Towards Robot Autonomy in Group Conversations: Understanding the Effects of Body Orientation and Gaze

Marynel Vázquez [1], Elizabeth J. Carter [2], Braden McDorman [2], Jodi Forlizzi [2,3], Aaron Steinfeld [1], Scott E. Hudson [2,3]

[1]Robotics Institute
Carnegie Mellon University
Pittsburgh, PA
{marynel,steinfeld}
@cmu.edu

[2]Disney Research Pittsburgh
4720 Forbes Avenue
Pittsburgh, PA
{liz.carter,braden.mcdorman}
@disneyresearch.com

[3]HCI Institute
Carnegie Mellon University
Pittsburgh, PA
{scott.hudson,forlizzi}
@cs.cmu.edu

ABSTRACT

We conducted a 2×2 between-subjects experiment to examine the effects of two orientation and two gaze behaviors during group conversations for a mobile, low degree-of-freedom robot. For this experiment, we designed a novel protocol to induce changes in the robot's group and study different social contexts. In addition, we implemented a perception system to track participants and control the robot's orientation and gaze with little human intervention. The results showed that the gaze behaviors under consideration affected the participants' perception of the robot's motion and that its motion affected human perception of its gaze. This mutual dependency implies that robot gaze and body motion must be designed and controlled jointly, rather than independently of each other. Moreover, the orientation behaviors that we studied led to similar feelings of inclusion and sense of belonging to the robot's group, suggesting that both can be primitives for more complex orientation behaviors.

Keywords

Social Human-Robot Interaction; Robot Motion; Gaze

1. INTRODUCTION

The idea that robots will interact around and with human groups has recently motivated significant research efforts within Human-Robot Interaction (HRI). For example, algorithms for group detection [34, 66, 47] and engagement classification [33] have been developed. Planning for navigation has begun to consider the social costs of crossing between group members [48, 13]. Experimental evidence further suggests that people establish similar spatial arrangements with robots as they do with people during free-standing group conversations [53, 76]. People position and orient themselves to maximize their opportunities to monitor each other as well as maintain their group as a spatially distinct unit from other nearby interactions. These spatial arrangements are known as *F-Formations* [29] and often emerge as a circle during conversations among multiple people (Fig. 1(a)).

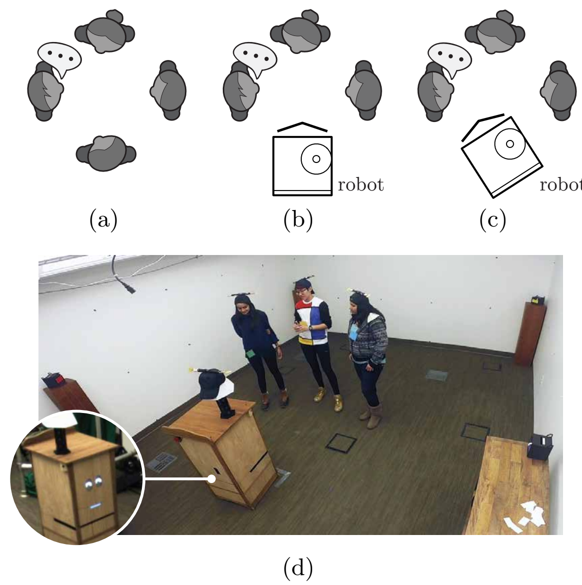

Figure 1: (a) Group conversation. (b) One member is a robot and it orients towards the group center. (c) The robot orients towards the speaker. (d) HRI experiment where we tested these orientation behaviors along with two gaze behaviors.

Although we know that a robot's body orientation may induce reconfigurations of dyadic F-formations [31], little is known about its effects during conversations with more people. How should robots orient themselves with respect to a group during a conversation? At first glance, one may think that a robot should mimic human behavior and orient its body towards the middle of the group (Fig. 1(b)). This approach was previously proposed by Althaus and colleagues [5] and was associated with more positive perceptions of the behavior of a telepresence robot [69]. However, this strategy is not the only reasonable one for mobile, low degree-of-freedom (DoF) robots. For robots with a fixed head, such as FROG [18], SPENCER [64], or the CoBots [68], it may be better to orient towards the focus of attention of the conversation, e.g., the speaker, as illustrated in Fig. 1(c). This behavior could convey attentiveness to the interaction and make users perceive the robot as more active or responsive.

HRI '17, March 06–09, 2017, Vienna, Austria

ISBN 978-1-4503-4336-7/17/03...$15.00

DOI: http://dx.doi.org/10.1145/2909824.3020207

To further our understanding of robot positioning during group conversations, we conducted an experiment in which a mobile robot interacted with small groups of people (Fig. 1(d)) while we manipulated its orientation and gaze behaviors. Gaze was manipulated because the robot that we used has a fixed head, but the direction of its eyes could still communicate mental states and attention [7, 2]. We expected gaze to affect the perception of the orientation of the robot.

Another contribution of this paper is a new protocol for studying group interactions in the context of HRI. In this protocol, the robot asks the participants to help it solve a problem in a brainstorming session in the laboratory. Even though this is not a public setting, our design makes the interaction naturalistic. Participants are free to move in the environment as desired and, periodically, are induced to leave the robot's conversational group to document their ideas. This dynamic creates a variety of group formations on a frequent basis, thereby generating numerous instances for studying multi-party interactions. For example, groups with one to four people emerged during our experiment as a result of the flow of the activity. In addition, the proposed brainstorming protocol does not require us to provide the participants with specific instructions on their roles. This property leads to increased interaction time during experiments in contrast to a prior role-playing game protocol that was used to study group interactions with a social robot [65].

The body and gaze behaviors of the robot were controlled automatically by a multi-modal perception system during most of the the brainstorming activity. This system relied on several off-the-shelf sensors and data fusion techniques to (1) track the users and the robot, (2) detect conversational groups by reasoning about spatial behavior, and (3) detect the current speaker in the environment. Although the core components of this system were developed previously, their integration is valuable. First, it allowed us to run parts of the experiment in an automated fashion. Second, it allowed us to collect a corpus of human spatial behavior with and around our robot, which we are using to develop more complex perception capabilities and evaluate group detection methods. We describe our system and its limitations for future replication and adaptation efforts.

2. RELATED WORK

Proxemics. Many factors can influence how people use physical spaces during interactions [22]. In HRI, these factors include social norms, familiarity among the interactants, users' personal characteristics, mutual gaze, a robot's voice and height, and the direction from which a robot approaches users [41, 61, 70, 73, 72, 71].

Proxemics have been used as a cue to estimate users' engagement level with robots [39, 40, 51, 54, 26]. For example, Michalowski and colleagues [39] proposed an engagement model based on spatial information and head pose. Satake and colleagues [51] used motion to estimate if users would accept interacting with a robot. In this work, we used position and body orientation features to detect F-formations [66] and infer who was conversing with our mobile platform.

Initiating interactions. Different strategies and motion controllers have been developed to appropriately approach users and initiate interactions [5, 27, 49, 51, 53]. In particular, the controller of Althaus and colleagues [5] reduced the speed of the robot as it approached a group of users. This controller also adjusted the robot's orientation towards the middle of the group. In our experiment, a wizard teleoperated our robot at the beginning of the interaction to achieve a similar result and orient the mobile platform towards the participants. Once the brainstorming activity started, the wizard switched to automatic orientation control, as further explained in Sections 3 and 4.

Robot orientation. The middle orientation behavior that we studied in our experiment was proposed earlier for social navigation [5]. Karreman and colleagues [26] implemented this behavior on a museum guide robot that gave short tours to visitors. Turning the robot towards visitors led to increased interest in the platform in contrast to turning towards points of interest, like art pieces. There is also evidence that suggests that orienting a telepresence robot towards the center of a group makes people comfortable [69].

