Perceptions of Conversational Group Membership based on Robots' Spatial Positioning: Effects of Embodiment

Joe Connolly
Yale University
joe.connolly@yale.edu

Nathan Tsoi
Yale University
nathan.tsoi@yale.edu

Marynel Vázquez
Yale University
marynel.vazquez@yale.edu

ABSTRACT

Robots' spatial positioning is a useful communication modality in social interactions. For example, in the context of group conversations, certain types of positioning signal membership to the group interaction. How does robot embodiment influence these perceptions? To investigate this question, we conducted an online study in which participants observed renderings of several robots in a social environment, and judged whether the robots were positioned to take part in a group conversation with other humans in the scene. Our results suggest that robot embodiment can influence perceptions of conversational group membership. An important factor to consider in this regard is whether robot embodiment leads to a discernible orientation for the agent.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in HCI**; • **Computing methodologies → Spatial and physical reasoning**.

KEYWORDS

Human-robot interaction; embodiment; proxemic interactions

ACM Reference Format:
Joe Connolly, Nathan Tsoi, and Marynel Vázquez. 2021. Perceptions of Conversational Group Membership based on Robots' Spatial Positioning: Effects of Embodiment. In *Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction (HRI '21 Companion), March 8–11, 2021, Boulder, CO, USA.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3434074.3447195

1 INTRODUCTION

Group interactions are an important area of investigation in Human-Robot Interaction (HRI) [15]. One way of determining whether robots are considered a social member of a particular group is their spatial positioning. In the case of conversational groups, for example, prior work in social psychology has shown that conversations lead to structured spatial patterns of behavior which are sustained during these interactions [8], e.g., face-to-face or circular spatial arrangements. Various studies have shown that these spatial patterns translate to the context of human-robot interactions [1, 7, 9, 24].

HRI '21 Companion, March 8–11, 2021, Boulder, CO, USA

ACM ISBN 978-1-4503-8290-8/21/03…$15.00
https://doi.org/10.1145/3434074.3447195

Recently, data-driven methods to recognize typical spatial patterns of behavior in human-robot conversations have been built upon datasets of human interactions [6, 19]. Likewise, methods for robots to conform to these conversational patterns have been proposed in HRI by modeling human spatial behavior [12, 17, 25, 28]. However, it is unclear how robots' embodiment may affect human perceptions of spatial behavior during social group conversations. In the context of proxemics [5], prior work in HRI suggests that factors such as a robot's humanoid appearance [26] and robots' gaze (along with their likeability) can influence human-robot distancing [14]. This makes us believe that embodiment can alter human perception of robots' spatial positioning in social contexts and, thus, influence perceptions of conversational group membership in HRI.

We conducted an online study to investigate how robot embodiment influences human perception of conversational group membership. We considered various types of robots in our study, from more anthropomorphic platforms like Pepper and Kuri to less human-like robots like the Turtlebot 3 or Jackal (Fig. 1). As a baseline, we compared perceptions of spatial positioning by robots with spatial positioning by virtual humans. Our results suggest that robot embodiment can influence human conception of personal space and the likelihood that people will consider various agents as part of conversational groups based on their spatial positioning.

2 RELATED WORK

Proxemics. Significant work has investigated how people socially perceive and use the physical space around them. Hall built a general framework for the primary spatial zones people find themselves in based on their interactions [5]. This framework includes zones of intimate space, personal space, social space, and public space. Research suggests that the size of these zones can vary across cultural contexts and based on human personal preferences [5, 18]. Furthermore, researchers have investigated how proxemic zones differ for robots as opposed to other humans [13, 14]. Walters et. al made adjustments to previously known proxemic distances based on factors such as robot appearance, preferences, interaction context, and situation [26]. Overall, they found that people generally preferred more humanoid appearance robots to keep a further distance away than mechanoid robots. However, the height of the robots considered in their study did not impact these preferences. Worth noting, users' gender may influence perceptions of proxemics in HRI [20].

