"I Would Like to Get Close to You": Making Robot Personal Space Invasion Less Intrusive with a Social Gaze Cue

Stefan-Daniel Suvei[1(✉)], Jered Vroon[3], Vella V. Somoza Sanchéz[2], Leon Bodenhagen[1], Gwenn Englebienne[3], Norbert Krüger[1], and Vanessa Evers[3]

[1] MMMI, University of Southern Denmark, 5230 Odense M, Denmark
{stdasu,lebo,norbert}@mmmi.sdu.dk
[2] IMM, University of Southern Denmark, 5230 Odense M, Denmark
vella@sam.sdu.dk
[3] HMI, University of Twente, P.O. Box 217, 7500AE Enschede, The Netherlands
{j.h.vroon,g.englebienne,v.evers}@utwente.nl

Abstract. How can a social robot get physically close to the people it needs to interact with? We investigated the effect of a social gaze cue by a human-sized mobile robot on the effects of personal space invasion by that robot. In our 2×2 between-subject experiment, our robot would approach our participants (n = 83), with/without personal space invasion, and with/without a social gaze cue. With a questionnaire, we measured subjective perception of warmth, competence, and comfort after such an interaction. In addition, we used on-board sensors and a tracking system to measure the dynamics of social positioning behavior. While we did find significant differences in the social positioning dynamics of the participants, no such effect was found upon quantitative analysis of perception of the robot. In a subsequent inductive analysis we further investigated these results, our findings suggesting that the social cue *did* play a role for the participants – particularly related to their perceived safety.

Keywords: Personal space invasion · Social robotics · Social cues

1 Introduction

As mobile robots are used increasingly often in everyday social settings, the design of appropriate robot behaviors has become more and more relevant. While the contexts can vary significantly, all these robots need to know how to position themselves in relation to people in a way that supports the intended interaction. Ideally, we can design robot social positioning behaviors that allow for smooth and efficient interactions, that feel natural to the people being interacted with.

Most prior work on social positioning for robots has treated this in a static way by trying to find *the* appropriate distance or positioning depending on a

variety of factors that can be established at the beginning of the interaction. This seems to be rooted in the idea that it should be avoided that the robot 'gets too close', i.e. conducts personal space invasion (**PSI**). For example, earlier work has resulted in tables giving stopping distances to be used depending on factors such as human-likeness and height of a robot [1], and suggested that appropriate stopping distance may also depend on agreeableness or previous experience with animals or pets of the person being interacted with [2].

At the same time, social positioning behaviors can be highly dynamic during interactions. We can get really close, up to the point of "PSI", to someone we do not know well if we are sharing a secret, or if the environment is noisy, or if they have hearing problems. Similarly, people have been found to take the perceptional limitations of a robot into account in their own proxemic preferences [3]. In line with the later work on social positioning in human-human interaction (see e.g. the extensive 1987 review from Aiello [4]), this suggests that the appropriate interaction distance is something dynamic that can change several times during an interaction. In all these interactions signals are given to indicate to ones communication partner the current appropriateness of a chosen position (**social feedback cues**), as we have argued previously [5].

In this work, we investigated how a mobile robot (see Fig. 1) can provide a brief feedback cue to make the social positioning dynamic more smooth. Specifically, we tried to use a brief gaze change to allow a robot to conduct "PSI" without making the interaction less comfortable for the human user. This allowed for the investigation of the effect of using such a brief social cue on the overall interaction smoothness and the social perception of the robot. While this specific task was designed for research purposes, similar tasks could be used in real-life applications, e.g. in airports, institution entrances (including university), stadiums, or concert halls, where a robot could be used to provide assistance, if such a desire is signaled by the human user. As such, our main research question can be formulated as follows:

What are the effects of a robot's PSI and/or gaze change on:

1. the way in which people respond to that robot's approach by moving away from (or towards the robot)?
2. the perception of the robot in terms of warmth, competence, and discomfort?

The remainder of this paper is structured as follows. We will first discuss related work on PSI, gaze change, and measures of social perception, resulting in a specification of our hypotheses (Sect. 2). We will then discuss our experimental set-up in detail, including a description of the autonomous robot behaviors we implemented (Sect. 3). While we did not find effects on the perception of the robot, as assessed by questionnaire, we did find effects on participants' behavior – which we further investigated in our qualitative analysis (Sect. 4). Our findings show that, indeed, even a brief social feedback cue, such as a gaze change, plays a role in how PSI is perceived and responded to – providing an argument for the use of such social feedback cues in designing behaviors for social robots (Sect. 5).
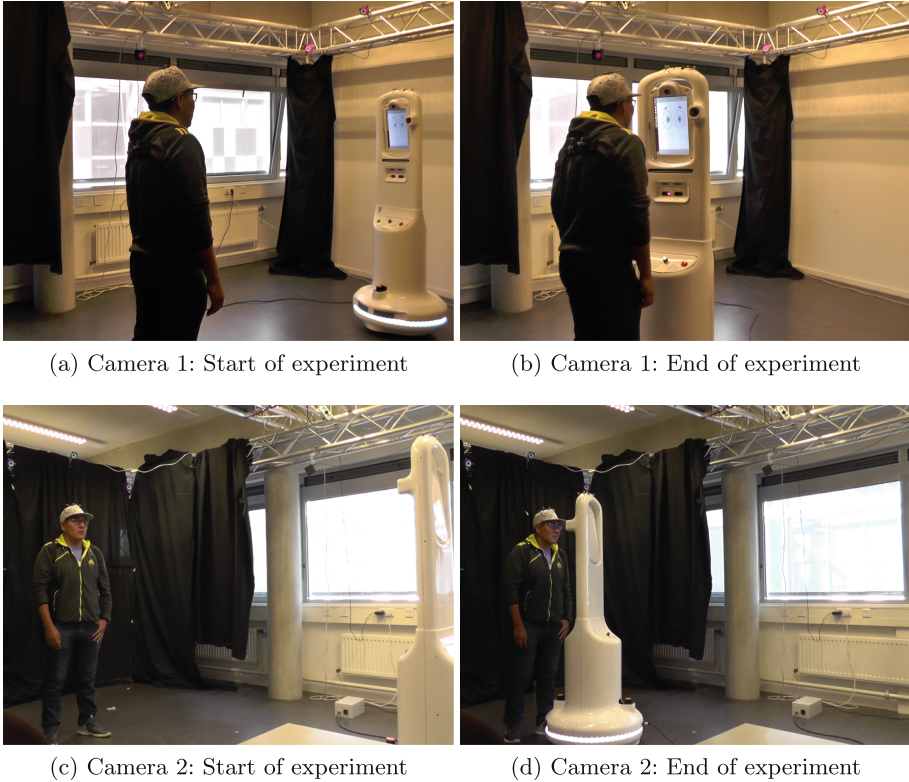
(a) Camera 1: Start of experiment   (b) Camera 1: End of experiment

(c) Camera 2: Start of experiment   (d) Camera 2: End of experiment

Fig. 1. The experiment, as seen from the two different cameras.