Social gaze. There is significant work in social eye gaze for human-computer and human-robot interaction [50, 2]. Related to our work, Garau and colleagues [20] found that synchronizing an avatar's head and eye animations with turn-taking patterns could improve its communication with humans in comparison to a random gaze behavior in which its head and eye animations were unrelated to conversational flow. As in our experiment, random gaze behaviors were also used in prior efforts to study robot gaze. For example, Yoshikawa and colleagues [75] compared a random gaze behavior versus three other gaze behaviors on a Robovie-R2 platform. Their experiment suggests that responsive robot gaze, e.g., gaze that communicates shared attention, induces stronger feelings of being looked at on users in comparison to non-responsive gaze. In addition, Skantze and colleagues [58] studied a random gaze behavior versus a human-inspired gaze behavior on a Furhat robot. This robot has back-projected eyes like the platform used in this work.

Other research has focused on analyzing gaze duration and frequency. For example, prior work [1] suggests that short, frequent fixations by a robot can give an observer stronger feelings of being looked at versus longer, less frequent stares. Also, a robot that looks towards users more often may be perceived as more extroverted than to one that looks more towards the task space [6]. Note that gaze can also influence people's roles in a conversation with a robot [30, 43] and their attitudes towards these machines [25]. Some gaze behaviors may work better than others, depending on the type of conversation [14].

Sense of groupness. Several efforts within HRI have investigated how much people perceive themselves as part of a group [23, 37, 43, 46]. Similar to prior work, we follow the approach of Mutlu and colleagues [43] to measure interpersonal closeness to our robot with the "Inclusion of Other in Self" (IOS) scale [8]. We use the survey by Williams and colleagues [74] to measure feelings of groupness and ostracism.

Multi-modal perception. Our perception system was inspired by work in multi-modal sensing [32, 10, 28, 56, 57, 63, 11, 44] and is an alternative to other approaches meant to enable human-robot interactions in controlled settings. In particular, our system estimates users' positions and body orientations by fusing ultra wide-band tracking information and skeleton data output by a Kinect. Even though prior work used ultra wide-band localization systems to track people [21, 9] or the Kinect to enable interactions [38, 3, 77, 24], we are the first to fuse these types of data for HRI to the best of our knowledge. The fusion offers key advantages: operation beyond the Kinect's range, better occlusion handling, and simple user identification.

Our perception system also builds on advances in localization [62] and human spatial analysis [29, 15, 66]. While recent efforts to detect social interactions based on spatial behavior have focused on analyzing users only [36, 34, 47], we opt to jointly reason about the users' and our robot's spatial configurations in a unified perspective.

3. ORIENTATION AND GAZE BEHAVIORS

We studied two orientation and two gaze behaviors during group conversations with the furniture platform Chester [67]. This robot has a differential drive base, a fixed face, and back-projected eyes, as shown in Fig. 1(d). Even though the robot's design led to specific decisions for the orientation and gaze behaviors, they can be easily adapted to other mobile platforms with expressive eyes. We detail our implementation to facilitate future explorations in this direction.

For the following explanations, assume that the robot has started a conversation and we know its position $\mathbf{r} = [r_x\ r_y]^T$ and orientation ρ (yaw angle) on the ground. Assume that we also know the position $\mathbf{p}^i$, the lower body orientation, and the velocity of any person i near the robot, so that we can detect its conversational group by reasoning about F-formations (e.g., using [15, 52, 66]). Finally, assume that we know who is speaking in the robot's conversation. Data collection methods are later described in Sec. 3.3.

3.1 Body Orientation Behaviors

For any member i in the robot's conversation, let $\mathbf{u}^i = [u_x^i\ u_y^i]^T = \mathbf{p}^i - \mathbf{r}$ be the direction from the robot to this person, and $\gamma^i = \text{atan2}(u_y^i, u_x^i)$ the corresponding angle. We used this angle to orient the robot as described below.

3.1.1 *Middle Orientation Behavior (MO)*

The robot oriented towards the middle of its conversational group G using the *mean direction* $\bar{\theta}$ of all γ^i [19]:

$$\bar{\theta} = \text{atan2}\Big(\sum_{i \in G} \sin(\gamma^i), \sum_{i \in G} \cos(\gamma^i)\Big) \tag{1}$$

3.1.2 *Attentive Orientation Behavior (AO)*

If the robot was speaking, it biased its orientation towards its addressee; otherwise, it biased its orientation towards the current speaker in its conversational group. Let γ^i be the orientation towards the speaker or the addressee, and $\bar{\theta}$ be the middle orientation in the group, as in eq. (1). At any given time, the orientation $\hat{\rho}$ of the robot was set as follows:

$$\hat{\rho} = \begin{cases} \bar{\theta} - \tau & \text{if } \text{minAngDiff}(\gamma^i, \bar{\theta}) < -\tau \\ \bar{\theta} + \tau & \text{if } \text{minAngDiff}(\gamma^i, \bar{\theta}) > \tau \\ \gamma_i & \text{otherwise} \end{cases} \tag{2}$$

where minAngDiff returns the signed minimum difference between two angles, and τ is a parameter that controls how much the robot rotates away from the middle orientation $\bar{\theta}$ (Fig. 2). In particular, we set $\tau = 60^\circ$ for our robot so that it would not turn its back to group members to its side.

If the robot was not addressing anyone and nobody had spoken for a significant time (10 seconds), the platform's orientation was set towards the middle direction as in the MO behavior. This also happened when the robot conversed with a single user, given that $\hat{\rho}$ in eq. (2) became $\bar{\theta}$.

3.2 Gaze Behaviors

We tested simple gaze behaviors to complement the effects of our orientation manipulation. These behaviors serve as a baseline for future investigations on the relationship between body motion and complex gaze patterns, e.g., involving discourse structure or fixations on the environment [12, 42].

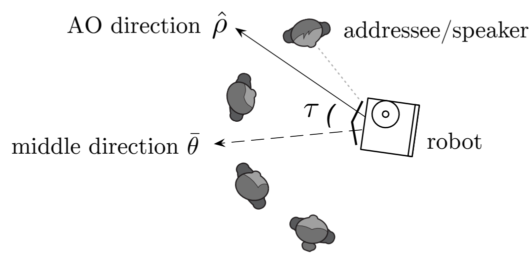

Figure 2: Geometric relations for the AO behavior. The $\wedge$ mark denotes the robot's front. The middle direction corresponds to eq. (1) and $\hat{\rho}$, τ to eq. (2).

Gaze was calibrated using a projective mapping from 3D world coordinates to 2D pupil positions. The mapping was used for both pupils of the robot, making their lines of sight parallel. While this constraint prevented vergence eye movements, it worked well in practice because the robot's eyes look cartoonish and have a slight curvature. This design makes users forgiving of gaze patterns that do not fully mimic human gaze and induces the Mona Lisa gaze effect [4]: users perceive mutual gaze more often than intended.

3.2.1 *Random Gaze Behavior (RG)*

The robot executed several pre-defined eye animations that helped communicate ideas while it spoke (Fig. 3). For example, referential gaze was used at times with verbal utterances to convey spatial information. When no pre-defined animation was scheduled for the eyes, they blinked occasionally or their pupils moved randomly at small intervals.

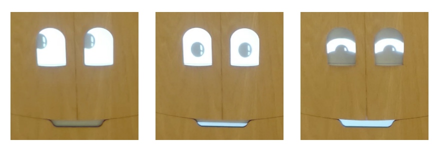

Figure 3: Left to right: eyes look up and to the right of the robot, look forward, and squint.

Our specific implementation of eye blinks was inspired by human blinking activity [16]. The duration of inter-blink intervals followed a normal distribution $\mathcal{N}(5.2, 3^2)$ in seconds.

Gaze shifts were scheduled by sampling time intervals in seconds from the uniform distribution $Unif(1.8, 3)$. When the timer triggered and no blink was set to occur, the pupils moved a small amount horizontally (d_x) and vertically (d_y), based on the size of the eyes:

$$d_x = eye\_width * \epsilon_1 \text{ and } d_y = eye\_height * \epsilon_2$$

with ϵ_1 and ϵ_2 sampled uniformly in a small interval. Any displacement (d_x, d_y) that rendered the pupils outside the limits of the eyes was considered invalid and was re-computed by sampling new values. Furthermore, we prevented Chester from fixating significantly downwards, towards the ground, so that it would not look extremely introverted.