Conversational Formations. Face Formations (or F-formations) are spatial patterns of human behavior that organically arise during group conversations [8]. They are a consequence of people needing to be close to each other to talk in conversation or engage in an interaction that requires a common, focused point of attention. Classic F-formations include face-to-face or shoulder-to-shoulder spatial arrangements in smaller groups. Circular arrangements are

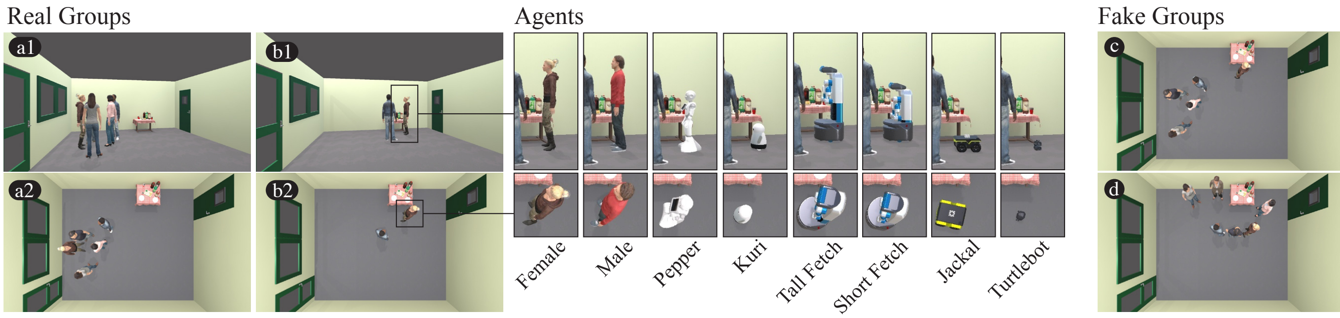

Figure 1: Images a, b, c and d show two real conversational groups from the Cocktail Party dataset and two fake groups, respectively. Image a2 shows a top-down view of a1, b2 shows a top-down view of b1, and c was generated from a's data. The middle area shows sections of b1 and b2 for all the agents considered in the study. The room is approx. 4.8 × 6.0 meters.

more typical for larger groups. Other factors like environmental constraints may also alter how people congregate socially in groups [2, 11, 21]. In our work, we leverage data of naturalistic human conversational groups to render visual stimuli for our study.

Robot Embodiment. Robots are known to be perceived differently based upon their physical presence and appearance [3]. For instance, Li et al. discovered that people were more comfortable with being close to real-life robots as opposed to robots rendered in virtual reality, and rated the appearance of real-life robots higher [10]. Moreover, people tend to like robots that have humanoid appearances more than those with basic mechanical appearances, both in static images and when robots performed simple dynamic actions [27]. Our study expands prior work on robot embodiment by investigating how it may affect human perception of robots' positioning in relation to conversational groups.

3 METHOD

We investigated how people perceive several robots and their membership to group conversations based on their poses in a social environment. To this end, we modeled in 3D the environment of the Cocktail Party dataset [29], a popular dataset for conversational group detection [16, 19, 23]. Then, we rendered 20 conversational groups from the dataset in the environment using the Unity game engine. This corresponded to 5 sets of 4 groups with 2, 3, 4, 5, and 6 people each. For the renderings, we varied the 3D model used to represent one of the agents. This agent could be a female character, a male character or 6 different types of robots (Fig. 1). Through these variations we aimed to study the effect of different embodiments on perceptions of spatial behavior and social interactions.

To gather human opinions in more diverse settings, we purposefully generated additional renderings of "fake" groups from the above data. For conversational groups with less than six members, we replaced the pose of the chosen agent with the pose of another agent in the dataset who was not part of the original group that was rendered in the scene. For groups with six members, which included all social interactants in the dataset, we rotated the orientation of the chosen agent such that it was opposite to the center of the group. This resulted in an atypical spatial pattern for social conversations [8]. Figure 1 (right) shows two example fake groups where the female character is oddly positioned in the scene.

Study Design and Hypotheses. We designed the study with a $20 \times 8 \times 2$ mixed design, considering *Groups* (4 x 5 group sizes), *Agent* (2 human agents and 6 robots), and rendering *Type* (real or fake group) as independent variables. All the participants evaluated all the groups, but only one agent and one type of rendering. We hypothesized that:

H1. People would more easily identify the orientation of the agents with a face (Female, Male, Pepper, Kuri, Fetch) than the orientation of the other agents (Jackal, Turtlebot) in the renderings.

H2. Perceptions of the agents being part of a conversational group would be higher with the real groups than with the fake groups. This assumption follows from the design of our visual stimuli.

H3. As a corollary of H1, the agents with a face (Female, Male, Pepper, Kuri, Fetch) would be more often identified as being part of real groups based on their pose in the scene than the other agents (Jackal, Turtlebot). Also, the latter agents would be more often identified as being part of fake groups than the former agents.