2   State of the Art

Much of the existing work in social positioning for robots is based on two theories on social positioning in humans. **Proxemics**, a term coined by Hall in [6], focuses on the distances people use in social interactions. **F-formations**, as introduced by Kendon [7], describe the different spatial arrangements used by people. As proxemics and F-formations would predict, many different social situations can be distinguished based on only position and orientation information (e.g. [8,9]).

Previous work has applied and investigated proxemics and F-formations in the context of robotics. This work often treats social positioning as something mostly static. For example, several authors have tried to establish the 'appropriate interaction distance' depending on a variety of factors, such as height and human likeness of the robot [1] as well as personality traits of the person being interacted with [2].

Interestingly, when we look at later work on social positioning in humans, it is treated more as a social dynamic, that can change during an interaction and that can carry communicative meaning within that interaction [4]. Information

can be (non-verbally) transmitted about social positioning via responsiveness, as we argued in [5], by the usage of social feedback cues.

In the literature, there are only a few examples of artificial agents utilizing social feedback cues to influence the dynamics of social positioning. Recent work by Mead and Matarić [3] found that robots can effectively signal their (perceptual) needs to influence the proxemic preferences of people with whom they are interacting. Using a virtual agent instead of a physical robot, Kastanis and Slater [10] have also investigated ways to influence the proxemic preferences of people; they trained an agent to position itself such as to most effectively cause participants to move to a particular position in a space. Work by Jung et al. investigated human-robot teamwork and found that when their robots used back-channeling, this improved team functioning, though it also decreased perceived competence [11].

Facial features represent a key component of any social interaction, because they are easily recognizable and offer important cues regarding the inner states of the involved social actors. There exists a range of work in which artificial 'facial' features are used effectively as social feedback cues for non-mobile robots. Consider, as an example, the Baxter robot, discussed in [12], which is equipped with a screen which displays a pair of animated eyes and eyebrows. Even though the animations are limited to changing the angles of the eyebrows, opening and closing the eyes and moving the pupils, this is enough to signal to the human user whether or not the robot has understood the task.

In addition to these animations of facial features, there has also been work looking more specifically into the effects of different gaze behaviors. Recent work found that users feel more at ease with a robot that maintains visual contact while the task is being performed [13]. In contrast, in a human-human object handing interaction, the approaching human was found to be expected to maintain eye contact just towards the end of the interaction [14]. Other work has investigated and found gender effects, showing that women tend to position themselves further from a robot if it is constantly looking at them [2].

In all, though much work remains to be done, the prior work shows that people are sensitive to the social feedback cues used by robots and suggest that the use of face-inspired features – specifically gaze – can be an especially effective social feedback cue.

2.1   Social Perception Measurement

Because robots are more and more present in our day-to-day life, being able to correctly measure the way they are perceived has become increasingly important. Existing research shows that people have the tendency of making social category judgments, similar to the way they do in human-human interactions, based on the physical appearance, gender, race or nationality of the robot [15]. Furthermore, there is strong indication that people anthropomorphize robots, assigning human characteristics to them, especially when human characteristics are accessible and applicable [16]. As such, an interesting aspect to measure is

the association of characteristics and traits based on the features or the perceived social category membership of the robot.

Literature on social psychology has also shown that two of the main drivers for the judgment of other humans are *warmth* and *competence*. The main reason for this is because these attributes link to basic survival instincts − we quickly evaluate if another person wants to harm us and if they are capable of doing so [17]. Furthermore, different combinations of competence and warmth will evoke an array of different other emotions, such as discomfort, pity or envy [18]. As shown in [19], there is strong indication that warmth and possibly other social attributes are important factors that can affect the perception of a robot. A method that offers the means to asses the central attributes implicated in human perception of robots and how such attributes can influence human-robot interactions is the Robotic Social Attributes Scale (**RoSAS**), presented in [20]. The RoSAS is a psychologically valid, 18-item scale, that was constructed using items from the Godspeed Scale [21] combined with key ideas taken from the psychological literature regarding social perception. The advantage of this scale is that it is psychometrically validated and it can be used to measure different types of robots, disregarding their purpose.

2.2   Hypotheses

Based on the existing literature on social positioning, we would expect that PSI has a negative effect on the interaction - causing the participant to step away and perceive it more negatively (which we will measure with the RoSAS). At the same time, there is work suggesting that we can leverage the dynamics of interactions to mediate this effect. More specifically, we show that by using gaze change as an appropriate social feedback cue, the robot can make its PSI less intrusive. Combined, this leads to the following hypotheses:

Hypothesis 1: When a robot conducts PSI without a social cue, as compared to not conducting PSI without a social cue, we expect people to:

a. step away from it, and;
b. perceive it more negatively in terms of warmth, competence, and discomfort.

Hypothesis 2: We expect an interaction effect between PSI and social cue (gaze change), such that the social cue 'compensates' for the PSI, leading people who experience both PSI *and* a social cue (gaze change) to:

a. step away from it less, and;
b. perceive it more positively in terms of warmth, competence, and discomfort.