3.2.2 *Attentive Gaze Behavior (AG)*

The robot used the same blinking pattern and pre-defined eye animations as in RG. When no animation was scheduled, the robot attempted to establish mutual gaze with the person who was the focus of attention. That is, the person that the robot addressed in particular, the current speaker if the robot was quiet, or anybody who moved with a speed of at least 0.5 m/s in the group when everybody was silent.

Once the robot gazed towards someone, gaze shifts were sampled as often as in RG but were biased towards the head of the focus of attention. If $\mathbf{q}$ is the 3D position of the head, then the new, biased positions for the pupils were set as:

(x,y) = lookAt($\mathbf{q}$) `// pupils position towards` q
$r \sim Unif(0, 1)$
if $r < 0.2$ **then** `// add noise` 20% `of the time`
  $x = x + eye\_width * \alpha_1$ with $\alpha_1 \sim \mathcal{N}(0, \sigma^2)$
  $y = y + eye\_height * \alpha_2$ with $\alpha_2 \sim \mathcal{N}(0, \sigma^2)$

where lookAt returned the 2D location of the pupils that made the robot look towards the desired direction, and σ controlled the amount of variation in gaze shifts. After 10 seconds of silence and no significant motion in the group, gaze shifts continued without the bias as in RG.

3.3 Multi-Modal Perception System

We implemented a real-time system to control the robot's orientation and gaze based on human behavior, as well as to collect data during the experiment. The system required instrumenting the environment with ultra wide-band (UWB) localization beacons[1] and a Kinect. Each participant wore an instrumented baseball cap with two UWB radio beacons for tracking and identification, as shown in Fig. 1(d). The robot also wore a cap to make it look like the participants.

3.3.1 System Components

Figure 4 shows the main components of the system. Grey boxes denote modules that ran on the robot; the rest executed on external computers. The boxes with thicker edges correspond to modules that were in charge of the manipulated behaviors. Note that the robot's speech was controlled by a hidden operator, as detailed in Sec. 4.1.

The system processed data as follows. First, the position of the UWB beacons carried by the participants was smoothed with a Kalman filter ("Filter" module in Fig. 4). The smoothed values were then aggregated to estimate the position and orientation of each hat ("Hat Pose" module) and fused with the skeleton output of a Kinect ("User Tracker" module). This fusion step output estimates of the position and orientation of each participant, taking advantage of both sensing modalities. The Kinect reduced localization error, which ranged up to 30 cm on average for the hats. The UWB

[1] We used DWUSB sensors by Ciholas, Inc.

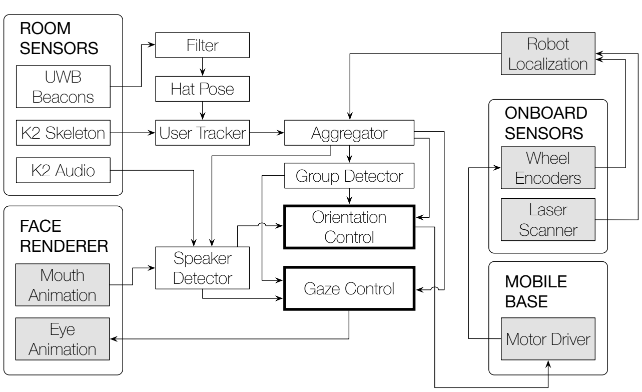

Figure 4: The system used to control the body orientation and gaze of the robot. "UWB" stands for ultra wide-band and "K2" stands for Kinect v2.

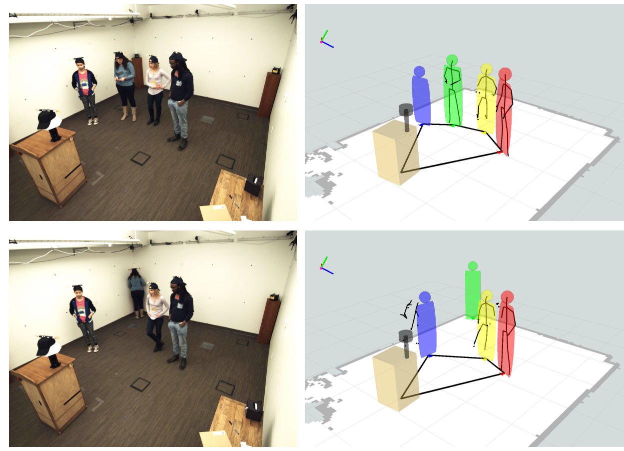

Figure 5: Experiment (left) and outputs of our perception system (right). A Kinect in the left corner of the room output skeleton data (shown in black near the participants). Colored markers denote participants' pose as output by the "User Tracker" module. The black lines on the ground connect the estimated members of the robot's conversational group.

data provided continuous tracking information throughout the environment as well as participants' identities.

While users were localized, the "Robot Localization" module estimated Chester's pose using an on-board laser scanner and a map of the environment [62]. The "Aggregator" program then grouped all this localization information and passed it to the "Group Detector" and "Speaker Detector" modules. The former module reasoned about conversational groups based on F-formations, as in [66] and illustrated in Fig. 5. The latter module was in charge of identifying the current speaker based on the interactants' positions, sound detections (output by "K2 Audio"), and information from Chester's dialog engine ("Mouth Animation" module). If Chester's mouth was moving, the robot was identified as the current speaker. Otherwise, the speaker was the person closest to the Kinect's audio beam (within 1 m) or nobody when no sound was localized.

Finally, the locations of the participants, the conversational groups, and the identity of the speaker were sent to the "Orientation Control" and "Gaze Control" modules. These programs output motion and gaze commands for the robot.

3.3.2 Limitations

Our system is a practical contribution of this work because it can enable human-robot interactions with little human intervention. However, it does not solve all perception problems in HRI, e.g., because it requires instrumentation and this may be impossible or undesirable in some cases.

Two types of errors due to shortcomings of the underlying technologies were the main factors that influenced the system's performance. First, tracking errors were common at the edges of the room due to the Kinect's limited range and field of view as well as a noticeable bias that affected UWB localization in these regions. As discussed in Sec. 4.6, these errors rarely affected group detection and the robot's orientation during the experiment because the participants were usually in the middle of the space. Second, sound localization errors were typically caused by simultaneous speech. These events were also infrequent in our experiment as interactants respected turn-taking.

4. METHOD

4.1 Study Design and Setup

We designed a 2×2 between-subjects experiment to test orientation (middle vs. attentive) and gaze (random vs. attentive) behaviors. The experiment followed a *Wizard with Oz* arrangement [59] in which the manipulated behaviors were autonomous, but the sequencing of events within the study and the robot's speech were managed by a hidden operator or "wizard". In a few instances, the wizard also re-configured the robot spatially with respect to the participants, as detailed in Sec. 4.6. The experiment was approved by our Institutional Review Board.

During the experiment, the robot led and participated in a brainstorming session with a small group of participants. Each session was performed under one of four conditions:

MO+RG condition. The robot oriented towards the middle of its conversational group and randomized its gaze.

AO+RG condition. The robot biased its orientation towards the focus of attention and randomized its gaze.

MO+AG condition. The robot oriented towards the middle and tried to establish mutual gaze with the person who was the focus of attention.

AO+AG condition. The robot biased its orientation and gaze towards the focus of attention.

Given these conditions, we hypothesized that:

H1. The gaze behaviors would affect the perception of the robot's motion, with AG increasing perceived naturalness.

H2. For the AO behavior, participants would find the robot more attentive and responsive than MO.

H3. The AO behavior would make the participants feel like the robot was more of a part of their group than MO.

H4. The AO+AG condition would lead to reduced feelings of ostracism or increased feelings of inclusion compared to MO+RG.

The experiment was conducted in a room with a free space of 4.4×4.4 meters (Fig. 1(d)). A table was placed adjacent to a wall for the participants to write down the brainstormed ideas on slips of paper. These slips then had to be deposited in different boxes in the room, according to the author.

