Approaches we additionally observed that the Visitor stopped 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 lowest rated 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 highest rated 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 highest and the lowest rated 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 highest rated Converses, but in the ten lowest rated the robot was usually far away from the group center or relatively close to at least one of the Interaction Targets.

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

These are only first results, to illustrate how this data could be used. Though it does not fit the size and scope of this paper, further analysis, for example based on mutual information, will likely reveal even more 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.

V. CONCLUSIONS AND DISCUSSION

In this paper, 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, all these patterns can be used as hypotheses for more general settings. These could be settings with a different task, different people, and a different robot. 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 a telepresence robot, there is no guarantee they

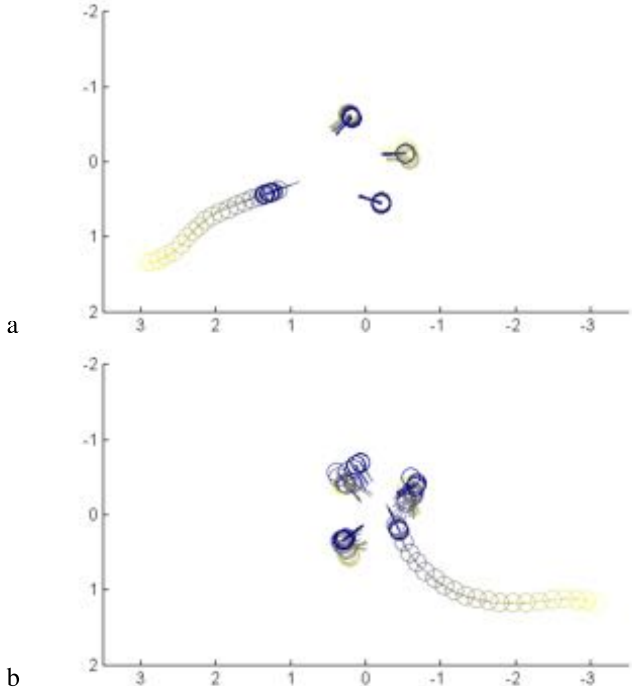


Fig. 3. Representation of head tracking data from two Approaches, one with a high average normalized rating (a) and one with a low average normalized rating (b). The circles with lines show the positions and orientations of the Visitor and Interaction targets in the interaction area. Indicators near 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.

will translate to other types of robots. It is for this limitation of inductive reasoning that it is important to realize that our findings are hypotheses only.

To demonstrate the use of our method, we have used this inductive reasoning to generate a variety of hypotheses on social positioning. These include, in line with what proxemics would pose [3], 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 a quantitative inductive study to robotics research and used it to generate various hypotheses that can guide the design of social positioning robot behavior. Our findings furthermore show that the temporal and social dynamics can play a role in the interaction between a (telepresence) robot and a group.

REFERENCES

[1] A. Kendon, *Conducting interaction: Patterns of behavior in focused encounters*. CUP Archive, 1990, vol. 7.

[2] A. Pentland, "Social dynamics: The voice of power and influence," MIT Media Lab, Tech. Rep. 579, 2004.

[3] E. T. Hall, *The Hidden Dimension*. Anchor Books New York, 1966.

[4] G. Groh, A. Lehmann, J. Reimers, M. R. Frieß, and L. Schwarz, "Detecting social situations from interaction geometry," in *Social Computing (SocialCom), 2010 IEEE Second International Conference on*. IEEE, 2010, pp. 1–8.

[5] B. Lau, K. O. Arras, and W. Burgard, "Multi-model hypothesis group tracking and group size estimation," *International Journal of Social Robotics*, vol. 2, no. 1, pp. 19–30, 2010.

[6] E. Torta, R. H. Cuijpers, J. F. Juola, and D. van der Pol, "Design of robust robotic proxemic behaviour," in *Social Robotics*. Springer, 2011, pp. 21–30.

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

[8] L. Takayama and C. Pantofaru, "Influences on proxemic behaviors in human-robot interaction," in *Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on*. IEEE, 2009, pp. 5495–5502.

[9] M. P. Joosse, R. W. Poppe, M. Lohse, and V. Evers, "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*. ACM, 2014, pp. 121–130.

[10] J. M. Beer and L. Takayama, "Mobile remote presence systems for older adults: acceptance, benefits, and concerns," in *Proceedings of the 6th international conference on Human-robot interaction*. ACM, 2011, pp. 19–26.

[11] J. Jung, T. Kanda, and M.-S. Kim, "Guidelines for contextual motion design of a humanoid robot," *International Journal of Social Robotics*, vol. 5, no. 2, pp. 153–169, 2013.

[12] A. Kristoffersson, K. S. Eklundh, and A. Loutfi, "Measuring the quality of interaction in mobile robotic telepresence: a pilots perspective," *International Journal of Social Robotics*, vol. 5, no. 1, pp. 89–101, 2013.

[13] T. van Oosterhout and A. Visser, "A visual method for robot proxemics measurements," in *Proceedings of Metrics for Human-Robot Interaction: A Workshop at the Third ACM/IEEE International Conference on Human-Robot Interaction (HRI 2008)*. Citeseer, 2008, pp. 61–68.

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

[15] R. Mead, A. Atrash, and M. J. Matarić, "Automated proxemic feature extraction and behavior recognition: Applications in human-robot interaction," *International Journal of Social Robotics*, vol. 5, no. 3, pp. 367–378, 2013.

[16] G. Hoffman and W. Ju, "Designing robots with movement in mind," *Journal of Human-Robot Interaction*, vol. 3, no. 1, pp. 89–122, 2014.

[17] T. Kruse, A. Kirsch, E. A. Sisbot, and R. Alami, "Exploiting human cooperation in human-centered robot navigation," in *RO-MAN, 2010 IEEE*. IEEE, 2010, pp. 192–197.

[18] M. Rehm, E. André, and M. Nischt, "Lets come together–social navigation behaviors of virtual and real humans," in *Intelligent Technologies for Interactive Entertainment*. Springer, 2005, pp. 124–133.

[19] H. Kuzuoka, Y. Suzuki, J. Yamashita, and K. Yamazaki, "Reconfiguring spatial formation arrangement by robot body orientation," in *Proceedings of the 5th ACM/IEEE international conference on Human-robot interaction*. IEEE Press, 2010, pp. 285–292.

[20] S. G. Hart, "Nasa-task load index (nasa-tlx); 20 years later," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 50, no. 9. Sage Publications, 2006, pp. 904–908.

[21] F. Biocca and C. Harms, "Guide to the networked minds social presence inventory v. 1.2," 2003.

[22] M. Heerink, B. Kröse, B. Wielinga, and V. Evers, "Studying the acceptance of a robotic agent by elderly users," *International Journal of Assistive Robotics and Mechatronics*, vol. 7, no. 3, pp. 33–43, 2006.

4th International Symposium on New Frontiers in Human-Robot Interaction
21–22 April 2015, Canterbury, UK
M. Salem, A. Weiss, P. Baxter, K. Dautenhahn (Eds.)