3   Method

To test the proposed hypotheses, we conducted a 2×2 between-subject experiment. Our autonomous robot (Sect. 3.1) approached the participant in the

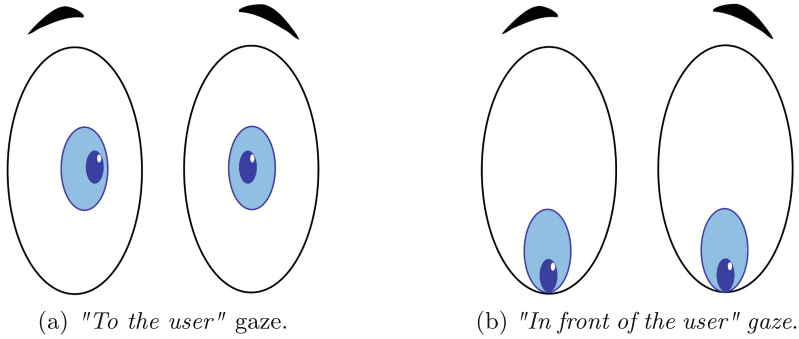(a) *"To the user"* gaze.   (b) *"In front of the user"* gaze.

Fig. 2. The displayed social gaze cue.

context of potentially providing assistance to the user (Sect. 3.2). During the approach we manipulated PSI by the robot, as well as the use of gaze change as a social cue (Sect. 3.3). To collect our data, we used a questionnaire, a semi-structured interview, and several sensors for objective measures (Sect. 3.4).

3.1   Robot

The robot used in the experiments is a human-sized mobile robot (see Fig. 1), build within the Teresa project [22]. The system was build on top of the Giraff robot platform, with a new shell, and additional sensors, including: two RGB-depth cameras (Asus Xtion) and two laser range scanners (Hokuya LIDAR). A full description of the platform can be found in the Teresa project deliverables [23]. The platform is further equipped with a screen at head height, which can be used for telepresence by showing a live feed of the remote user. For the purpose of our experiment, the screen was used to signal the social feedback cue in the form of a short animation which changes the gaze of the robot's eyes, as shown in Fig. 2.

The behavior displayed by the robot was autonomous, combining person detection with navigation. The method works by detecting the participant in the scene, computing his/her position in the map coordinate frame and generating an appropriate goal position that will allow the robot to approach the human with PSI or without, depending on the condition (see Fig. 3). Before driving to the generated goal position, the robot displays the gaze-change animation on the screen, depending on the condition. Four seconds after the robot reached the goal position the LEDs on the base of the robot would light up green to signal the end of the experiment. Our person detection algorithm was build based on existing code and combined leg detection based on the laser range finder data (wide range) with upper-body tracking based on the RGB-depth data (also providing information on posture). We used map-based navigation. The map of the environment was built beforehand, using the laser range scanners and the Simultaneous Localization and Mapping [24] method. The robot uses the
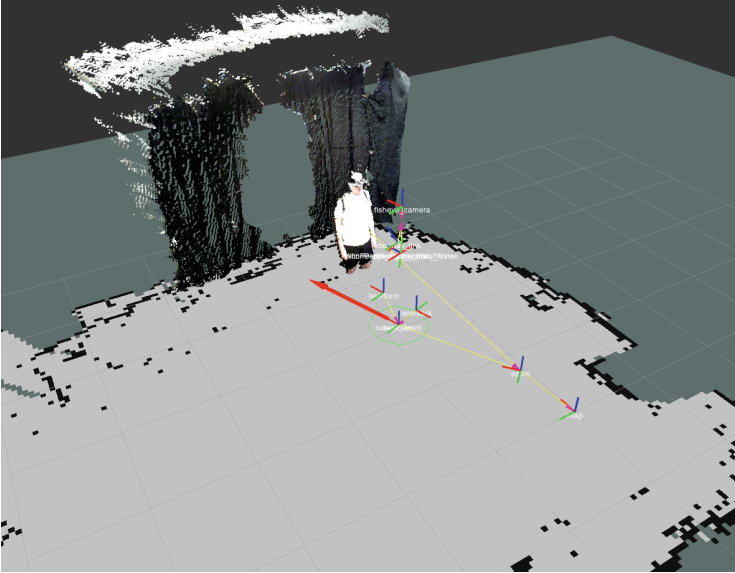
Fig. 3. The behavior of the robot, as displayed in Rviz. The red arrow represents the goal position for the robot and it is computed relative to the position of the detected participant. (Color figure online)

Monte Carlo [25] algorithm to self-localize in the map and the ROS move_base framework for navigation. The speed of the robot was limited to 1 m/s.

3.2   Task

Given our research questions, the task necessitated a form of 'PSI'. To avoid an effect of any additional motions, we aimed for a task in which the only required action was a spatial positioning by the robot. For this reason, we designed a close-up approaching task, with the purpose of potentially providing assistance to the user. In the task, participants were told that the robot would approach them to potentially offer assistance (see Fig. 4). While free to move around the test area, the participants were advised to wait for the robot to approach them.

3.3   Manipulations

We manipulated two factors; approach distance [PSI/noPSI], and the giving of a social cue [GAZE/noGAZE]. This resulted in four conditions, PSI-GAZE, PSI-noGAZE, noPSI-GAZE, and noPSI-noGAZE.

Approach Distance [PSI/noPSI]: During the approach the robot would navigate to a location at a given distance to the participant. To have the robot conduct PSI (or not), we manipulated this distance. Based on the related work discussed
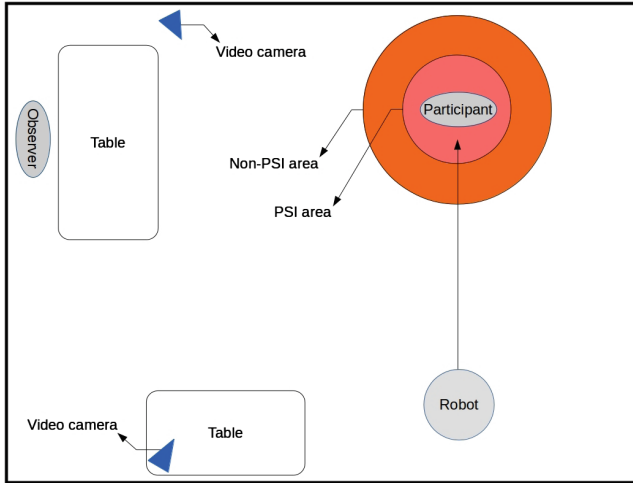
Fig. 4. Overview of the experiment area.

above and pilot testing, we expected distances below 45 cm to be perceived as a (strong) PSI. Therefore, we used a distance of 20 cm for the **PSI** conditions, and a distance of 50 cm for the **noPSI** conditions.