The room was equipped with a UWB sensor network, a Kinect v2 and four RGB cameras near the ceiling. The UWB sensors and the Kinect were used to localize the participants, identify them, and detect speakers, as described in Sec. 3.3. The cameras recorded the interaction from multiple views and allowed the wizard to monitor the experiment remotely.

4.2 Participants

We recruited 20 groups (5 per condition) of 3 or 4 people using a participant pool, word of mouth, and fliers. The participants were at least 18 years of age, fluent in English, and had grown up in the U.S. The last restriction was imposed to reduce the effects of cultural biases in spatial behavior.

Table 1 shows details of the 69 participants that interacted with our robot. In general, most participants were university students, and their average age was 24.8 years old (SE = 1.0). In 7 sessions, two or more participants knew each other.

Before the interaction, the participants indicated how often they used a computer and their familiarity with robots on a 7 point Likert responding format (1 being lowest). Most participants used computers daily (M = 6.97, SE = 0.02) but were not very familiar with robots (M = 3.38, SE = 0.20).

Table 1: Participant characteristics per condition. "G", "F", "M", and "P" are used to abbreviate groups, female, male, and participants, respectively.

Condition	#G	#F	#M	#P	Age (Std Err)
MO+RG	5	8	10	18	22.2 (0.8)
MO+AG	5	9	9	18	23.5 (0.8)
AO+RG	5	11	5	16	24.4 (1.3)
AO+AG	5	9	8	17	29.3 (3.6)

4.3 Procedure

First, an experimenter gave a colored badge to each participant for identification purposes and administered a demographics survey. She then asked the participants to wear instrumented baseball caps with UWB beacons, and explained that each of them had a box in the room with their same color identifier. The experimenter introduced the robot, gave it an instrumented cap to make it look like the participants, and stepped away. The robot opened its eyes, and started a semi-scripted conversation with three phases:

1. *Introduction.* Chester presented itself to the group. The robot explained that the laboratory wanted to retire him, but people might keep him around if they found him useful. Chester encouraged the participants to think of how it could help in the lab and explained its sensors and capabilities. To facilitate brainstorming, the robot provided a first example and explained how it delivered souvenirs to lab visitors in the past. Chester then opened the floor to new ideas.

2. *Brainstorming.* The robot encouraged the group to brainstorm tasks that it could do in the lab for 6 min. Chester replied favorably to useful ideas and requested that authors write them on a slip of paper and deposit the slip in their corresponding box. The robot also asked for more details or discouraged unrealistic and complicated tasks. When people ran out of ideas, Chester provided more suggestions.

3. *Closing.* Chester asked a participant to count the ideas in the boxes and write the color of the box on each slip to help keep track of them. Meanwhile, the robot asked other people about their favorite ideas and gave his opinion. Chester thanked everybody for helping and said good-bye.

Finally, the experimenter adminstered a post-test survey, paid the participants, and debriefed them about the wizard. During debriefing, the experimenter also explained that the requests to deposit paper slips on boxes were an excuse to induce people to leave the robot's conversation and re-enter in natural ways. These requests were motivated by our prior experience, where we found that we had little chance of observing varied spatial behaviors without a task like this one.

4.4 Dependent Measures

We considered subjective and objective measures. The post-test survey asked people about their impressions of:

– the robot's motion and gaze;
– closeness to the robot using the IOS scale [8];
– the robot's and the participants' feelings of belongingness and ostracism in the brainstorming group [74];
– Chester with respect to a set of attributes, e.g., perceived intelligence, responsiveness, and entertainment value;
– Chester's ability to lead the brainstorming session and whether it should be decommissioned; and
– any unusual behavior for the robot [55].

Objective measures included the distance that the participants kept from the robot, the participants' membership in the robot's conversational group, and the number of paper slips collected during the brainstorming activity.

4.5 Pilot Sessions

Before starting the experiment, we recruited 35 people to conduct two types of pilot sessions. First, we ran 3 human-only sessions to evaluate the dynamics of the brainstorming activity and collect example tasks for the robot. Second, we ran 8 human-robot pilot sessions to test the Chester's dialog and the manipulated behaviors. During these sessions, we also simplified the wizard's teleoperation interface and the protocol of the experiment to avoid confusing procedures.

We considered studying a random orientation behavior for our robot as a baseline. However, the pilot sessions quickly showed that people are highly sensitive to inappropriate or unexpected orientations. These motions often halted interactions because people did not know how to interpret them.

4.6 Confirmation of Autonomy and Behaviors

The robot moved autonomously for most of the interaction as defined by the experimental condition. The exceptions were (1) when the robot started conversing, (2) when it said good-bye, and (3) during a handful of situations due to technical difficulties. In the first case, the wizard reconfigured the robot to show that it could move and tacitly induce an F-formation. In the second, the wizard moved Chester away to end the interaction. In the third, the wizard corrected for slight undesired changes in the robot's orientation, e.g., because of people-tracking failures in our perception system. During the brainstorming phase – the main part of the experiment – sporadic reconfigurations of this sort happened in 16 sessions out of 20. In these sessions, total teleoperation time while brainstorming was 9.37 sec on average (SE = 2.04), which represented only 2.4% of the duration of this phase (M = 383.53 sec, SE = 5.43, N = 16). REstricted or REsidual Maximum Likelihood (REML) analyses [45, 60] on the number of teleoperation events and teleoperation time while brainstorming showed no significant differences for the effects of Orientation (Attentive, Middle) and Gaze (Attentive, Random).

To confirm that the robot oriented as expected during the brainstorming phase, two coders annotated the members of the robot's conversation.[2] Using this ground truth and the logs from our perception system, we then computed the *ideal* middle orientation of the robot at 1 Hz. As expected, the absolute angular difference between the robot's orientation and this ideal middle direction was smaller for MO (M = 7.04°, SE = 0.12, N = 3686) than for AO (M = 14.33°, SE = 0.27, N = 3644). Note that these differences were induced in part by the robot's motion planner and a small bias in the robot's orientation towards the table in the room. The planner prevented Chester from jittering by ignoring turns of 5° or less. The bias (1.63° on average) was generated by occasional user tracking errors that made Chester believe that some people were still conversing with it when they left to write paper slips. Interestingly, the motion induced by these errors was interpreted as though Chester was checking that participants were following its instructions.

We transcribed when people spoke in the robot's group and its specific addressees in 2 sessions per condition. We then used the data to check when Chester adjusted its orientation towards these people. As expected, the robot turned more towards these foci of attention with AO (47% of 280 annotated events) than with MO (25% of 319). The robot did not move in many cases because the target was within 5° of its orientation (22% of the events for AO; 21% for MO).

We also inspected Chester's eye fixations during the experiment to confirm that the gaze behaviors worked as expected. As can be seen in Fig. 6, the positions of the pupils were less concentrated for RG than for AG because the robot tried to establish mutual gaze with the focus of attention in the latter case. Also, the robot had a tendency to look forward because several of our pre-defined eye animations positioned the pupils towards the middle of the eyes.

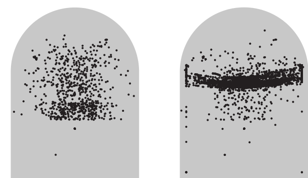

Figure 6: Chester's eye fixations in the 20 sessions of the experiment.

5. RESULTS

We first analyse survey results and the spatial behavior of the participants around our robot. Then, we discuss the implications of our findings in terms of our hypotheses.

5.1 Survey Results

We ran REML analyses to evaluate survey responses. Unless noted, analyses used Participant as a random effect nested by Session, and Orientation (Attentive, Middle), Gaze (Attentive, Random), and Gender as main effects. Student's t-tests and Tukey HSD tests (with significance thresholds of $p < 0.05$) were used for post-hoc analyses of two sample and multiple pair-wise comparisons, respectively. Ratings were on 7-point Likert responding formats and responses were grouped only when Cronbach's alpha was above 0.7.