Robots Guiding Small Groups: The Effect of Appearance Change on the User Experience

Michiel Joosse, Robin Knuppe, Geert Pingen, Rutger Varkevisser, Josip Vukoja, Manja Lohse and Vanessa Evers [1]

Abstract. In this paper we present an exploratory user study in which a robot guided small groups of two to three people. We manipulated the appearance of the robot in terms of the position of a tablet providing information (facing the group that was guided or the walking direction) and the type of information displayed (eyes or route information). Our results indicate that users preferred eyes on a display that faced the walking direction and route information on a display that faced them. The study gave us strong indication to believe that people are not in favor of eyes looking at them during the guiding.

1 Introduction

Social robots are designed to interact with humans in human environments in a socially meaningful way [3]. As a logical consequence, the design of robots often includes human-like features, e.g., heads or arms in order to generate social responses. It has been found that by using such anthropomorphic cues, people automatically have expectations of the robot's behavior [4].

However, the capabilities of robots differ from those of humans which allows them to use the anthropomorphic cues in different ways. For example, robot eyes can face the user while walking because the robot has other means (e.g., laser range finders) to detect the way to go. Thus, robots can walk backward. As eye contact has been shown to impact our image of others, and whether positive or negative, this being a sign of potential social interaction [6], robots facing users while guiding might actually be beneficial. On the other hand, literature indicates that people use a combination of head and eye movement to non-verbally indicate their direction [1] and users might expect robots to do the same.

Robots can also use non-anthropomorphic cues in different ways than humans, e.g. in the guiding context they can display route information rather than eyes. Related work found that visitors in historic places prefer a guide, as they would not have to worry about the route, or carry a map [2]. Therefore this could be beneficial for robots as well.

In the FP7-project SPENCER[2] we aim at developing a guide robot for a public place (airport) which will have a head and a screen. In this context, the questions arise which direction the head and screen should face when guiding a small group and what content should be displayed on the screen.

In related work, Shiomi et al. [5] conducted an experiment with the Robovie robot that drove either forward or backward while guiding participants in a mall (over a short distance). The overall finding in this experiment was that more bystanders joined when the robot moved backwards compared with frontwards, and that more people were inclined to follow the robot the entire time when moving backwards. In our work we are not so much interested in attracting people, but more in guiding people over a longer distance. Thus the question we pose here is how these design decisions impact the user experience in the process of guiding.

In this paper we present an exploratory study, in which we asked participants to follow a guide robot through a public lab space. This robot was equipped with a tablet (facing forwards, or facing the user) providing information to the participants. We were specifically interested in finding out which combination of tablet direction and type of information provided (eyes or route information) would yield the most positive user experience.

2 Method

In order to answer our research question, we designed an exploratory user study in which small groups of two to three participants were given a short guided tour by a robot.

2.1 Robot platform

For this study we attached a shell on top of a remote-controlled Robotino robot platform[3]. The height of the robot was 170cm and it drove at a speed of approximately 0.7 m/s. For purposes of this exploratory study, it was not deemed necessary to have the robot drive the path autonomously. Furthermore, the location of obstacles in the DesignLab changed from time to time (e.g. couches, chairs). As we were primarily interested in user experience ratings, the robot was remotely operated by an experimenter. Participants were not made aware of this before participating in the experiment.

2.2 Manipulations

We manipulated the direction of the tablet mounted on top of the robot and the information displayed on the tablet (Figure 1 and Table 1). In conditions A (Figure 1a) and B (Figure 1c) a set of blinking eyes was displayed on the tablet either facing the participants or the walking direction. In condition C we programmed the tablet to display route information, i.e., the remaining distance to the target (Figure 1e). A condition having the tablet mounted on the front of the robot, while displaying route information was deemed unnecessary as this would neither provide information for the participants following the robot, nor for other people present in the laboratory.

[1] Human Media Interaction group, University of Twente, the Netherlands, email: {r.a.knuppe, g.l.j.pingen, r.a.varkevisser, j.vukoja}@student.utwente.nl, {m.p.joosse, m.lohse, v.evers}@utwente.nl

[2] http://www.spencer.eu

[3] http://www.festo-didactic.com/int-en/learning-systems/education-and-research-robots-robotino/

(a) Condition A front (b) Condition A back (c) Condition B front (d) Condition B back (e) Condition C front (f) Condition C back

Figure 1: The appearance of the robot in the three conditions, showing the front and back side of the robot

Table 1: Overview of study conditions and number of participants

Condition	A	B	C
Tablet direction	Front	Back	Back
Tablet display	Eyes	Eyes	Time to destination
N	9	8	8
Group distribution	3x 3-person	1x 2-person 2x 3-person	1x 2-person 2x 3-person

2.3 Measures

In the post-experiment questionnaire user experience was assessed using a variety of measures.

All questions (except demographic- and open questions) were formulated as 5-point Likert-scaled items. General experience was assessed with eleven questions measuring among others if participants trusted that the robot knew where it was going, if it was clear where the robot was going and whether or not the robot was helpful in guiding someone. In this set of questions also the speed of the robot and volume of the audio messages were evaluated.

Five questions related to the physical appearance assessed the design, and specifically the height of the robot. Usability questions included questions related to users' expectancies of system capabilities and whether or not they were satisfied with the overall performance of the robot. Depending on the condition, this section included 5 (condition A), 6 (condition B), or 7 (condition C) questions.

Eight questions were included related to demographic information (age, gender, educational background) and familiarity with robots, social robots, and the premises where the test was conducted. A control question about the position of the tablet was included, and finally, we were interested in knowing whether or not the instructions provided were clear. Overall, this resulted in 30-32 questions

2.4 Procedure

Small groups of participants were recruited to participate in a guided tour of the DesignLab, a recently-opened lab of the University of Twente. Participants were given a briefing, after which they were given a tour of about five minutes through the lab. Participants were requested to follow the robot. No specific instructions were provided regarding the distance they should keep to the robot (Figure 4). The tour went past two points of interest (Figure 2, point B and C) where the robot provided a brief statement about the purpose using a text-to-speech engine. For example, when arriving at waypoint A, participants would see a tray with kinetic sand, and the robot would state that ”The kinetic sand is made up of 98 percent sand, and 2 percent polyminethyl siloxane which gives it its elastic properties.”

Afterwards the robot returned to the starting position where participants were requested to fill out the post-experiment questionnaire (Figure 2 point A). Following debriefing, participants were provided some candy as reward for their participation.

2.5 Participants

A total of 25 participants (14 males, 11 females) participated in the user study, with ages ranging from 17 to 40 (M=23.76, sd=5.93). All participants were students and staff from the University of Twente, primarily of Dutch (68%), German (8%) and Greek (8%) nationality. Participants had average experience with robots in general (M=2.84, sd=.90) and little experience with social robots (M=2.12, sd=1.09).

2.6 Data analysis

We calalculated means for all items. To compare between conditions, the data were first tested for normality. In case of normally distributed data, we report ANOVA's and T-tests in the results section, otherwise Kruskal-Wallis and post-hoc Mann-Whitney tests are reported.