Social Cue [GAZE/noGAZE]: During the approach, an animation of two eyes would be shown on the screen of the robot. To have the robot provide a social cue, we manipulated the apparent gaze direction of these eyes (see Fig. 2). Since the eyes were shown on a 2D screen, we did not manipulate actual gaze direction; instead we used an emphasized animation of the pupil position to suggest apparent gaze direction.

In our prior empirical tests, we saw that a change of the apparent gaze direction from *"to the user"* to *"down, in front of the user"* felt like the robot was giving a social cue – which aligns with the related work discussed above. For the **noGAZE** condition, we used the eyes with a gaze "to the user", which would not change during the interaction. For the **GAZE** condition, we would start the interaction with eyes "to the user", changing to "in front of the user" as the robot would start its approach. This change was fast, which meant that it could potentially be missed if participants were not paying attention, but our informal prior tests suggested that making the apparent gaze change slower would feel more as awkward than as a social cue.

3.4 Measures

Objective Measures. To allow us to determine how much the participants stepped away from the robot, we recorded the movement of the participants. For this we used both the on-board sensors, and an external system using markers (OptiTrack). In addition, we also collected video data.

On-Board Sensors. For each experiment, the Xtion sensor and the laser range scanner were tracking and recording the movement of the participant. The tracking data is based on the 3D sensor depth information, by computing the XY-position of the human in the map, relative to the robot's base, once the robot has reached its goal position.

OptiTrack. The participant was equipped with two markers (one on the back of the chest, one on a cap), which were tracked using the OptiTrack motion capture system using 8 infra red cameras. The robot also had a marker, located on its screen. The system used allows sub-centimeter level precision tracking of the position and orientation of those markers. In our set-up we found that the system did not always reliably detect the markers, presumably because of occlusions and reflections on the shell of the robot.

Video Recordings. Two cameras recorded the interaction, providing a side view and a view on the face and upper body of the participant. These cameras were also used to record the interviews.

Subjective Measures. To measure perception of the robot, we used a questionnaire and a semi-structured interview. While the questionnaire gave more quantified data, the interview was intended to give us rich qualitative insights.

Questionnaire. The employed questionnaire asked participants to rate the robot and their experience in terms of 18 items on a 9-point Likert scale. These 18 items follow the RoSAS and are used for the social perception analysis of the robot [20]. They are grouped into 3 main factors, as follows:

Warmth: Happy, Feeling, Social, Organic, Compassionate and Emotional.
Competence: Capable, Responsive, Interactive, Reliable, Competent and Knowledgeable.
Discomfort: Scary, Strange, Awkward, Dangerous, Awful and Aggressive.

Furthermore, the participants are asked to describe their interaction in at least 140 characters. To ensure the fact that the social cue was perceived correctly, when signaled, the questionnaire also contained four manipulation checks which asked the user to mention what was the gaze of the robot at the beginning and at the end of the interaction, as well as the color of the LEDs. Additionally, the questionnaire also contained questions regarding the age, gender and nationality of the participant. The participants were also asked about their prior experience with robots and pets.

Interview. After the experiment and filling out the questionnaire, a short semi-structured interview was done with each participant. The two main questions asked were about their experience and their general feeling about the robot, as a result of the interaction. Depending on the answers, follow-up questions were asked, as for example "What do you think is the purpose behind this behavior?" or "How did you feel about the approach?".

3.5 Procedure

The participants were greeted at the entrance of the lab and were lead inside, where the context of the experiment was explained by the experimenter. They were then given an informed consent form, which they were asked to read and sign. The participants were then guided to the experiment area and equipped with the OptiTracker markers used for tracking their movements. They were then told to wait for the robot to approach them. We emphasized that while they were free to move, they should really let the robot approach them, and not the other way around. The robot would then approach them to a distance of 20–50 cm, exhibiting the behavior as appropriate for the condition the participant was in. The participants were recorded using two video cameras. Additionally, the OptiTracker system and the robot's sensors (the laser scanner and the ASUS camera) were recording their movement. After the interaction with the robot, the participants were asked to fill in the questionnaire on the interaction, and a brief interview was conducted by the experimenter.

3.6 Participants

We recruited 83 students and staff members from the campus of the University of Twente. Participants were recruited using word of mouth, and adverts. For agreeing to participate in the experiment, participants were compensated with chocolate. The average age of the participants was 24.25 (range 18–45). 66.2% of the participants were Dutch. 34 of the participants were women and 49 were men. Regarding previous experience with robots, 6% of the participants have built a robot before, 43.3% stated that they have worked with robots, 25.3% said that they have seen robots before, 16.8% have played with a robot before and 7% had no previous experience with robots. Additionally, 84.3% of the participants had previous experience with pets.

4 Results

To investigate the proposed hypotheses we employ both a deductive and a qualitative analysis of the data available from the experiments - this means sensor data, the OptiTracker data, questionnaire data, interview transcripts and video data.

4.1 Deductive Analysis

As described in Sect. 3.3, the movement of the participants relative to the robot was tracked with the help of the robot's own sensors and the OptiTracker system. The deductive analysis refers to the behavioral analysis shown by the participants, as measured with the help of these sensors, and to the evaluation of the questionnaire results, as proposed by the RoSAS model.

Due to robot failures, the experiment data for 3 of the participants had to be excluded. Additionally, for the statistical analysis of Hypothesis 2, we removed

all the participants that either did not perceive the feedback social cue or that falsely considered that there was such a cue signaled. The elimination criterion was based on the answers that the participants gave to the two manipulation checks regarding the robot's gaze at the beginning and the end of the experiment from the questionnaire. As a result, the data for 58 participants out of 83 were used in our final investigation regarding Hypothesis 2 – 14 in the noPSI-GAZE and the PSI-GAZE conditions, and 15 participants in each of the other two.

Computing Distance (Lower)body-robot (From On-Board Sensors). We used the data from the Xtion sensor (RGB-depth) and the laser range scanners to determine the displacement of participants during the interaction. To do so, we first used the person tracking (see Sect. 3.1) based on a combination of both sensors. This resulted, for each time-stamp, in the position of the participant relative to the robot. Since some of detections were based only on the leg-detection from the laser range scanners (due to the participants moving outside the field of view of the camera), we used only the position on the floor plane. From this time-stamped list of positions, we then selected the 3 or 4 s of data to use based on the time-stamp at which the robot reached its goal position and stopped its movement. Within this subset, we then found the participant's minimum and maximum distance to the robot. The displacement was then calculated as the difference between this minimum and maximum. For 10 participants the data from the on-board sensors was not of sufficient quality to reliably compute the displacement.