H4. The robots with the widest bases (Pepper, Fetch and Jackal) would be more often perceived as standing too close to virtual humans to socially engage with them in comparison to the other agents. This hypothesis is complementary to H3, as it focuses more on proxemics [5] than group formations. The assumption was motivated by some of the robots considered in our study being wider than the humans that originally generated the Cocktail Party data.

Procedure. As approved by our Institutional Review Board, we ran our study as an online survey. The survey first gathered demographics data. Then, it showed renderings of the 20 (real or fake) groups chosen for our study, with one of the 8 agents displayed in them. For each group, the survey asked the participants to visually identify the agent in the rendered scene. Next, it asked to rate a number of statements about the pose of this agent relative to the other humans. At the end of the survey, the participants provided their opinion about how hard it was to complete the survey based on the appearance of the specific agent that they experienced. The participants were paid $2 USD for completing the survey.

Image Renderings. We leveraged tools from the Social Environment for Autonomous Navigation [22] to generate our renderings in Unity. More specifically, we created two separate cameras in the

Unity scene to capture the Cocktail Party environment from overhead and side angles, such that participants could easily perceive the agents' spatial positioning relative to one another. We then used a ROS script to load the pose of the agents in the 20 groups, pass these poses to Unity via ROS#, receive the rendered images in return, and save them to disk for use in our survey. We utilized the Microsoft Rocketbox avatar library to render the human agents [4]. For the robots, we used 3D models from open-source Universal Robot Description Files (URDFs).

Measures. We gathered three types of measures via the survey:
(1) For H1, we asked the participants to indicate for each rendered group with a given agent X if "*X is oriented towards other human(s) in the scene.*" They could choose among 3 answers: "Yes", "No", and "I cannot tell the orientation of X from the scene views." Also, we asked the participants at the end of the survey to indicate if "*the survey was difficult to complete because of the appearance of X*" using a 7 point responding format. We gave participants the option to further explain their answer to the latter question via a text box.

(2) For H2 and H3, we gathered participant opinions in regards to whether they agreed with the following statements about the given agent X: "*X is too far from the human(s) in the scene to engage naturally in a group conversation with them*"; "*X is in a location that makes it look like (s)he is in a group conversation with everybody else in the scene*"; "*X is positioned to socially engage with the human(s) in the scene*"; and "*X is orienting in an unusual way to be having a conversation with everybody else in the scene.*" Ratings were obtained using a 7 point Likert responding format from "strongly disagree" to "strongly agree." We combined these ratings into an "In Group" measure based on the position and orientation of the agents (Cronbach's alpha was 0.87).

(3) For H4, we asked the participants to indicate their agreement with "*X is uncomfortably too close to a human to socially engage with him/her in the scene*" using a 7 point Likert responding format.

Participants. We recruited 480 participants via Prolific, with 240 female participants, 238 male participants, and two participants who indicated "Other" gender. We excluded the last two participants from our analyses (Sec. 4) because their gender did not fit the two prescreening categories of male/female. We randomly assigned males, and then females, to each combination of Agent and Type such that there was roughly the same amount of participants per condition. Overall, each Agent/Type combination had about 30 participants split roughly evenly between males and females.

4 RESULTS

We conducted analyses on our measures considering Agent (8 levels), Type (2 levels), and participant Gender (2 levels) as main effects. For image ratings, we also considered Groups (20 levels) as main effect, and Participant ID as random effect. We used Student's t-tests or Tukey HSD tests for post-hoc comparisons when appropriate.

Agent's Orientation. The ratings for "*the survey was difficult to complete because of the (agent's) appearance*" were heavily skewed towards low ratings, so we analyzed them using non-parametric Kruskal-Wallis tests. The tests indicated that only Agent had a significant effect on the ratings, $X^2(7) = 47.81$, $p < 0.0001$. A Tukey HSD post-hoc test indicated that the ratings for the Turtlebot (M= 3.13, SE= 0.267) were significantly higher than for all other agents except for the Jackal (M= 2.66, SE = 0.244). Also, Jackal had significantly higher ratings than the Female agent (M= 1.70, SE= 0.163), Kuri (M= 1.67, SE= 0.146), the Tall Fetch (M= 1.63, SE= 0.118), and the Male agent (M= 1.5, SE= 0.138). These significant differences were further confirmed with a non-parametric Steel-Dwass test. These results provide partial support for H1.