3 Results

Overall, participants indicated they were quite satisfied with the robot: they believed the robot was helpful (M=4.47, sd=0.78), it

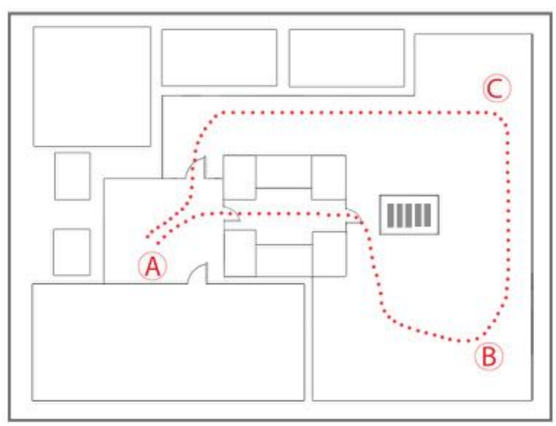


Figure 2: Layout of the laboratory showing start/end position (A) and two points of interest (B and C)

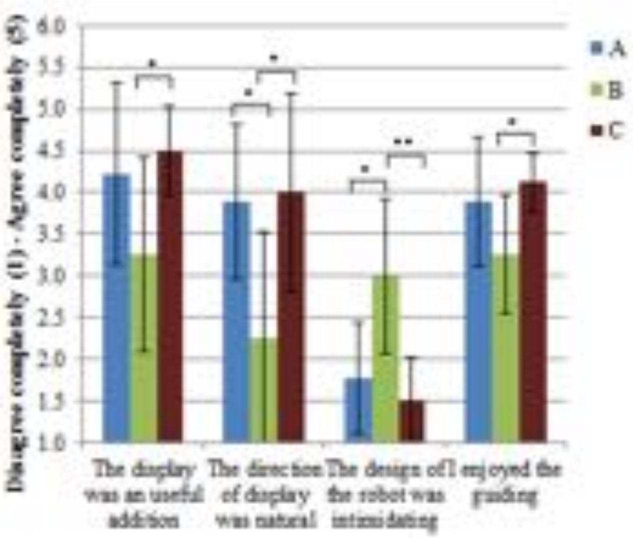


Figure 3: User experience ratings in the conditions; * indicates significance at the 0.05 level, ** at the 0.01 level

moved at a comfortable speed (M=3.12, sd=1.37), and participants trusted that the robot knew where it was going to (M=4.47, sd=0.78). These ratings did not differ significantly between conditions. Participants were moderately positive about the usability of the system: they felt comfortable using it (M=3.67, sd=1.05) and were satisfied by its performance (M=3.56, sd=0.77). No main effects or correlations were found including gender, age, robot experience and/or educational background.

Between conditions, Kruskal-Wallis tests indicated there were significant differences which were mostly due to the location of the tablet, thus between conditions A and C, versus condition B where the tablet was mounted on the front of the robot.

Post-hoc Mann-Whitney's indicated participants felt the direction of the screen was more appropriate in condition A (M=3.89, sd=.928) compared with B (M=2.25, sd=1.28), U=11.5; Z=-2.459, $p<0.05$. A similar effect was found between conditions B and C (M=4.0, sd=1.20), U=10.0, Z=-2.36, $p<0.05$ Furthermore, the design in condition B was more intimidating (M=3.00, sd=.97) compared with condition A (M=1.78, sd=.68), U=11.5, Z=-2.51, $p<0.05$ and condition C (M=1.50, sd=.54), U=6.00, Z=-2.885, $p<0.01$. Participants in condition C enjoyed the guiding more (M=4.13, sd=.35) compared with those in condition B (M=3.25, sd=.71), U=10.5, Z=-2.62, $p<0.05$.

With respect to the robot's appearance, participants felt that the body design matches the robot's function (M=2.71, sd=0.94). One of the interesting findings was that participants indicated the height was appropriate (M=4.21, sd=0.82). Informal sessions with participants indicated the robot would be too tall for a guiding robot, but in the end this was not the case. One of the reasons for this could be that participants' own average height was 177cm (sd=8.5cm), thus, most of them being taller than the robot.

4 Discussion & Conclusion

In this paper we presented an exploratory study into the effect of a robot's physical appearance on usability and user experience. Small groups of people were provided a short tour by a guide robot. Our results indicate that the location of the screen can be either forward or backward, depending on the information displayed. In the case of eyes facing participants, our results showed that this was considered to be very unnatural and intimidating. On the other hand, when the tablet faced participants and route information was provided this was again evaluated as more useful. This might seem to be in contrast with the results of Shiomi et al. [5] who found that eyes facing participants are more effective to attract bystanders. However, we think this could be explained because in our setup the participants had already been introduced to the robot and asked to follow it.

Figure 4: A small group of participants being guided by the robot

Neither gender, age or experience with robots influenced the evaluation of the robots significantly, which could be due to small sample size.

Our future work will include a more interactive setup (e.g. provide participants some choices) during the tour. A second area of interest would be robot speed, and to investigate whether or not the speed of a guiding robot could be slower when guiding small groups compared with individual people. To conclude: the appearance of a guide robot can greatly influence user experience, something subtle as two eyes facing participants significantly decreases a robot's evaluation. Hence, more research is needed to even better understand how to design acceptable guide robots.

ACKNOWLEDGEMENTS

This research has been partly supported by the European Commission under contract number FP7-ICT-600877 (SPENCER).

REFERENCES

[1] Mark A Hollands, AE Patla, and JN Vickers, 'look where youre going!: gaze behaviour associated with maintaining and changing the direction of locomotion', *Experimental Brain Research*, **143**(2), 221–230, (2002).

[2] Daphne E Karreman, Elisabeth MAG van Dijk, and Vanessa Evers, 'Using the visitor experiences for mapping the possibilities of implementing a robotic guide in outdoor sites', in *RO-MAN, 2012 IEEE*, pp. 1059–1065. IEEE, (2012).

[3] Kwan Min Lee, Wei Peng, Seung-A Jin, and Chang Yan, 'Can robots manifest personality?: An empirical test of personality recognition, social responses, and social presence in human–robot interaction', *Journal of communication*, **56**(4), 754–772, (2006).

[4] Manja Lohse, 'The role of expectations and situations in human-robot interaction', *New Frontiers in Human-Robot Interaction*, 35–56, (2011).

[5] Masahiro Shiomi, Takayuki Kanda, Hiroshi Ishiguro, and Norihiro Hagita, 'A larger audience, please!: encouraging people to listen to a guide robot', in *Proceedings of the 5th ACM/IEEE international conference on Human-robot interaction*, pp. 31–38. IEEE Press, (2010).

[6] Michael von Grünau and Christina Anston, 'The detection of gaze direction: A stare-in-the-crowd effect', *Perception*, **24**(11), 1297–1313, (1995).

Short Duration Robot Interaction at an Airport: Challenges from a Socio-Psychological Point of View

Michiel Joosse, Manja Lohse, and Vanessa Evers

Human Media Interaction, University of Twente, the Netherlands
m.p.joosse, m.lohse, v.evers@utwente.nl

1 Introduction

This extended abstract concerns the FP7-project Spencer[1]. As part of the Spencer project, a demonstrator robot will be developed which provide services to passengers at a major European airport. Example services include (1) guiding transfer passengers from their arrival gate to the so-called Schengen barrier, and (2) assisting in the transfer process by printing boarding passes. The goal of the robot is to make sure that passengers will make their connecting flight, with our own focus being on the human-robot interaction. In the following, we describe a sample use case of the project scenario. Based on this we identify possible challenges that are of interest with respect to interactive robots in public spaces.

2 Use case

The Spencer project aims to develop and deploy a demonstrator service robot which can provide assistance to transferring passengers at a major airport in Europe. The industrial partner has about 25.000 transfer passengers daily. A large portion of these passengers transfer from (intercontinental) non-Schengen flights to (European) Schengen flights, requiring passengers to go through a passport control before arriving at their next departure gate. This process costs time, and is one of the major sources why passenger miss their connecting flights.