Computing Head Distance to the Robot (From OptiTrack Data). We additionally used the data from the OptiTrack system to compute displacement of the participants during the interaction. Where the on-board sensors computed position based on center of mass (with Xtion) or on position of legs (with laser range finders), the OptiTrack allowed us to instead track head position. Additionally, with the OptiTrack only for 1 participant the data from the on-board sensors was not of sufficient quality to reliably compute the displacement. The OptiTrack system did occasionally detect the location of markers wrongly, presumably because of reflections. For this reason, we first smoothed the tracked positions, using a 500-frame moving average (data recorded at 120 fps). We manually coded the end of each robot approach based on the speed profile of the robot, and used a time window from the end of the approach until 4 s later. Within this time frame, we then computed the minimum and the maximum distance on the floor plane between the (smoothed) position of the marker on the base and the (smoothed) position of the marker worn on the participant's head. The displacement was then calculated as the difference between this minimum and maximum.

Scoring Competence, Warmth, and Discomfort from Questionnaire Data. A principal components analysis (PCA) was run on the 18-item RoSAS questionnaire results. The suitability of PCA was assessed prior to analysis.

All variables had at least correlation coefficient greater than 0.3, but Kaiser-Meyer-Olkin measure for three items was below 0.5; Awkward (0.439), Dangerous (0.474), and Scary (0.477). We thus ran the PCA excluding these three items, resulting in an overall Kaiser-Meyer-Olkin measure of 0.830. Bartlett's test of sphericity was statistically significant ($p = .000$).

PCA with Varimax orthogonal rotation revealed four components that had eigenvalues greater than one, which together explained 71.660% of variance (39.130%, 16.964%, 8.764%, and 6.801%). The interpretation of the data was in some respects consistent with the constructs that RoSAS intended to measure. The first component was very close to the construct 'Competence' from the RoSAS, with the main items loading onto it being Competent, Knowledgeable, Capable, Reliable, Responsive, and Interactive The second component was close to the construct 'Warmth' from the RoSAS, with the main items loading onto it being Compassionate, Feeling, Happy, Emotional, and Social. We did not find a single component that was close to the construct 'Discomfort' from the RoSAS, though the 3rd and 4th component seem to capture differing aspects of it. The 3rd component –with Responsive, Interactive, Social, not Aggressive, and not Awful loading onto it– we would label as 'Discomfort-Appropriateness'. The 4th component – with not Awful, Organic, and not Strange loading onto it– we would label as 'Discomfort-Weirdness'.

Hypothesis 1: Effects of PSI vs noPSI (noGAZE). A Mann-Whitney U test was run to determine if there were differences in stepping away and perception of the robot between PSI-noGAZE and noPSI-noGAZE (hypothesis 1). Distributions of the variables were similar, as assessed by visual inspection.

For moving away, as computed from the OptiTrack data, there was a statistically significant difference in the median score ($U = 44$, $z = -2.662$, $p = .008$, $n = 29$) between PSI-noGAZE (.83 m) and noPSI-noGAZE (.47 m). There was no statistically significant difference for stepping away as computed from the on-board sensors ($U = 60$, $z = -.693$, $p = .488$, $n = 24$), or for perception of the robot in terms of warmth ($U = 121$, $z = .353$, $p = .724$, $n = 30$), competence ($U = 152$, $z = 1.641$, $p = .101$, $n = 30$), discomfort-appropriateness ($U = 132.5$, $z = .832$, $p = .405$, $n = 30$), or discomfort-weirdness ($U = 121$, $z = .354$, $p = .723$, $n = 30$). We found the exact same pattern of results when including the participants who wrongly perceived the social cue, with a significant different median score on stepping away, as computed from the OptiTrack data ($U = 97$, $z = -2.427$, $p = .015$, $n = 38$) between PSI-noGAZE (.80 m) and noPSI-noGAZE (.48 m).

Hypothesis 2: Effects of Approach Distance x Social Cue. We conducted two-way ANOVAs to examine the effects of approach distance and social cue on our different measures. We suspect that our sample was a bit too small, as for some of these measures, we found outliers (as assessed by boxplot), residuals that were not normally distributed (assessed by Shapiro-Wilk), and no homogeneity of variances (Levene's test).

We found no statistically significant interaction effects of approach distance and social cue for any of our measures; not on our measures of stepping away, and neither on our measures of robot perception. Subsequent analysis of main effects also did not find a significant main effect for any of the constructs used to measure the perception of the robot.

We did find a statistically significant main effect of approach distance on stepping away as measured by OptiTrack $(F(1,53) = 8.070$, $p = .006$, partial $\eta^2 = .132)$ between PSI (mean $.796\,\text{m} \pm .070\,\text{m}$) and noPSI $(.521\,\text{m} \pm .071\,\text{m})$, as shown in Fig. 5b. No significant main effect was found for social cue $(F(1,53) = .051$, $p = .551$, partial $\eta^2 = .007)$.

We also found a statistically significant main effect of social cue on stepping away as measured by on-board sensors $(F(1,44) = 8.246$, $p = .006$, partial $\eta^2 = .158)$ between GAZE $(.060 \pm .013)$ and noGAZE $(.111 \pm .013)$, as seen in Fig. 5a. No significant main effect was found for approach distance $(F(1,44) = .034$, $p = .855$, partial $\eta^2 = .001)$.