Robot's gaze. In general, Chester's gaze looked natural to the participants (M = 4.93, SE = 0.16). They did not feel like Chester was staring at them (M = 2.91, SE = 0.17) nor avoiding looking at them (M = 2.01, SE = 0.12). These results had no significant main effects.

The robot's orientation led to significant differences on how much the participants felt that Chester looked at them ($F[1, 68] = 7.47$, $p < 0.01$). As shown in Fig. 7, the AO behavior (M = 4.72, SE = 0.18) had significantly higher ratings than the MO behavior (M = 4.11, SE = 0.15) in this respect. The fact that the results were not significantly different for Gaze may be explained by the Mona Lisa gaze effect and the tendency of the robot to look forward.

Robot's motion. Gaze had a significant effect on the ratings for "Chester's motion looked natural during the interaction" ($F[1, 68] = 4.08$, $p = 0.05$). As shown in Fig. 8, the AG behavior elicited significantly higher agreement with the statement relative to RG (M = 4.71, SE = 0.28 vs. M = 4.00, SE = 0.23). No significant differences were found for "Chester's motion was distracting" (M = 2.03, SE = 0.13), "I felt confident that the robot was not going to hit me" (M = 6.42, SE = 0.18), nor "Chester's motion made me anxious" (M = 1.57, SE = 0.12).

[2]Inter-coder reliability was computed for 4 sessions (20%). Two annotations were misaligned; Cohen's kappa for the other 75 annotations was 1.0, indicating perfect reliability.

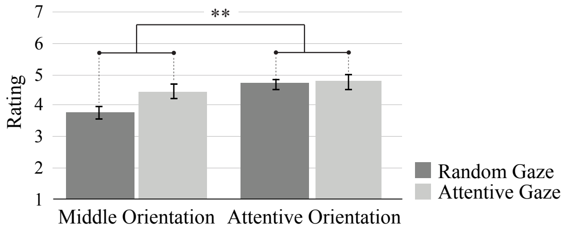

Figure 7: Ratings for how much the participants felt that Chester looked at them. (**) denotes $p < 0.01$.

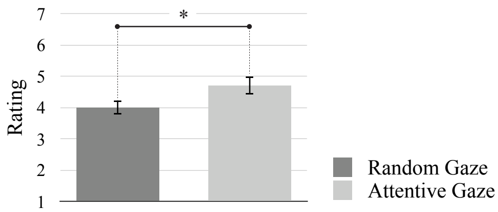

Figure 8: Ratings for how natural the robot's motion looked based on its gaze. (*) denotes $p < 0.05$.

Robot's attentiveness. Participants rated how much they thought that Chester paid attention to what they said (M = 5.33, SE = 0.16) and to what the other participants said (M = 5.48, SE = 0.14). A REML analysis with Participant as random effect nested by Session, and Orientation, Gaze, Gender, and Speaker (Me, Others) as main effects showed significant differences for the interaction between Orientation and Gender only ($F[3, 66] = 4.94$, p = 0.03). The post-hoc test then showed no significant pair-wise differences, but the tendency was interesting: male participants thought that the robot paid more attention with AO than with MO (M = 5.73, SE = 0.20 vs. M = 4.92, SE = 0.25).

Inclusion and ostracism: IOS scale [8] ratings indicated that the participants did not feel close to Chester (M = 2.57, SE = 0.14). However, they thought that both they (M = 5.03, SE = 0.17) and the robot (M = 5.33, SE = 0.19) belonged to the brainstorming group.

We found low perceptions of being ignored or excluded by the robot (M = 1.57, SE = 0.12) or the other participants (M = 1.42, SE = 0.09). REML analyses for these results resulted in no significant differences, but Orientation was close for the former (p = 0.06). The trend suggested that MO could lead to higher feelings of ostracism from the robot than AO (M = 1.77, SE = 0.17 vs. M = 1.33, SE = 0.14).

Other perceptions of the robot. The participants generally thought that Chester was a good leader for the brainstorming activity (M = 5.00, SE = 0.17) and had significantly different impressions of how much the robot and the other participants liked them ($F[1, 68] = 4.98$, p = 0.03). In particular, the participants thought that the robot liked them significantly more than did the other people in the experiment (M = 5.16, SE = 0.14 vs. M = 4.93, SE = 0.13).

Chester was not perceived as anti-social (M = 1.58, SE = 0.09). The only trend in this respect (p = 0.06) suggested that RG could make the robot look more anti-social than AG (M = 1.76, SE = 0.16 vs. M = 1.4, SE = 0.09).

Table 2 shows a factor analysis on a series of additional attributes for the robot. Factor I was associated with interactivity, Factor II with competence, and Factor III with

Table 2: Ratings for the factors resulting from factor analysis. Machine-like was reversed (R) for the analysis and for computing Chronbach's alpha.

Attribute	Mean (SE)	Cronbach's α	Factor
Responsive Interactive	5.33 (0.11)	0.786	I
Useful Knowledgeable Intelligent Competent	4.62 (0.13)	0.791	II
Entertaining Funny	5.47 (0.15)	0.846	III
Lifelike Machine-like (R)	4.14 (0.18) 4.20 (0.15)	0.623	-

entertainment. These factors explained 18.3%, 26.4%, and 20.3% of the variance, respectively. Their ratings were positive in general with no significant main effects of condition.

Only 8 participants of 69 indicated that Chester should be decommissioned in the post-survey. Their responses were typically associated with the robot's usefulness (e.g., *"I can't see a practical use for it, but the robot was entertaining"*).

Interaction: In general, the interaction with Chester was enjoyable (M = 5.45, SE = 0.14). Desire to brainstorm for longer was correlated with the number of paper slips written per session ($r(67) = 0.48$, p< 0.01), which tended to be just a few, or ten or more. This result motivated a REML analysis on the ratings for wanting to brainstorm for longer with Slip Count (1 if ten or more slips, 0 otherwise), Orientation, Gaze, and Gender as main effects, and Participant as random effect within Session. Not surprisingly, Slip Count had a significant effect ($F[1, 68] = 15.09$, p< 0.01). Ratings in sessions with many slips were significantly higher than the rest (M = 4.65, SE = 0.28 vs. M = 3.00, SE = 0.24). Also, the interaction between Gender and Slip Count was significant ($F[3, 66] = 6.78$, p= 0.01). Male participants wanted to brainstorm significantly more when there were at least ten slips (M = 5.27, SE = 0.37) than in other cases (M = 2.53, SE = 0.30). Female ratings were more uniform and neutral.

Note that the robot did not use a balancing criteria to ask people for ideas during the brainstorming activity. A REML analysis on the number of ideas proposed by the participants did not result in any significant differences for the main effects of Orientation, Gaze, and Gender, suggesting that this aspect of the protocol did not generate a confound. Moreover, all but one of the 69 participants proposed ideas. The only person that stayed quiet during the brainstorming phase took part in the activity towards the end, when the robot asked him to count the number of slips in the boxes.

5.2 Human Spatial Behavior

We analyzed proxemics during the brainstorming phase, when the participants often moved to write ideas at the table. For the analyses, we used the spatial information output by our perception system (sampled at 1Hz) and the group membership annotations described in Sec. 4.6.

When the participants conversed with the robot in the brainstorming phase, their average separation from Chester was typical of social encounters (M = 2.15m, SE = 0.04, N = 69) [22]. Because people often adjusted their position as they became familiar with the robot and the activity, we decided to further analyze proxemics during the last minute of the brainstorming part of the experiment. We performed a REML analysis on the distance between the robot and the participants during this period, considering Orientation, Gaze, and Gender as main effects and Participant as random effect nested by Session. Gaze was significant ($F[1, 68] = 5.67$, $p = 0.02$): participants stood significantly farther away from the robot with RG (M = 2.29, SE=0.05) than with AG (M = 2.09, SE = 0.06). The interaction between Gaze and Orientation was also significant ($F[3, 66] = 4.27$, $p = 0.04$). The members of the robot's group were significantly farther away from it with MO+RG than with MO+AG (M = 2.40, SE=0.07 vs. M = 2.03, SE=0.10), as shown in Fig. 9.