The Spencer robot is envisioned to collect a group of transfer passengers with a minimal connection time at the gate (for example they could make their connection if they hurry and go to the fast-track Schengen lane). The Spencer robot will guide them towards the fast-track Schengen lane, after which they can proceed to their departure gate. An average group of transfer passengers constitutes between 20-40 people, who do not necessarily know each other; their only common denominator is their next flight.

Scientific challenges for the Spencer project include (among others) socially intelligent navigation and the detection of groups of people and relations within these groups. The scientific challenge we are focusing on, is the *evaluation and design of (spatial) robot behaviors that are experienced as (socially) normative.*

[1] http://www.spencer.eu

3 HRI challenges for the Spencer robot

The use case as described above contains both technical and scientific challenges. In this section, we will focus on what we believe as being the most important challenges for the interaction between passengers and a robot.

3.1 Normative behavior: it is about intention recognition

Independent of the ways by which the robot conveys its intentions, the robot should be perceived as behaving in a normative way. Thus, the behavior of the robot should conform to the social norms expected by the *current* passengers. Examples of these normative behaviors could include adjusting the speed to the group, and giving way to people approaching from the right. While the implementation of these issues could be considered technical ones, we believe the identification of the norms is a socio-psychological problem.

We believe that because - for most people - flying is not considered to be an everyday activity; many people consider it to be hectic, and are sometime unsure of what (not) to do or where to go. This makes it especially important that the messages a robot transmit, for instance those which convey its movement, are clear and predictable.

At airports identifying normative behavior is particularly complicated in part because the robot will have to deal with people with different cultural backgrounds. These might even form part of one group that has to be guided at the same time. Also the fact that we have to deal with groups as such is a challenge for behavior planning and other technical requirements such as robust spoken language processing and person tracking.

Thus, there are two distinct different normative behaviors we consider in this extended abstract. One the one hand we argue that the movement of the robot should be legible and conveying towards the passengers. On the other hand, the robot has to behave in a normative way in the sense that it abides with the (un)written conventions of pedestrian traffic.

3.2 Communication Modalities

To address the issue of legible and conveying movement, one can think of different modalities which could be useful to communicate intent. Whereas humans can use non-verbal communications to exchange social signals when approaching one another [2], robots are not (yet) capable of this. We propose to evaluate two different communication modalities for the robot, each having pro's and con's in the context of an airport.

Speech or sound in general, could be one of these modalities. Due to the multicultural mix of passengers these messages would ideally be universal. A solution can be to implement a text-to-speech engine in the robot, or a noise-like level as described in [1]. Since an airport is a noisy environment, the robot has to repeat the messages.

A graphical interface could also be used to convey movement intentions, for instance a screen, indicating the robot's speed or acceleration. This would be limited in that it can only supply information to those who can see the screen; people who are moving behind, or next to, the robot.

Both modalities could be used to communicate intent to passengers. We intend to test both modalities synchronous and asynchronous in order for the robot to communicate as effective as possible.

3.3 Research approach

The Media Equation states that people treat computers, and related media, as if they were people [3]. Based upon this work of Nass and colleagues, our approach is to first identify what people do, implement similar behaviors on a robot, and evaluate whether human norms hold for human-robot interaction.

We do not expect that human normative behavior will unequivocally carry over to normative robot behavior, however, we will use it as a starting point.

Based upon a literature review and a contextual analysis (systematic observation of what really happens), we will design and implement normative behaviors for a robot. These behaviors will first be tested in lab studies, followed by experiments at the site of the industrial partner; the airport in order to get an idea of the experiences of the passengers.

To get insight in the experiences of passengers at the airport, we can employ several methods for user studies. Examples include self-reported questionnaires, coding of video data and analysis of one's galvanic skin response.

For our experiments, we will primarily collect video data, as well as subjective questionnaires or -interviews. Objective video data makes it easy to capture certain behavioral responses from multiple people in a short time. However, legal and organizational issues (such as privacy and security) could hinder this method when used outside the lab. Interviews and questionnaires should be able to capture the required data in only a few questions, given that passengers will be likely be in a hurry. These languages should be unambiguous for passengers with different cultures; this raises the question if the language should be native, or universal (read: English). Different languages would require multiple iterations of translation and back translation to ensure the questions truly ask the same.

In light of the issues described above we may have to rethink our data collection methods. This also holds for the data the robot collects for its own perception. Prior to evaluation in a real-world setting, behavior will be evaluated in a more controlled setting. This could be in a lab setting as for example in [1, 4], but also by other ways, for instance by using videos of a robot interacting with people [5].

4 Conclusion

In this extended abstract we have described challenges from a socio-psychological point of view when a robot interacts with users at an airport. We argue the it is especially important to convey the robot's intentions toward users in an appropriate manner with respect to social norms while taking into account constraining environmental factors (such as noise levels). In part due to their safety-critical nature, airports in general cause specific challenges with regards to data collection practices.

5 Acknowledgements

This work has partly been supported by the European Commission under contract number FP7-ICT-600877 (SPENCER)

References

1. M. Lohse, N. van Berkel, E. M. A. G. van Dijk, M. P. Joosse, D. E. Karreman, and V. Evers. The influence of approach speed and functional noise on the perception of a robot. In *Proceedings of the 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2013.
2. R. W. Picard. *Affective computing*. MIT press, 2000.
3. B. Reeves and C. Nass. *How people treat computers, television, and new media like real people and places*. CSLI Publications and Cambridge University Press, 1996.
4. A. Sardar, M. Joosse, A. Weiss, and V. Evers. Don't stand so close to me: users' attitudinal and behavioral responses to personal space invasion by robots. In *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*, pages 229–230. ACM, 2012.
5. M. L. Walters, M. Lohse, M. Hanheide, B. Wrede, D. S. Syrdal, K. L. Koay, A. Green, H. Hüttenrauch, K. Dautenhahn, G. Sagerer, et al. Evaluating the robot personality and verbal behavior of domestic robots using video-based studies. *Advanced Robotics*, 25(18):2233–2254, 2011.

Lost in Proxemics: Spatial Behavior for Cross-Cultural HRI

Michiel Joosse
Human Media Interaction
University of Twente
Enschede, the Netherlands
m.p.joosse@utwente.nl

Manja Lohse
Human Media Interaction
University of Twente
Enschede, the Netherlands
m.lohse@utwente.nl

Vanessa Evers
Human Media Interaction
University of Twente
Enschede, the Netherlands
v.evers@utwente.nl

ABSTRACT
Socio-psychological research hints to the fact that people from different cultures have different preferences with respect to proxemics. Thus, what might be considered normal for one person, could be a violation of a norm for another person. If cultural background influences spatial behaviors, a logical follow-up question would be if a robot should be equipped with different sets of normative motion behaviors for guiding people. In this paper, we provide an overview of research into cultural differences in proxemics and human-robot social norms. We will address culture not at a national level (i.e. Dutch vs. German national culture), but instead at a clustered, supranational level based upon work by [13]. We conclude with foreseen challenges and solutions for analyzing the appropriateness of HRI behaviors in the context of different cultures.