To test our assumption that participants who did not see the cue should be excluded, we investigated if this effect on stepping away would also occur when comparing, within the subset of PSI-GAZE participants, people who perceived the social cue with people who did not. Though it should be noted that this was a very small sample $(n = 18)$, we did find a similar effect of perceiving the social cue $(.048 \pm .054)$ or not $(.161 \pm .117)$ on stepping away as measured by the on-board sensors $(U = 55$, $z = 2.218$, $p = .027)$.
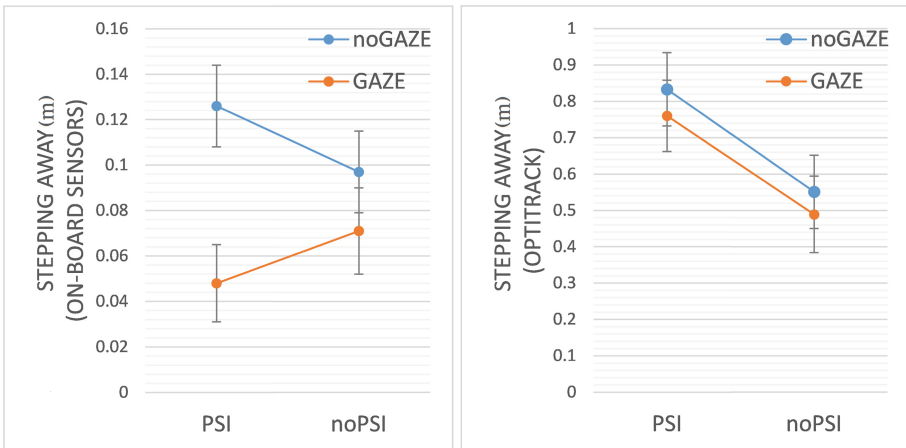
(a) Lower-body movement average (on-board sensors)

(b) Upper-body movement average (Opti-Track sensor)

Fig. 5. The statistical analysis of the sensor data. The values represent the average movement in meters.

The statistical results show that the social cue has significant effects on the positioning dynamic. The OptiTrack data shows that in the case of PSI, the

participants lean backwards (thus moving their head more) to ensure a comfortable interaction distance. Additionally, as shown by the on-board sensors data, when PSI is conducted and no social gaze cue is provided, there is a significant increase in the lower-body movement of the participants.

4.2   Qualitative Analysis

Our deductive analysis showed that the robot behavior had effects on the non-verbal reactions of our participants. However, much to our surprise, we found no particular effects on social perception of the robot as measured by the questionnaire. To further investigate if other aspects could have played a role, we performed an inductive analysis, based on the responses of the participants in the conducted semi-structured interviews. Additionally, we used this analysis to derive managerial recommendations for the robotic system and the tested social cue.

The selected analytical method is a mixture of hermeneutics, as proposed by [26]. This method allows for study interpretation and content analysis in order to establish categories which can be counted and linked. In order to overcome the problem of 'anecdotalism' regarding the results of the analysis, we used a 'triangulation of methods' technique [26]. Therefore, in our analysis we used a combination of the open questions from the questionnaires, interview transcripts *and* non-verbal behaviors of the respondents as observed from the videos. Additionally, we also used other available sources (positioning information from OptiTrack and on-board sensors, statistical interpretations, and the quantitative questionnaire data) to interpret the different findings.

When coding and analyzing the set of interviews, the starting point of the investigation were the three factors proposed by RoSAS scale: *Warmth*, *Competence* and *Discomfort*. However, because we systematically found cases that did not fit any of the three initial factors, new categories (emerging from the data) were created, in order to produce a holistic interpretation of all the phenomena, in a comprehensive data treatment [26]. As a result, the hermeneutics analytical approach led us to the inductive categories of *Incompetence*, *Problems and Solutions* and *Intentionality*.

We thus ended up with a total of six categories: *Warmth* represents the respondent's positive feelings regarding the interaction; *Discomfort* relates to the respondent's feeling of uneasiness regarding the interaction with the robot; *Competence* reflects the participant's feeling that the robot is considered to have the skills necessary to perform the task of the experiment; *Incompetence* is the opposing category, where the robot is considered as not having the appropriate skills; The category of *Intentionality* contains those quotations in which the respondent shows signs of appreciating specific intentions of the actions performed by the robot; The final category, *Problems and Solutions*, is a managerial theme where the respondents evaluate problems in the robot and give possible solutions for them. All these categories have an interpretative quality used to reflect how the participant's understanding and interpretation of the robot are influenced by the experiment.

Table 1. Frequencies of the different encoded categories from our inductive analysis, in each of the different conditions

Case	Warmth	Competence	Discomfort	Incompetence	Intentionality	Problems and Solutions
PSI-GAZE	7	8	38	17	5	11
PSI-noGAZE	13	7	39	23	3	14
noPSI-GAZE	24	26	33	12	12	18
noPSI-noGAZE	10	9	23	19	3	18

Because it follows a constructionist approach, the used triangulation method is limited by the fact that it has to accept the idea that an interview and its meaning is negotiated between the respondent and the interviewer and therefore accepts that different sources of data may bring different truths or interpretations, as discussed in [26]. This happens due to the fact that the truth generated through each interview is negotiated in its context and therefore its conclusions may not be repeated through a different source of data [27]. In order to address this limitation, additional measures where taken in order to ensure the validity of the results, as advised by [26,27]. These measures follow the refutability principle presented in [26] which seeks for critical thinking when analyzing data, using a constant comparative method. To this end, the analysis of the data focused initially on a smaller part, comprised of only a few of the conducted experiments and interviews, where their findings were systematically compared and discussed amongst the coders, in order to ensure that all the cases and any possible deviations are covered and explained accordingly in the established analysis [28].

After we established our categories, independent coders coded the data, and we measured their agreement to establish the reliability of our analysis [28]. For the first two coders, we found a very strong inter-rater agreement (Cohen's kappa = 0.81). We wish to report that when a third coder also coded 20% of the selected interview transcripts, the inter-rater agreement was much lower (Cohen's kappa = 0.45). This surprised us, though there are several possible explanations – e.g. it might have been caused by the third coder using only the transcripts. An alternative explanation is an effect of culture (similar to what was discussed in [29,30]), as the third coder lived in the Netherlands, while the first two coders both lived in Denmark. Given these results, including the strong inter-rater agreement between the first two coders, we discarded the codes of the third coder; we used only the codes of the first two coders for our further analysis.

In the following paragraphs we offer a detailed description of the outcomes of our analysis (see Table 1 for an overview).