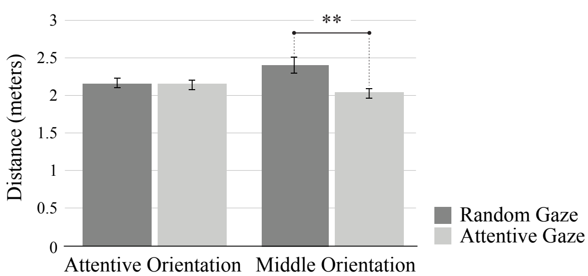

Figure 9: Distance to Chester during the last minute of brainstorming. (**) denotes $p < 0.01$.

Throughout the experiment, we observed qualitatively circular or side by side spatial arrangements. In a few cases in which the robot engaged in a dyadic interaction and it was not oriented as expected, participants proactively changed their positions to stand in front of it. These efforts to establish appropriate spatial arrangements suggest that people may be willing to collaborate with robots to establish F-formations and adapt to unforeseen circumstances.

5.3 Hypotheses Support and Implications

Attentive Gaze made the participants think that Chester's motion looked more natural in comparison to Random Gaze. This result supported our first hypothesis (H1) and is related to prior findings on the influence of gaze on the perception of a robot's head motion [35]. Furthermore, the orientation behaviors also affected the perception of the robot's gaze. With Attentive Orientation, the participants perceived that the robot looked at them more. These outcomes suggest that robot gaze and body motion should be designed and controlled jointly, rather than independently of each other.

We expected the Attentive Orientation behavior to make the robot seem more attentive and responsive than the Middle Orientation behavior (H2). While we did not find that the robot's orientation altered how responsive it looked, the participants thought that the robot gazed at them more with AO than with MO, as mentioned before. There was also a trend that suggested that male participants thought that the robot paid more attention to what people said with AO.

Opinions on how close the participants felt to the robot and whether they perceived it as part of their group were not significantly affected by the orientation behaviors, as hypothesized in H3. Interestingly, the distance between the participants and the robot varied significantly with MO based on the robot's gaze, but did not vary as much with AO. This finding might indicate more subtle effects of the manipulation than can be gleaned from questionnaires. Also, the lack of support opens up possibilities for developing more complex orientation behaviors and fulfilling other non-social tasks during multi-party interactions. Given that both MO and AO were acceptable and did not affect the perception that the robot was part of the group, both behaviors could be used by robots depending on other factors besides the interaction. For example, robots could switch between MO and AO to reduce uncertainty about the environment.

In terms of H4, the AO+AG condition did not reduce feelings of ostracism or increase feelings of inclusion relative to MO+RG. Nonetheless, there was a trend that suggested that MO could lead to higher feelings of ostracism than AO. This finding should be explored further in future research.

Finally, we learned an important lesson from the pilot sessions: people are sensitive to inappropriate or unexpected robot orientations. If users do not understand why a robot moves, interactions can easily be disrupted. This outcome is related to prior work on legible and predictable motion [17].

6. DISCUSSION

Limitations. Our work was limited in several ways. First, Chester's dialog was scripted and, thus, it could not respond appropriately in all circumstances. Second, the physical appearance and capabilities of our robot could have influenced our results and biased some aspects of the design of the behaviors under consideration. For example, the differential drive base of the robot constrained the complexity of its spatial behavior and, in turn, this could have affected the perception of its motion. Third, the perception system that we implemented for the experiment required instrumentation. While this system enabled autonomous robot behaviors, we are now interested in shifting towards on-board computation. This includes improving robots' capabilities so that they can reason about social contexts using on-board sensors only and, therefore, interact more casually.

Methodology. Overall, the perception of the brainstorming activity used in the experiment was positive. The protocol successfully created opportunities for changes in conversation group size, which allowed us to study the behaviors under consideration in different social contexts. In the future, this protocol could be used to study turn-taking patterns and collaboration in HRI. Similar to social games [65], brainstorming activities are customizable (e.g., the topic of the conversation can be easily adapted) and can be conducted with groups of strangers. In contrast, brainstorming sessions are less adversarial and do not require teaching very specific instructions.

Findings. The gaze of the robot affected the participants' perception of its motion and its motion affected the perception of its gaze. This dependency implies that robots should reason about and control their gaze and body motion jointly. Furthermore, some trends implied that the Attentive Orientation could be preferred over the Middle Orientation (e.g., AO could make the robot look more attentive and less antisocial). However, these behaviors led to similar feelings of inclusion and belonging to the group, suggesting that both AO and MO could be used as primitives for more complex orientation behaviors.

The implications of these findings are particularly important for mobile and low DoF robots, like ours, that engage in multi-party interactions, e.g., during social gatherings, or as they travel in human environments. In these circumstances, appropriate orientation and gaze behaviors can lead to more effective human-robot communication and user adoption.

7. ACKNOWLEDGMENTS

We thank Disney Research for continued support of this research effort, as well as Peter Carr and Alanson Sample for helping us set up the perception system used in this work.

References

[1] H. Admoni, B. Hayes, D. Feil-Seifer, D. Ullman, and B. Scassellati. Are You Looking at Me?: Perception of Robot Attention is Mediated by Gaze Type and Group Size. In *Proc. of the 8th ACM/IEEE Int'l Conf. on Human-robot Interaction*, HRI, 2013.

[2] H. Admoni and B. Scassellati. Social Eye Gaze in Human-Robot Interaction: A Review, 2017. Online access: hennyadmoni.com/documents/jhri17.pdf.

[3] S. Al Moubayed, J. Beskow, G. Skantze, and B. Granström. *Cognitive Behavioural Systems: COST 2102 International Training School, Revised Selected Papers*, chapter Furhat: A Back-Projected Human-Like Robot Head for Multiparty Human-Machine Interaction, pages 114–130. Springer Berlin Heidelberg, 2012.

[4] S. Al Moubayed, J. Edlund, and J. Beskow. Taming Mona Lisa: Communicating Gaze Faithfully in 2D and 3D Facial Projections. *ACM Trans. on Interactive Intelligent Systems*, 1(2):25, 2012.

[5] P. Althaus, H. Ishiguro, T. Kanda, T. Miyashita, and H. Christensen. Navigation for human-robot interaction tasks. In *Proc. of the 2004 IEEE Int'l Conf. on Robotics and Automation*, ICRA, 2004.

[6] S. Andrist, B. Mutlu, and A. Tapus. Look Like Me: Matching Robot Personality via Gaze to Increase Motivation. In *Proc. of the 33rd Annual ACM Conf. on Human Factors in Computing Systems*, CHI, 2015.

[7] M. Argyle. *Bodily Communication*. University paperbacks. Methuen, 1988.

[8] A. Aron, E. N. Aron, and D. Smollan. Inclusion of Other in the Self Scale and the structure of interpersonal closeness. *Journal of personality and social psychology*, 63(4):596, 1992.

[9] B. Bodt, R. Camden, H. Scott, A. Jacoff, T. Hong, T. Chang, R. Norcross, T. Downs, and A. Virts. Performance Measurements for Evaluating Static and Dynamic Multiple Human Detection and Tracking Systems in Unstructured Environments. In *Proc. of the 9th Workshop on Performance Metrics for Intelligent Systems*, PerMIS, 2009.

[10] H.-J. Böhme, T. Wilhelm, J. Key, C. Schauer, C. Schröter, H.-M. Groß, and T. Hempel. An approach to multi-modal human-machine interaction for intelligent service robots. *Robotics and Autonomous Systems*, 44(1):83 – 96, 2003.

[11] D. Bohus and E. Horvitz. Dialog in the Open World: Platform and Applications. In *Proc. of the 2009 Int'l Conf. on Multimodal Interfaces*, ICMI-MLMI, 2009.

[12] J. Cassell, O. E. Torres, and S. Prevost. Turn taking versus discourse structure. In *Machine conversations*, pages 143–153. Springer, 1999.

[13] I. Chatterjee. A low-cost, human-inspired perception approach for dense moving crowd navigation. Master's thesis, Robotics Institute, Carnegie Mellon University, August 2016.