Categories and Subject Descriptors
J.4 [**Computer Applications**]: Social and Behavioral Sciences

General Terms
Human Factors

Keywords
Human-Robot Interaction, Cultural Differences, Public Space, Proxemics.

1. INTRODUCTION
The phrase "as robots start entering our life" might be an understatement, especially in this field of research. It is not so much a question of if, but more when, and how social robots will enter our daily lives. Over a decade ago, Fong et al. [10] provided an overview of the then-current state of robotics, and distinguished six major application areas. In this paper we focus on culture-aware robotics within the service application field, and specifically short-term public interaction robots.

As part of the EU FP7-project Spencer[1], we intend to elicit and evaluate socially normative motion behaviors for a robot which navigates through a crowded environment. The crowded environment is an international airport, where the robot will guide delayed, culturally diverse, passengers from their intercontinental flight to their connecting continental (European) flight. We do not attempt to trivialize the underlying technical challenges to navigate such an environment in an effective and safe way, but we will focus on the aspect of cultural normative behavior.

Research has pointed to evidence suggesting that people explain machine behavior in terms of human behavior. People anthropomorphize, or have "the tendency to imbue the real or imagined behavior of nonhuman agents with humanlike characteristics, motivations, intentions, or emotions" [9]. Examples include a preference for a specific (static) robot head, given a certain task [12], or the perception of cameras as eyes.

In this paper, we will first provide a short overview of human social norms in general, and cross-cultural social norms research specifically (Section 2). We will then discuss human-robot social norms (Section 3), and discuss challenges for cross-cultural human-robot interaction (HRI) research (Section 4).

2. ON SOCIAL NORMS
Social norms are unwritten norms, sustained by feelings of embarrassment and guilt when violated [8], the disapproval of other people, and social sanctions [32]. These norms are situational dependent; norms governing appropriate conduct during a soccer game differ from those which govern a funeral [1]. The definition of social norms we use in this paper is "Rules and standards that are understood by members of a group and that guide and/or constrain social behavior without the force of laws" [6].

Examples of research into human adherence to social norms include series of experiments by Cialdini et al. and Keizer et al. [24]. The norm researched was the social norm of littering in public space. The main findings include that a) people tend to litter more in an already-littered environment, b) littering increased when the norm was made salient, and c) that the violation of one norm (a littered environment) makes violation of others norms more likely – the latter also called a cross-inhibition effect. Similar results have been found for other social norms, such as the norm of "being silent in the library" [1].

While above research provides insightful results, these are not necessarily the social norms that are automatically relevant or applicable for the Spencer project. A norm that is relevant, is the norm concerning the adherence to one's personal space. Personal space is one of the four proxemics zones defined by Hall [14], and refers to the semi-circular shaped protective bubble people keep around themselves that cannot be invaded without causing some sort of discomfort. In his book, *the Hidden Dimension* [14], Hall indicated the size of one's personal space to be around 45 cm., this being applicable to Northern Americans, and indicating this size to be different for, for instance, Chinese people.

2.1 Personal space is dependent on culture
Several experiments showed that people with different cultural backgrounds have a different sized personal space zone. One

HRI'14 Workshop on Culture Aware Robotics, 3 March, 2014, Bielefeld, Germany.

[1] http://www.spencer.eu

example dimension to explain cultural differences is the dimension, or maybe division, of cultures into "contact" and "noncontact" cultures. Based upon observations, Hall [14] noted that people from noncontact cultures (Northern European, Northern American countries) maintain a larger personal space compared with their counterparts from contact cultures (Southern European, Southern American, Arab countries).

In one of the experiments, 105 students from three different ethnical groups (Japanese, American and Venezuelan) had a (seated) five-minute conversation with a same-sex, same-nationality confederate [34]. Either in their native language, or in English. They found, when speaking English, participants from the non-contact culture (Japan) sat further apart from each other compared to the contact culture (Venezuela). Within the ethnical groups male participants sat further apart than female participants. Furthermore, when speaking their native language, contact culture participants sat closer together.

Other experiments looking at cross-cultural proxemics distances include the work by Little [27], who used the placement of dolls to infer at which distance people from either the U.S., Sweden, Scotland, Italy and Greece would place people in 19 different social situations, and found similar differences between countries.

Likewise, Høgh-Olesen [19] looked at proxemic differences between cultures, but also at similarities. Based upon the work of Pike [31], he differentiated between two terms; proxethics and proxemics. Proxethics refers to the behaviors and dynamics which are shared by humans – thus being universal. In contrast, proxemics looks at the differences [19]. Høgh-Olesen found six cross-cultural proxethics conventions within six cultures (Greenland, Finland, Denmark, Italy, India and Cameroon). For instance, people leave more room between two strangers compared with one stranger, and the personal space is smaller in social spaces (a café) as compared with non-social spaces (library).

With the knowledge that social norms exist for humans, and these norms can be different for people with different cultural background, a question arises what culture is, and what research has been conducted with regards to cross-cultural human-robot interaction. However, before discussing this in Section 3, we will take a look at the current research in HRI with respect to social norms.

2.2 Human-Robot Social Norms

HRI work related to social norms has mostly been concerned with physical norms, such as approaching someone. Work by Walters [38] focused primarily on the identification of the size of humans' personal space bubble. Takayama & Pantofaru [35] looked at the effect of robot gaze on the approach distance humans keep. They found that when the robot would gaze towards one's legs, men and woman would approach equally close (M=0.28 / 0.30m). However, when the robot gazed towards the participants face, woman maintained a significant larger personal space (M=0.30 m.) compared with men (M=0.24m).

Related to personal space, Dautenhahn et al. [7] looked at the angle of robot approach. In a between-subjects experiment, the majority of participants indicated the robot should bring a remote control from a right-frontal side approach, instead of a full-frontal approach. Koay et al. [25] found comparable results in a longitudinal study, however, over time, participants allowed the PeopleBot to approach equally close from the full-front as from the front-side.

Pandey & Alami (2009) developed and tested a framework for a social robot which (autonomously) conformed to four different social conventions, these being: (1) Maintain right-half portion in a narrow passage, (2, 3) pass and overtake a person from his / her left side. (4) Avoid very close sudden appearance from behind a wall. In a between-subjects experiment (N=8), a 84.7% reduction in unwanted behavior was found [29].

From this we conclude that social norms exist for humans, and that, if equipped with social norms, acceptance and user experience of social robots can be improved.

3. THERE'S CULTURE AND THERE'S CULTURE

Culture is an ambiguous concept. We use the following definition of culture: "a fuzzy set of attitudes, beliefs, and behavioral norms, and basic assumptions and values that are shared by a group of people, and that influence each member's behavior and his/her interpretations of the 'meaning' of other people's behavior" [33]. Triandis divided culture into a subjective and material culture. *Material culture* consists of elements, for instance food, houses and tools. *Subjective culture*, on the other hand refers to the characteristic way in which a specific group perceives its environment [36]. When referring to culture, we are referring to subjective culture.

Usually, when scholars are looking at a culture – and the differences between cultures, the level of analysis is the nation, or sometimes subcultures within a nation. Karahanna et al. [22] defined different levels of cultures, these being supranational, national, and levels within a nation, such as the professional, organizational and the group level.