Warmth: Regarding the perceived warmth of the interaction, the qualitative results show a clear impact of the social cue. This impact can be observed through the increased number of opinions regarding positive points on the interaction, that is then described as *nice*, *fine*, *natural* and *friendly*. When the PSI is added

to the interaction, the respondents see an incoherence of those signals and therefore the positive effect of the cue shrinks. Therefore, since the gaze can be viewed as an element that shows the robot's perception of the human, when the PSI occurs some participants interpret this as a malfunction of the robot, as stated in the following: *"Mhm … but the point that it is looking down, I think, I mean … I doubt it is working ok…"* (Interview 83)

Competence: Regarding 'competence', the results are conclusive regarding the social cue when it is not accompanied by PSI. The perceived skills relate to the speed, the closeness, the natural movement and the robot's ability to "see" the human. The expressed opinion regarding the skills and therefore competence of the robot triples. An example of this can be seen in the following extract: *"The closeness is fine. How it drives towards you, the speed… the distance was ok, the rest was ok."* (Interview 44)

Discomfort: When no social cue is signaled and there is PSI, the robot is perceived as to be lacking the necessary skills relating to the adequate closeness, speed, distance and in general is perceived as dangerous: *"So, it felt like a bit dangerous, that it did not perceived you well, so it can drive over you."* (Interview 13)

Incompetence: Other less objective faults are attributed to the robot when there is no social cue and there is PSI, for example the time in response is perceived as slower and as having the inability to properly locate the human: *"the robot approached me after a while, but he seemed to be kind of aiming a little bit next to me, so it wasn't really completely approaching and looking in the right direction."* (Interview 57)

Intentionality: A surprising result of the experiment is the fact that the social cue increases the impression of the robot as having intentions. The perceived intentionality decreases if there is PSI, but even in that case it scores higher than without the social cue. *"It felt like he then saw me, just standing there and turned towards me and came towards me and then turned a bit when it was in front of me to look more directly at my face."* (Interview 60)

Problems and Solutions: Some of the most relevant findings for the managerial implications in the cases in which the social cue was signaled are linked to the fact that the robot also displayed a task bar on the screen, which disrupted the image of the face of the robot: *"Well it really was a robot. I think the eye was not sophisticated enough to bring some human feelings, to me. Especially because there was also taskbar next do in his eyes, so. That kind of broke the image of a face"* (Interview 32).

Another common concern of the respondents was how to deploy the robot in a place with many people. The proposed solution, as shown in the following quote, is to ensure the minimum quantity of movements of the robot as to ensure its safety to a potential crowd surrounding it: *"… maybe like the fact that he could move like with you, but on the other hand if it gets too much, people will*

freak out, maybe. So it was properly good that it just stays there and doesn't do too many movements." (Interview 32)

For the cases in which no social cue is provided, the topics brought up by the respondents change. The speed and space become major topics, as they are perceived as more relevant problems to handle before worrying, for example about how the face is displayed on the screen. One possible suggestion for dealing with the PSI is adding a voice or a sound to the robot, as a form of signaling the interaction that is about to happen. One such example is given in: *"Like it really doesn't have any reasons to stand that close to my face. Like, if it at least have been talking to me or had another way of notifying me that it saw me and what it's intention was, maybe it would have been more ok. But now it was a bit random." (Interview 49)*

All in all, the analysis shows that, even though PSI is the main source of discomfort of the experiment carried out, the social cue decreases the level of discomfort of the participants. The main purpose attributed to the social cue is regarding safety, since the participants consider that the robot perceives their presence and therefore they trust in the fact that it will stop in an appropriate manner for the interaction. For example, the following quote reveals the concern of a participant that was not exposed to the social cue: *"I actually didn't have the idea that he was seeing me, so that was also the reason that I stepped aside. Because I thought: Ok he does not see me, he will just roll this way" (Interview 37)*. This indicates that a social gaze cue can be used to improve the perception of the overall interaction and safety of the robot.

5   Conclusions and Discussion

In this paper, we investigated if personal space invasion by a robot can be made more 'smooth' by introducing social gaze cues. We manipulated both personal space invasion and the use of social gaze cues, in a between-subjects set-up. To investigate the effects on the interaction dynamic, we measured distance to the robot, both of their lower body (using on-board sensors), and of their head (using a tracking system). To investigate the effects on the way in which the robot was perceived, we used a questionnaire with items on warmth, competence, and comfort.

We found significant main effects of our manipulations on the positioning dynamic; (1) participants increased their head distance after the approach more when the robot got close to them, i.e. conducted PSI, and (2) participants increased their lower body distance after the approach less when the robot gave a social gaze cue. This suggests that different aspects of people's positioning dynamic reflect different things. One interpretation, that was further supported by our inductive analysis, is that the social gaze cue increases peoples' feeling of safety, reducing the need to step back – and that people will lean back to ensure a comfortable interaction distance.

Given the richness of this kind of positioning data, and the promising results, a more comprehensive analysis of movement could be an interesting direction for

future work. We found it valuable to find that the different sensors we used, onboard and a tracker, yielded different aspects of stepping away, focused on lower body and head, respectively. Beyond stepping away after an approach, there is a range of other movement aspects that could be extracted, e.g. stepping away during an approach, or orientation of body or head. When doing so, it would be important to keep limitations such as sensor noise in mind, and to avoid looking for significant results (which is also why we deliberately limited our own analysis).

To our surprise, we found no significant effects of our manipulations on the questionnaire intended to measure the way in which participants perceived the robot – especially since the used approach distances were well above and below what is perceived as PSI in the literature, but also since we *did* find an effect on the interaction dynamic. A possible cause is that, during the experiment, the participants had sufficient space and time to adjust their position (by stepping back or to the side) and thus weaken the effect that the personal space invasion could have on them and their perception of the robot. This would be in line with our idea that aspects of the interaction dynamic can influence perception of PSI. Additionally, since our participants were not native in English, they may have missed some nuances in the meaning of different items, perceiving them as "similar" – for example Capable/Competent or Strange/Awkward. This would of course have an impact on the "strength" of these descriptors.

To further explore the perception of the robot, we conducted a qualitative analysis based on the semi-structured interviews, the open questions from the questionnaire, and the video recordings. We identified various categories in which the comments of our participants could be encoded, including discomfort and intentionality. While our findings were rich and qualitative (see Sect. 4.2), we will here discuss one specific pattern; when the social cue was signaled, there was a sense among participants that the robot was aware of their presence, and, as a result, participants would trust its actions more – also when it conducted PSI.

Overall, our findings show that social cues might indeed be used to make personal space invasion by a robot more smooth, especially in terms of the interaction dynamic and perceived safety.