[14] J. J. Choi, Y. Kim, and S. S. Kwak. Have You Ever Lied?: The Impacts of Gaze Avoidance on People's Perception of a Robot. In *Proc. of the 8th ACM/IEEE Int'l Conf. on Human-robot Interaction*, HRI, 2013.

[15] M. Cristani, L. Bazzani, G. Paggetti, A. Fossati, D. Tosato, A. Del Bue, G. Menegaz, and V. Murino. Social Interaction Discovery by Statistical Analysis of F-Formations. In *Proc. of the British Machine Vision Conf.*, BMVC, 2011.

[16] M. J. Doughty. Consideration of three types of spontaneous eyeblink activity in normal humans: during reading and video display terminal use, in primary gaze, and while in conversation. *Optometry & Vision Science*, 78(10):712–725, 2001.

[17] A. Dragan, S. Bauman, J. Forlizzi, and S. Srinivasa. Effects of Robot Motion on Human-Robot Collaboration. In *Proc. of the 10th ACM/IEEE Int'l Conf. on Human-Robot Interaction*, HRI, 2015.

[18] V. Evers, N. Menezes, L. Merino, D. Gavrila, F. Nabais, M. Pantic, P. Alvito, and D. Karreman. The Development and Real-world Deployment of FROG, the Fun Robotic Outdoor Guide. In *Proc. of the 9th ACM/IEEE Int'l Conf. on Human-robot Interaction*, HRI, 2014.

[19] N. Fisher. *Statistical Analysis of Circular Data*. Cambridge University Press, 1995.

[20] M. Garau, M. Slater, S. Bee, and M. A. Sasse. The Impact of Eye Gaze on Communication Using Humanoid Avatars. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems*, CHI, 2001.

[21] A. Godil, R. Bostelman, K. Saidi, W. Shackleford, G. Cheok, M. Shneier, and T. Hong. 3D Ground-Truth Systems for Object/Human Recognition and Tracking. In *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2013.

[22] E. T. Hall. *The Hidden Dimension*. Doubleday, 1966.

[23] Z. Henkel, C. Bethel, R. Murphy, and V. Srinivasan. Evaluation of Proxemic Scaling Functions for Social Robotics. *IEEE Transactions on Human-Machine Systems*, 44(3):374–385, 2014.

[24] C.-M. Huang, M. Cakmak, and B. Mutlu. Adaptive Coordination Strategies for Human-Robot Handovers. In *Proc. of Robotics: Science and Systems*, R:SS, 2015.

[25] D. Karreman, G. Bradford, B. Van Dijk, M. Lohse, and V. Evers. What happens when a robot favors someone? How a tour guide robot uses gaze behavior to address multiple persons while storytelling about art. In *Proc. of the 8th ACM/IEEE Int'l Conf. on Human-Robot Interaction*, HRI, 2013.

[26] D. E. Karreman, G. D. Ludden, E. M. van Dijk, and V. Evers. How can a tour guide robot's orientation influence visitors' orientation and formations? In *Proc. of the 4th Int'l Symposium on New Frontiers in Human-Robot Interaction*, 2015.

[27] Y. Kato, T. Kanda, and H. Ishiguro. May I Help You?: Design of Human-like Polite Approaching Behavior. In *Proc. of the 10th ACM/IEEE Int'l Conf. on Human-Robot Interaction*, HRI, 2015.

[28] M. Katzenmaier, R. Stiefelhagen, and T. Schultz. Identifying the Addressee in Human-human-robot Interactions Based on Head Pose and Speech. In *Proc. of the 6th Int'l Conf. on Multimodal Interfaces*, ICMI, 2004.

[29] A. Kendon. *Conducting Interaction: Patterns of Behavior in Focused Encounters*. Cambridge UP, 1990.

[30] N. Kirchner, A. Alempijevic, and G. Dissanayake. Nonverbal Robot-group Interaction Using an Imitated Gaze Cue. In *Proc. of the 6th Int'l Conf. on Human-robot Interaction*, HRI, 2011.

[31] H. Kuzuoka, Y. Suzuki, J. Yamashita, and K. Yamazaki. Reconfiguring Spatial Formation Arrangement by Robot Body Orientation. In *Proc. of the 5th ACM/IEEE Int'l Conf. on Human-robot Interaction*, HRI, 2010.

[32] S. Lang, M. Kleinehagenbrock, S. Hohenner, J. Fritsch, G. A. Fink, and G. Sagerer. Providing the Basis for Human-robot-interaction: A Multi-modal Attention System for a Mobile Robot. In *Proc. of the 5th Int'l Conf. on Multimodal Interfaces*, ICMI, 2003.

[33] I. Leite, M. McCoy, D. Ullman, N. Salomons, and B. Scassellati. Comparing Models of Disengagement in Individual and Group Interactions. In *Proc. of the 10th Annual ACM/IEEE Int'l Conf. on Human-Robot Interaction*, HRI, 2015.

[34] T. Linder and K. Arras. Multi-model hypothesis tracking of groups of people in RGB-D data. In *Proc. of the 17th Int'l Conf. on Information Fusion*, FUSION, 2014.

[35] C. Liu, C. T. Ishi, H. Ishiguro, and N. Hagita. Generation of Nodding, Head Tilting and Eye Gazing for Human-robot Dialogue Interaction. In *Proc. of the 7th ACM/IEEE Int'l Conf. on Human-Robot Interaction*, HRI, 2012.

[36] M. Luber and K. O. Arras. Multi-Hypothesis Social Grouping and Tracking for Mobile Robots. In *Proc. of Robotics: Science and Systems*, R:SS, 2013.

[37] Y. Matsuyama, I. Akiba, S. Fujie, and T. Kobayashi. Four-participant group conversation: A facilitation robot controlling engagement density as the fourth participant. *Computer Speech & Language*, 33(1):1 – 24, 2015.

[38] R. Mead, A. Atrash, and M. J. Mataric. Recognition of Spatial Dynamics for Predicting Social Interaction. In *Proc. of the 6th Int'l Conf. on Human-Robot Interaction Extended Abstracts*, HRI EA, 2011.

[39] M. P. Michalowski, S. Sabanovic, and R. Simmons. A Spatial Model of Engagement for a Social Robot. In *Proc. of the 9th Int'l Workshop on Advanced Motion Control*, AMC, 2006.

[40] S. Muller, S. Hellbach, E. Schaffernicht, A. Ober, A. Scheidig, and H.-M. Gross. Whom to talk to? Estimating user interest from movement trajectories. In *Proc. of the 17th IEEE Int'l Symposium on Robot and Human Interactive Communication*, RO-MAN, 2008.

[41] J. Mumm and B. Mutlu. Human-robot Proxemics: Physical and Psychological Distancing in Human-robot Interaction. In *Proc. of the 6th Int'l Conf. on Human-robot Interaction*, HRI, 2011.

[42] B. Mutlu, J. Forlizzi, and J. Hodgins. A Storytelling Robot: Modeling and Evaluation of Human-like Gaze Behavior. In *Proc. of the 2006 6th IEEE-RAS Int'l Conf. on Humanoid Robots*, Humanoids, 2006.

[43] B. Mutlu, T. Shiwa, T. Kanda, H. Ishiguro, and N. Hagita. Footing in human-robot conversations: How robots might shape participant roles using gaze cues. In *Proc. of the 4th ACM/IEEE Int'l Conf. on Human-Robot Interaction*, HRI, 2009.

[44] M. A. Nicolaou, H. Gunes, and M. Pantic. Continuous Prediction of Spontaneous Affect from Multiple Cues and Modalities in Valence-Arousal Space. *IEEE Trans. on Affective Computing*, 2(2):92–105, 2011.

[45] H. D. Patterson and R. Thompson. Maximum likelihood estimation of components of variance. In *Proc. of the 8th Int'l Conf. on Biochem.*, 1974.

[46] I. Rae. Using Robot-mediated Communication to Improve Remote Collaboration. In *Proc. of CHI'13 Extended Abstracts on Human Factors in Computing Systems*, CHI EA, 2013.