Over the years, there have been several scholars like Hofstede [18] and Pelto [30] who described differences between national cultures according to different dimensions. In a study by Gelfand et al. [11] participants (N=6823) from over 33 countries were asked to rate the appropriateness of twelve behaviors in fifteen everyday situations, and, whether or not there were clear rules for appropriate behaviors in these situations. It was found that there was a high within-nation agreement about the level of constraint in everyday situations, and a high level of variability between-nations. The nation as unit of analysis appears to have proven to be an useful unit of analysis.

A common belief is that society is becoming more and more individualistic, in part due to IT advances. As Jones [21] puts it: "[…] many researchers find culture to be a dynamic, constantly changing field. Cultures are merging, technology is changing the way we communicate, and globalization is changing the way we trade and interface". Thus, the question arises if cultures as a whole are also becoming more individualistic. Hamamura [15] compared national studies studying individualism-collectivism in the U.S. and Japan over time. In contrast to the common belief they concluded both cultures did not become significantly more individualistic. Similar, Gelfand et al. [11] concluded that social constraint appeared to be more or less stable over time in the United States.

Due to various reasons, some of the 196 countries on this planet will have inhabitants with similar cultural backgrounds. We intend to analyze cultures at the supranational level, here being regional clusters of countries.

3.1 Supranational Level: Clusters of Cultures

According to Gupta et al. [13], three major forces have been used historically to cluster countries, these being (1) geographic

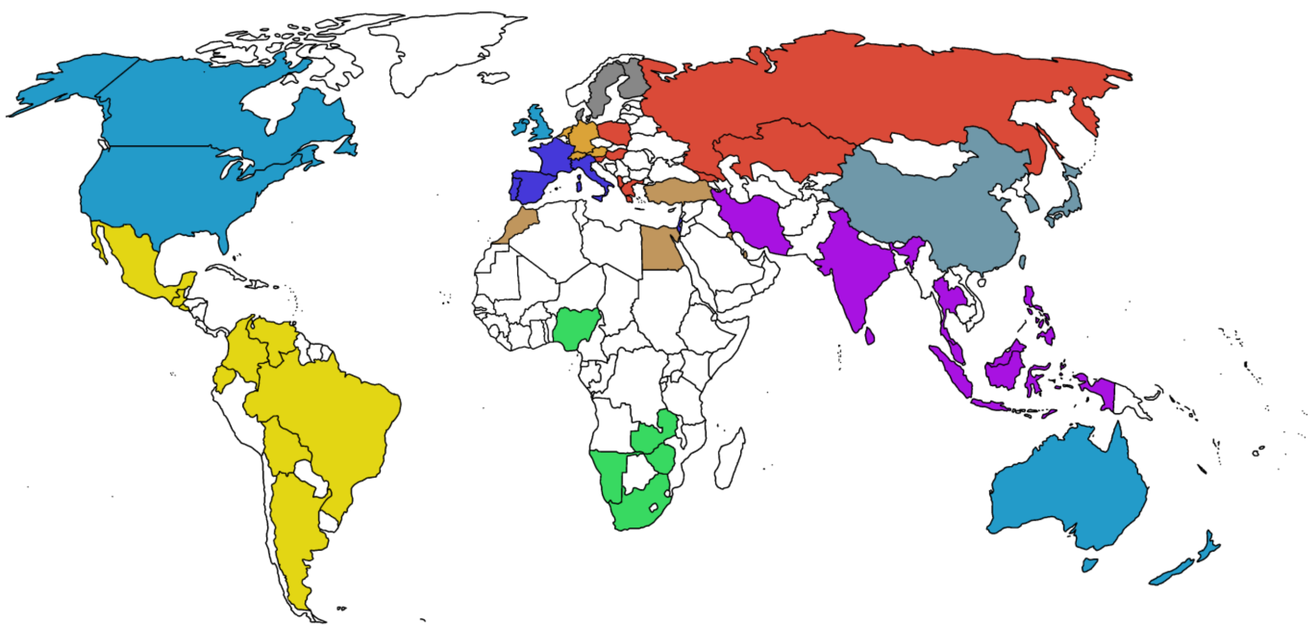


Figure 1. Ten clusters of cultures, figure based upon [13].

Legend: Anglo, Latin Europe, Nordic Europe, Germanic Europe, Eastern Europe, Latin America, Sub-Saharan Africa, Middle East, Southern Asia, Confucian Asia

proximity, (2) mass migration & ethic social capital, and (3) religious and linguistic communality. Societal clustering is a part of the GLOBE project. One of the goals of the authors was to understand similarities and differences among the countries studied within the GLOBE project [20]. As part of this project, 61 nations were clustered into 10 clusters of cultures (see Figure 1, and Appendix I) [13]. Examples include the *Nordic European* cluster containing Finland, Sweden and Denmark, and the *Germanic European* cluster with Austria, Switzerland, the Netherlands and Germany. Appendix I provides the countries contained within each of the ten regional clusters. The remainder of this section will discuss the methodology by which the measures underlying this clustering were developed in more detail.

Among the measures were nine dimensions of culture. These dimensions (*performance orientation, assertiveness, future orientation, humane orientation, institutional collectivism, in-group collectivism, gender egalitarianism, power distance, and uncertainty avoidance)* are the primary measures of interest for us. For each of these scales, questions assessed participants' idea regarding both the practices *(as is)* as well as the values *(should be)* in *organizations* and *society*.

As high wind blows on high hills, there are limitations with the GLOBE project as with any other research paper. Hofstede [17] provides an overview of similarities and differences between the GLOBE study and his own work [16]. One of his major concerns is that the questionnaire items might not have captured what the researchers had in mind, and, that the complete GLOBE questionnaire has not been published. Hofstede is well-known for his work on national value differences while employed by IBM. Five dimensions of national culture were identified based upon results from a survey completed by 117.000 IBM employees. Both GLOBE and Hofstede's IBM studies make sense of culture within an industrial setting. On the other hand, the GLOBE involved managers, whereas the IBM study involved seven categories of employees, of which two were managerial categories [17] of employees. While it can be expected that the GLOBE project will either be loved or hated by scholars, in a way like the IBM study [21], for us the most important fact is that both studies provide empirical evidence that there are differences between cultures.

The next section will provide an overview of cross-cultural research in HRI.

3.2 Human-Robot Cultural differences

Several studies have been conducted in order to explain cultural differences in different situations involving robots. These situations range from a plain, general attitude to robots, to experiments involving human-robot teamwork.

Bartneck et al. [3] distributed a survey among internet users from different countries in which participants were asked to complete the Negative Attitudes towards Robots Scale (NARS) questionnaire. Results indicated cultural background significantly effected attitude towards robots.

In an unpublished experiment by Sau-Lai Lee, reported by Kiesler [24], Chinese participants viewed a video of robot interaction with an experimenter, they were asked whether or not the robot would know certain landmarks. The "cultural background" of the robot was manipulated by having the robot talk either English or Cantonese, and informing participants the robot was created in either China or New York. Based upon the origin of the robot, people had a different mental model of the robot. Lee found two relevant results providing evidence for this. First, people expected the robot to have more knowledge about famous landmarks in both countries, than about not so famous landmarks. The second, perhaps the most important: participants expected the "Chinese" robot to know more about Chinese landmarks than the "American" robot, and vice versa. In a similar way, Trovato et al. [37] found that Egyptian and Japanese participants preferred a robot displaying a similar cultural background. A robot was programmed to greet participants in the English language with either an Arabic or Japanese accent, and performing a greeting gesture also performed by humans in that culture. It was found that Japanese participants preferred the Japanese robot, and Egyptians the Arabic robot.