References

1. Walters, M.L., Dautenhahn, K., Boekhorst, R.T., Koay, K.L., Syrdal, D.S., Nehaniv, C.L.: An empirical framework for human-robot proxemics. In: Proceedings of New Frontiers in Human-Robot Interaction (2009)
2. Takayama, L., Pantofaru, C.: Influences on proxemic behaviors in human-robot interaction. In: IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2009, pp. 5495–5502. IEEE (2009)
3. Mead, R., Matarić, M.J.: Robots have needs too: people adapt their proxemic preferences to improve autonomous robot recognition of human social signals. In: New Frontiers in Human-Robot Interaction, p. 100 (2015)

4. Aiello, J.R.: Human spatial behavior. In: Stokols, D., Altman, I. (eds.) Handbook of Environmental Psychology, Chap. 12, pp. 389–504. Wiley, New York (1987)
5. Vroon, J., Englebienne, G., Evers, V.: Responsive social agents. In: Agah, A., Cabibihan, J.-J., Howard, A.M., Salichs, M.A., He, H. (eds.) ICSR 2016. LNCS (LNAI), vol. 9979, pp. 126–137. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-47437-3_13
6. Hall, E.T.: The Hidden Dimension. Anchor Books, New York (1966)
7. Kendon, A.: Conducting Interaction: Patterns of Behavior in Focused Encounters, vol. 7. CUP Archive (1990)
8. Groh, G., Lehmann, A., Reimers, J., Frieß, M.R., Schwarz, L.: Detecting social situations from interaction geometry. In: 2010 IEEE Second International Conference on Social Computing (SocialCom), pp. 1–8. IEEE (2010)
9. Lau, B., Arras, K.O., Burgard, W.: Multi-model hypothesis group tracking and group size estimation. Int. J. Soc. Robot. **2**(1), 19–30 (2010)
10. Kastanis, I., Slater, M.: Reinforcement learning utilizes proxemics: an avatar learns to manipulate the position of people in immersive virtual reality. ACM Trans. Appl. Percept. **9**(1), 3:1–3:15 (2012)
11. Jung, M.F., Lee, J.J., DePalma, N., Adalgeirsson, S.O., Hinds, P.J., Breazeal, C.: Engaging robots: easing complex human-robot teamwork using backchanneling. In: Proceedings of the 2013 Conference on Computer Supported Cooperative Work, pp. 1555–1566. ACM (2013)
12. Guizzo, E., Ackerman, E.: The rise of the robot worker. IEEE Spectr. **49**(10), 34–41 (2012)
13. Fischer, K., Jensen, L.C., Suvei, S.-D., Bodenhagen, L.: The role of the timing between multimodal robot behaviors for joint action. In: 3rd Workshop on Joint Action (at ICSR 2015) (2015)
14. Basili, P., Huber, M., Brandt, T., Hirche, S., Glasauer, S.: Investigating human-human approach and hand-over. In: Ritter, H., Sagerer, G., Dillmann, R., Buss, M. (eds.) Human Centered Robot Systems. Cognitive Systems Monographs, vol. 6, pp. 151–160. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-10403-9_16
15. Carpenter, J., Davis, J.M., Erwin-Stewart, N., Lee, T.R., Bransford, J.D., Vye, N.: Gender representation and humanoid robots designed for domestic use. Int. J. Soc. Robot. **1**(3), 261–265 (2009)
16. Kuchenbrandt, D., Eyssel, F., Bobinger, S., Neufeld, M.: When a robot's group membership matters. Int. J. Soc. Robot. **5**(3), 409–417 (2013)
17. Fiske, S.T., Cuddy, A.J.C., Glick, P.: Universal dimensions of social cognition: warmth and competence. Trends Cogn. Sci. **11**(2), 77–83 (2007)
18. Willis, J., Todorov, A.: First impressions: making up your mind after a 100-ms exposure to a face. Psychol. Sci. **17**(7), 592–598 (2006)
19. Ho, C.-C., MacDorman, K.F.: Revisiting the uncanny valley theory: developing and validating an alternative to the godspeed indices. Comput. Hum. Behav. **26**(6), 1508–1518 (2010)
20. Carpinella, C.M., Wyman, A.B., Perez, M.A., Stroessner, S.J.: The robotic social attributes scale (RoSAS): development and validation. In: Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction, pp. 254–262. ACM (2017)
21. Bartneck, C., Kulić, D., Croft, E., Zoghbi, S.: Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. Int. J. Soc. Robot. **1**(1), 71–81 (2009)

22. Shiarlis, K., Messias, J., van Someren, M., Whiteson, S., Kim, J., Vroon, J., Englebienne, G., Truong, K., Evers, V., Pérez-Higueras, N., Pérez-Hurtado, I., Ramon-Vigo, R., Caballero, F., Merino, L., Shen, J., Petridis, S., Pantic, M., Hedman, L., Scherlund, M., Koster, R., Michel, H.: Teresa: a socially intelligent semi-autonomous telepresence system. In: ICRA 2015: Proceedings of the IEEE International Conference on Robotics and Automation, Workshop on Machine Learning for Social Robotics. IEEE (2015)
23. D6.5: Third version of semi-autonomous telepresence system. Teresa project deliverable (2016). teresaproject.eu/project/deliverables/
24. Thrun, S., Leonard, J.J.: Simultaneous localization and mapping. In: Siciliano, B., Khatib, O. (eds.) Springer Handbook of Robotics, pp. 871–889. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-30301-5_38
25. Dellaert, F., Fox, D., Burgard, W., Thrun, S.: Monte carlo localization for mobile robots. In: Proceedings of the 1999 IEEE International Conference on Robotics and Automation, vol. 2, pp. 1322–1328. IEEE (1999)
26. Silverman, D.: Doing Qualitative Research: A Practical Handbook. SAGE Publications Limited, Thousand Oaks (2013)
27. Russell Bernard, H.: Research Methods in Anthropology: Qualitative and Quantitative Approaches. Rowman & Littlefield, New York (2011)
28. Russell Bernard, H., Wutich, A., Ryan, G.W.: Analyzing Qualitative Data: Systematic Approaches. SAGE Publications, Thousand Oaks (2010)
29. Antaki, C., Rapley, M.: 'Quality of life' talk: the liberal paradox of psychological testing. Discourse Soc. **7**(3), 293–316 (1996)
30. Sveningsson, S., Alvesson, M.: Managing managerial identities: organizational fragmentation, discourse and identity struggle. Hum. Relat. **56**(10), 1163–1193 (2003)