[47] O. A. I. Ramírez, G. Varni, M. Andries, M. Chetouani, and R. Chatila. Modeling the dynamics of individual behaviors for group detection in crowds using low-level features. In *Proc. of the 2016 IEEE Int'l Symposium on Robot and Human Interactive Communication*, RO-MAN, 2016.

[48] J. Rios-Martinez, A. Spalanzani, and C. Laugier. Understanding human interaction for probabilistic autonomous navigation using Risk-RRT approach. In *Proc. of the 2011 IEEE/RSJ Int'l Conf. on Intelligent Robots and Systems*, IROS, 2011.

[49] V. Rousseau, F. Ferland, D. Létourneau, and F. Michaud. Sorry to interrupt, but may I have your attention? Preliminary design and evaluation of autonomous engagement in HRI. *Journal of Human-Robot Interaction*, 2(3):41–61, 2013.

[50] K. Ruhland, S. Andrist, J. Badler, C. Peters, N. Badler, M. Gleicher, B. Mutlu, and R. Mcdonnell. Look me in the eyes: A survey of eye and gaze animation for virtual agents and artificial systems. In *Eurographics 2014 - State of the Art Reports*, 2014.

[51] S. Satake, T. Kanda, D. Glas, M. Imai, H. Ishiguro, and N. Hagita. How to approach humans? - strategies for social robots to initiate interaction. In *Proc. of the 4th ACM/IEEE Int'l Conf. on Human-Robot Interaction*, HRI, 2009.

[52] F. Setti, C. Russell, C. Bassetti, and M. Cristani. F-formation Detection: Individuating Free-standing Conversational Groups in Images. *CoRR*, abs/1409.2702, 2014.

[53] C. Shi, M. Shimada, T. Kanda, H. Ishiguro, and N. Hagita. Spatial Formation Model for Initiating Conversation. In *Proc. of Robotics: Science and Systems*, R:SS, 2011.

[54] M. Shiomi, T. Kanda, H. Ishiguro, and N. Hagita. A Larger Audience, Please!: Encouraging People to Listen to a Guide Robot. In *Proc. of the 5th ACM/IEEE Int'l Conf. on Human-Robot Interaction*, HRI, 2010.

[55] E. Short, J. Hart, M. Vu, and B. Scassellati. No fair!! An interaction with a cheating robot. In *Proc. of the 5th ACM/IEEE Int'l Conf. on Human-Robot Interaction*, HRI, 2010.

[56] C. Sidner, C. Lee, C. Kidds, N. Lesh, and C. Rich. Explorations in Engagement for Humans and Robots. *Artificial Intelligence*, 2005.

[57] C. L. Sidner, C. Lee, L.-P. Morency, and C. Forlines. The Effect of Head-nod Recognition in Human-Robot Conversation. In *Proc. of the 1st ACM Conf. on Human-Robot Interaction*, HRI, 2006.

[58] G. Skantze, A. Hjalmarsson, and C. Oertel. Exploring the effects of gaze and pauses in situated human-robot interaction. In *Proc. of the 14th Annual Meeting of Special Interest Group on Doscourse and Dialogue*, SIGDial, 2013.

[59] A. Steinfeld, O. C. Jenkins, and B. Scassellati. The Oz of Wizard: Simulating the human for interaction research. In *Proc. of the 4th ACM/IEEE Int'l Conf. on Human-Robot Interaction*, HRI, 2009.

[60] W. W. Stroup. *Generalized Linear Mixed Models: Modern Concepts, Methods and Applications.* CRC Press, 2012.

[61] L. Takayama and C. Pantofaru. Influences on Proxemic Behaviors in Human-Robot Interaction. In *Proc. of the 2009 IEEE/RSJ Int'l Conf. on Intelligent Robots and Systems*, IROS, 2009.

[62] S. Thrun, W. Burgard, and D. Fox. *Probabilistic Robotics.* Intelligent Robotics and Autonomous Agents. MIT Press, 2005.

[63] J. G. Trafton, M. D. Bugajska, B. R. Fransen, and R. M. Ratwani. Integrating vision and audition within a cognitive architecture to track conversations. In *Proc. of the 2008 3rd ACM/IEEE Int'l Conf. on Human-Robot Interaction*, HRI, 2008.

[64] R. Triebel, K. Arras, R. Alami, L. Beyer, S. Breuers, R. Chatila, M. Chetouani, D. Cremers, V. Evers, M. Fiore, H. Hung, O. A. I. Ramírez, M. Joosse, H. Khambhaita, T. Kucner, B. Leibe, A. J. Lilienthal, T. Linder, M. Lohse, M. Magnusson, B. Okal, L. Palmieri, U. Rafi, M. van Rooij, and L. Zhang. SPENCER: A Socially Aware Service Robot for Passenger Guidance and Help in Busy Airports. In *Proc. of the 10th Conf. on Field and Service Robotics*, FSR, 2015.

[65] M. Vázquez, E. J. Carter, J. A. Vaz, J. Forlizzi, A. Steinfeld, and S. E. Hudson. Social Group Interactions in a Role-Playing Game. In *Proc. of the 10th ACM/IEEE Int'l Conf. on Human-Robot Interaction Extended Abstracts*, HRI EA, 2015.

[66] M. Vázquez, A. Steinfeld, and S. E. Hudson. Parallel Detection of Conversational Groups of Free-Standing People and Tracking of their Lower-Body Orientation. In *Proc. of the 2015 IEEE/RSJ Int'l Conf. on Intelligent Robots and Systems*, IROS, 2015.

[67] M. Vázquez, A. Steinfeld, S. E. Hudson, and J. Forlizzi. Spatial and Other Social Engagement Cues in a Child-robot Interaction: Effects of a Sidekick. In *Proceedings of the 2014 ACM/IEEE Int'l Conf. on Human-robot Interaction*, HRI, 2014.

[68] M. M. Veloso, J. Biswas, B. Coltin, and S. Rosenthal. CoBots: Robust Symbiotic Autonomous Mobile Service Robots. In *Proc. of the 24th Int'l Joint Conf. on Artificial Intelligence*, IJCAI, 2015.

[69] 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. In *Proc. of the 2015 24th IEEE Int'l Symposium on Robot and Human Interactive Communication*, RO-MAN, 2015.

[70] M. Walters, K. Dautenhahn, K. Koay, C. Kaouri, R. Boekhorst, C. Nehaniv, I. Werry, and D. Lee. Close encounters: spatial distances between people and a robot of mechanistic appearance. In *Proc. of the 5th IEEE-RAS Int'l Conf. on Humanoid Robots*, 2005.

[71] M. Walters, D. Syrdal, K. Koay, K. Dautenhahn, and R. te Boekhorst. Human approach distances to a mechanical-looking robot with different robot voice styles. In *Proc. of the 17th IEEE Int'l Symposium on Robot and Human Interactive Communication*, RO-MAN, 2008.

[72] 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 *Proc. of the IEEE International Workshop on Robot and Human Interactive Communication*, RO-MAN, 2005.

[73] 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.

[74] K. D. Williams, C. K. Cheung, and W. Choi. Cyberostracism: effects of being ignored over the Internet. *J. of Personality and Social Psychology*, 79(5):748, 2000.

[75] Y. Yoshikawa, K. Shinozawa, H. Ishiguro, N. Hagita, and T. Miyamoto. Responsive robot gaze to interaction partner. In *Proc. of Robotics: Science and Systems*, R:SS, 2006.

[76] M. A. Yousuf, Y. Kobayashi, Y. Kuno, A. Yamazaki, and K. Yamazaki. Development of a Mobile Museum Guide Robot That Can Configure Spatial Formation with Visitors. In *Intelligent Computing Technology*, volume 7389 of *Lecture Notes in Comput. Sci.* 2012.

[77] Z. Yumak, J. Ren, N. M. Thalmann, and J. Yuan. Tracking and Fusion for Multiparty Interaction with a Virtual Character and a Social Robot. In *SIGGRAPH Asia 2014*, SA, pages 3:1–3:7, 2014.