Wang et al. [39] conducted a 2x2 experiment involving robots, manipulating culture and robot communication style. 320 participants, 80 Chinese dyads and 80 U.S. dyads, interacted with a robot providing advice either implicitly or explicitly. The underlying hypothesis was that since the Chinese typically prefer and implicit communication style, and U.S. people a more explicit, a robot displaying a matching communication style would be seen as a more in-group member and thus more trusted and perceived as more credible. Supporting their hypothesis, Chinese participants preferred the implicit robot whereas U.S. participants preferred the explicit robot. Furthermore, when the robot communicated in the preferred way, participants were more likely to change their decisions in order to align with the robot.

Li et al. [26] also found evidence in a HRI trial that participants from a low-context culture (Germany) had different scores with respect to the evaluation of the interaction than those from high-context cultures (Chinese and Korean).

From the above we expect people from different cultures will have different views on which behaviors are normative for a robot. Previous work with regards to cultural aspects in HRI has been limited mostly to human-robot collaborative teamwork. The work in HRI on proxemics has not yet taken culture into account, which could become a shortcoming when robots are going to interact in public spaces with people having different cultural backgrounds.

4. TOWARDS A METHODOLOGY

In this section, we will describe two major challenges we see for HRI research researching cross-cultural robot behavior. These challenges are:

1) Choosing a research methodology

2) Sampling of cultures of interest

We will describe both challenges, insofar as not discussed before, and offer our ideas to solve this in Section 4.2.

4.1 Overview of methodologies

Different methodologies have been employed in order to gather data from participants from different cultures. In this section, we will first provide an overview of different methods which have been used to find answers with regards to cross-cultural differences, both in human-human, and human-robot interaction. We will then conclude with an experimental setup.

A number of studies manipulated culture by using native students and exchange students in a lab experiment. ([4], [26], [34]). Already in the 80s, Baldassare & Feller [2] hinted that the frequent comparison of U.S. versus exchange students of a culture decreases ecological validity, because a) the students are not observed in their natural culture, b) they have been influenced by North American proxemics patterns for an undisclosed time, and c) they are not a representative sample. Wang et al. [39] collected data at two separate sites; thus using native students in both settings. However, this sample was also not representative because it only included students.

Woods et al. [40] used a method called "video-based human-robot interaction" (VHRI) in which participants viewed videos of a human interacting with a robot. Results between this video-based methodology and a lab experiment with real participants were found to be comparable.

Self-reported measures, such as questionnaires, were also frequently employed. The advantage here being able to use participants from geographically distributed locations. ([3], [11], [5]). All reported studies report having the questionnaires translated and back-translated into the participants' native language.

Two experiments made use of either scaled dolls or silhouettes in order to capture people's impression of appropriate interpersonal distance in different situations ([27], [28]). Like a lab experiment, the use of dolls does require some sort of physical location when collecting data at different sites.

All these methods have advantages and disadvantages. The first method, experiments with an actual embodied robot, would be preferred for HRI since it would provide the most realistic setting. An ideal situation would be an experiment, be it a Wizard-of-Oz experiment with one type of robot, shipped all over the world to various data collection sites. This is an utopian experiment design in a world not constrained by resources like time, money and man-hours. The other methods (VHRI studies and scaled figures) could provide a solution, albeit generalizability of the results to a real-world setting could be questioned. In the next section we propose a hybrid approach to tackle these issues.

4.2 Proposed methodology

At this moment, we are conducting a survey with this setup using stills of 3D people. This survey is currently being distributed to three countries. While data collection has not yet been finished, one of the possible issues we might face is that the results are not generalizable enough because when you approach a group, the formation of the group is going to change as soon as you approach. Therefore, the use of 3D pictures might not be a sufficient methodology to investigate cross-cultural robot spatial behavior.

Based on this insight, we propose a combination of a lab- and video study to increase ecological validity while investigating the following questions:

1) "From which angle should a robot approach a small group of people?"

2) "Do people from different cultures have significant different preferences when a robot approaches a small group of people?"

3) "Do survey-based HRI studies provide reliable results when used in lieu of experiments when evaluating robot spatial behavior?"

In our situation, we have access to two robots of similar design, at two different sites – a site in the Netherlands, and a site in Spain. We propose to run a between-groups field experiment at both locations, thus having two different cultures. In the experiment, we will ask small groups of people (3-5) to stand in a room and discuss a topic. Participants will be informed that after a minute a robot will approach the group and bring the new discussion topic. The robot will approach the group from various angles, and stop at different distances.

At one of these locations, we will make a video recording of the different experiment conditions with actors. In order to test if the behaviors are perceived equally (un-)appropriate in videos compared with the field experiment, we will distribute the video to participants from the same countries as those in the field experiment. If it turns out to be true, the questionnaire can be distributed to participants with cultural backgrounds not investigated in the field experiment.

5. CONCLUSION

Service robots start entering our daily lives. When real social robots do, an important question will be if culturally different motion behaviors are necessary for a robot guiding people with distinct different backgrounds. Previous HRI research focusing on cultural aspects does not provide indisputable results, though we find it likely these results could surface when evaluating motion behaviors with respect to different cultures.

Based upon an overview of previously used methods to evaluate cross-cultural differences we have proposed a mixed-methods method in order to evaluate cross-cultural HRI behavior preferences in a resource-efficient way.

6. ACKNOWLEDGMENTS

This work has partly been supported by the European Commission under contract number FP7-ICT-600877 (SPENCER).

7. REFERENCES

[1] Aarts, H. and Dijksterhuis, A. 2003. The silence of the library: Environment, situational norm, and social behavior. *Journal of Personality and Social Psychology*. 84, 1 (2003), 18–28.

[2] Baldassare, M. and Feller, S. 1975. Cultural Variations in Personal Space: Theory, Methods, and Evidence. *Ethos*. 3, 4 (1975), 481–503.

[3] Bartneck, C., Suzuki, T., Kanda, T. and Nomura, T. 2006. The influence of people's culture and prior experiences with Aibo on their attitude towards robots. *Ai & Society*. 21, 1-2 (May 2006), 217–230.

[4] Beaulieu, C. 2004. Intercultural study of personal space: A case study. *Journal of Applied Social Psychology*. 34, 4 (2004), 794–805.

[5] Brauer, M. and Chaurand, N. 2010. Descriptive norms, prescriptive norms, and social control: An intercultural comparison of people's reactions to uncivil behaviors. *European Journal of Social Psychology*. 499, June 2009 (2010), 490–499.

[6] Cialdini, R.B., Reno, R.R. and Kallgren, C.A. 1990. A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology*. 58, 6 (1990), 1015–1026.

[7] Dautenhahn, K., Walters, M.L., Woods, S.N., Nehaniv, C.L., Sisbot, E., Alami, R. and Siméon, T. 2006. How may I serve you?: a robot companion approaching a seated person in a helping context. *Proceedings of the 2006 IEEE Conference on Human-Robot Interaction* (2006), 172–179.

[8] Elster, J. 1989. Social norms and economic theory. The *Journal of Economic Perspectives*. 3, 4 (1989), 99–117.

[9] Epley, N., Waytz, A. and Cacioppo, J.T. 2007. On seeing human: a three-factor theory of anthropomorphism. *Psychological review*. 114, 4 (Oct. 2007), 864–86.

[10] Fong, T., Nourbakhsh, I. and Dautenhahn, K. 2003. A Survey of Socially Interactive Robots: Concepts, Design, and Applications. *Robotics and autonomous systems*. 42, 3 (2003), 143–166.

[11] Gelfand, M.J. et al. 2011. Differences between tight and loose cultures: a 33-nation study. *Science*. 332, 6033 (May 2011), 1100–4.

[12] Goetz, J., Kiesler, S. and Powers, A. 2003. Matching Robot Appearance and Behavior to Tasks to Improve Human-Robot Cooperation. *Proceedings of the 12th IEEE International Symposium on Robot and Human Interactive Communication* (RO-MAN 2003) (2003), 55–60.

[13] Gupta, V., Hanges, P. and Dorfman, P. 2002. Cultural clusters: Methodology and findings. *Journal of World Business*. 37, (2002), 11–15.

[14] Hall, E.T. 1966. *The Hidden Dimension*. Anchor Books.

[15] Hamamura, T. 2012. Are cultures becoming individualistic? A cross-temporal comparison of individualism-collectivism in the United States and Japan. *Personality and Social Psychology Review*. 16, 1 (Feb. 2012), 3–24.

[16] Hofstede, G. 2001. *Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations across Nations*. Sage Publications, Inc.

[17] Hofstede, G. 2006. What did GLOBE really measure? Researchers' minds versus respondents' minds. *Journal of International Business Studies*. 37, 6 (2006), 882–896.

[18] Hofstede, G. and Hofstede, G.J. 2005. *Cultures and organizations: software of the mind*. McGraw-Hill.

[19] Høgh-Olesen, H. 2008. Human Spatial Behaviour: The Spacing of People, Objects and Animals in Six Cross-Cultural Samples. *Journal of Cognition and Culture*. 8, 3 (Aug. 2008), 245–280.

[20] House, R., Hanges, P., Javidan, M., Dorfman, P. and Gupta, V. 2004. *Culture, leadership, and organizations: The GLOBE Study of 62 Societies*. Sage Publications, Inc.

[21] Jones, M. 2007. Hofstede-culturally questionable? *Oxford Business & Economics Conference* (Oxford, U.K., 2007).

[22] Karahanna, E., Evaristo, J.R. and Srite, M. 2005. Levels of Culture and Individual Behavior: An Integrative Perspective. *Journal of Global Information Management*. 13, 2 (2005).

[23] Keizer, K., Lindenberg, S. and Steg, L. 2008. The spreading of disorder. *Science*. 322, 5908 (Dec. 2008), 1681–5.

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

[25] Koay, K.L., Syrdal, D.S., Walters, M.L. and Dautenhahn, K. 2007. Living with Robots: Investigating the Habituation Effect in Participants' Preferences During a Longitudinal Human-Robot Interaction Study. *Proceedings of the the 16th IEEE International Symposium on Robot and Human Interactive Communication(RO-MAN 2007)* (2007), 564–569.

[26] Li, D., Rau, P.L.P. and Li, Y. 2010. A Cross-cultural Study: Effect of Robot Appearance and Task. *International Journal of Social Robotics*. 2, 2 (May 2010), 175–186.

[27] Little, K.B. 1968. Cultural variations in social schemata. *Journal of Personality and Social Psychology*. 10, 1 (Sep. 1968), 1–7.

[28] Lomranz, J. 1976. Cultural variations in personal space. *The Journal of Social Psychology*. 99, 1 (1976), 21–27.

[29] Pandey, A. and Alami, R. 2009. A framework for adapting social conventions in a mobile robot motion in human-centered environment. *International Conference on Advanced Robotics (ICAR 2009)* (Munich, 2009), 1–8.

[30] Pelto, P. 1968. The differences between "tight" and "loose" societies. *Trans-action*. 5, 5 (1968), 37–40.

[31] Pike, K.L. 1966. Etic and Emic Standpoints for the Description of Behavior. *Communication and Culture: readings in the Codes of Human Interaction*. A.G. Smith, ed. Holt, Rinehart and Winston, Inc. 152–163.

[32] Prentice, D.A. 2012. The psychology of social norms and the promotion of human rights. *Understanding Social Action, Promoting Human Rights*. R. Goodman, D. Jinks, and A.K. Woods, eds. Oxford University Press, USA. 23–46.

[33] Spencer-Oatey, H. 2000. *Culturally Speaking: Managing Rapport Through Talk Across Cultures*. Continuum.

[34] Sussman, N.M. and Rosenfeld, H.M. 1982. Influence of culture, language, and sex on conversational distance. *Journal of Personality and Social Psychology*. 42, 1 (1982), 66–74.

[35] Takayama, L. and Pantofaru, C. 2009. Influences on proxemic behaviors in human-robot interaction. *IEEE/RSJ International Conference on Robots and Systems (IROS2009)* (2009), 5495–5502.

[36] Triandis, H. 2002. Subjective culture. *Online Readings in Psychology and Culture*. 2, 2 (2002), 1–12.

[37] Trovato, G., Zecca, M., Sessa, S., Jamone, L., Ham, J., Hashimoto, K. and Takanishi, A. 2013. Cross-cultural study on human-robot greeting interaction: acceptance and discomfort by Egyptians and Japanese. *Paladyn, Journal of Behavioral Robotics*. 4, 2 (2013), 83–93.

[38] Walters, M.L. 2008. *The design space for robot appearance and behaviour for social robot companions*. University of Hertfordshire.

[39] Wang, L., Rau, P. and Evers, V. 2010. When in Rome: the role of culture & context in adherence to robot recommendations. *Proceedings of the 2010 IEEE Conference on Human-Robot Interaction* (Mar. 2010), 359–366.

[40] Woods, S.N., Walters, M.L., Koay, K.L. and Dautenhahn, K. 2006. Methodological Issues in HRI : A Comparison of Live and Video- Based Methods in Robot to Human Approach Direction Trials. *Proceedings of the the 15th IEEE International Symposium on Robot and Human Interactive Communication*(RO-MAN 2006 (2006), 51–58.

8. Appendix I

Table 3 provides the ten GLOBE clusters of societies and the respective countries within each cluster.

Table 3. GLOBE clusters. Source [13]

Anglo Cultures England, Australia, South Africa (White sample), Canada, New Zealand, Ireland, United States
Confucian Asia China, Hong Kong, Japan, Singapore, South Korea, Taiwan
Eastern Europe Albania, Georgia, Greece, Hungary, Kazakhstan, Poland, Russia, Slovenia
Germanic Europe Austria, Germany, Netherlands, Switzerland (German speaking)
Latin America Argentina, Bolivia, Brazil, Colombia, Costa Rica, Ecuador, El Salvador, Guatemala, Mexico, Venezuela.
Latin Europe France, Israel, Italy, Portugal, Spain, Switzerland (French speaking)
Nordic Europe Finland, Sweden, Denmark
Southern Asia India, Indonesia, Iran, Malaysia, Philippines, Thailand
Sub-Sahara Africa Namibia, Nigeria, South Africa (Black sample), Zambia, Zimbabwe
Middle East Egypt, Kuwait, Morocco, Qatar, Turkey