Among those who found the survey difficult, the responses fell into two broad categories: some of the agents had orientations that were difficult to discern based on their appearance, and participants had concerns about the agents' appearance and capabilities to be socially engaged with humans. The former category was more apparent for Turtlebot and Jackal, which respectively had 16 and 9 responses out of 34 commentaries (Short Fetch had 5, Kuri had 2, Pepper had 1, and Female had 1). For example, a participant said about Jackal that "*It was difficult to see where the front and the back is*," and another said about Turtlebot that "*It was hard to tell where it was turned to.*" In regard to concerns about social engagement, Jackal garnered most of the responses in this category with 8 out of 23 comments (Short Fetch had 5, Tall Fetch had 4, Kuri had 3, and Turtlebot had 2, and Pepper had 1). People said things about Jackal such as "*When picturing a bunch of humans engaged in social interaction, throwing in a shin-height robot with no humanlike characteristics made it hard to imagine the scenarios,*" "*It looks like a roomba-type thing so its not very believable that it can even socially interact with people in any capacity,*" and "*It's boxlike appearance and lack of anthropomorphic features made it hard to imagine conversing with.*" Taken together, these results align with H1, as Jackal and Turtlebot were the least human-like agents and seemed to cause the most confusion. However, we acknowledge that they might have been influenced by our specific choice of visual stimuli. Images are static and may not fully convey how the agents could take part in social conversations.

Answers to whether the agent of interest was oriented towards the other humans in each scene provided further support for H1. Because these ratings were a repeated measure, we analyzed them using a Binomial Generalized Mixed Linear Model with a logit link function. To this end, we computed whether the given agent was orienting towards the other humans in the scene using the agents' poses: this was assumed to be true when the agent was oriented within $\pm 60^\circ$ towards the average position of the other humans. Then, we counted Yes/No answers that matched our definition as correct, and labelled both incompatible and unsure responses as incorrect. Interestingly, Agent had significant effects on identifying the orientation correctly, $F[7, 482.5] = 10.71$ ($p < 0.0001$). As shown in Fig. 2 (left), Jackal and Turtlebot led to significantly fewer correct answers than all other agents according to a Tukey HSD post-hoc test. Also, Groups ($F[19, 9057] = 45.30$, $p < 0.0001$), the interaction between Gender and Type ($F[1, 432.8] = 5.35$, $p = 0.02$), and the interaction of Agent and Type ($F[7, 433.3] = 6.27$, $p < 0.0001$) were significant. The post-hoc tests showed pairwise differences across groups, which we attribute to our specific choice for the study. In addition, there were significant differences in the responses by Type for male participants and for the Turtlebot. We omit further details due to limited space.

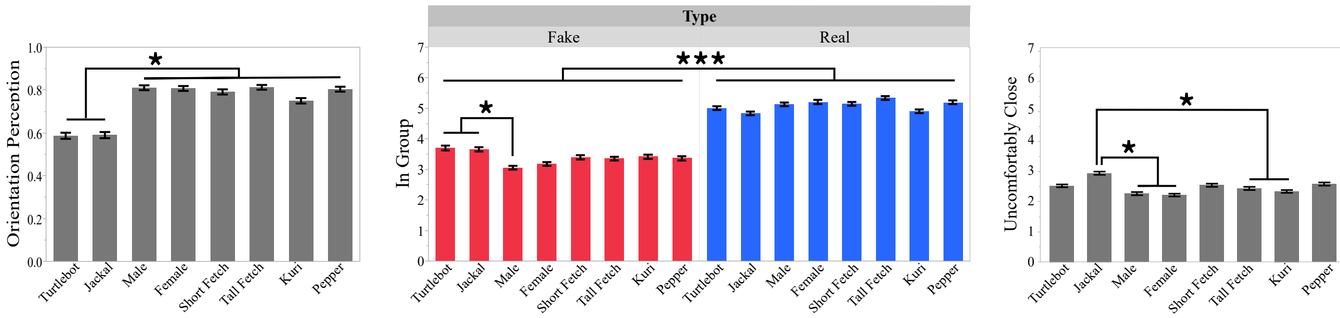

Figure 2: The proportion of images in which the participants identified correctly whether the agents oriented toward the other humans in the rendering (left), participants' combined ratings for the agent being in a group with the humans (middle), and participant ratings for agents being "*uncomfortably too close to a human to socially engage with him/her in the scene*" (right).

Perceptions of Group Membership. We conducted a Restricted Maximum Likelihood (REML) analysis on the In Group measure to evaluate H2 and H3. We found that Type (F[1, 453]=824.97, p<0.0001) and Groups (F[19, 9063]=222.65, p<0.0001) had significant effects on the results. A t-test on Type supported H2: the real groups led to significantly higher In Group perceptions (M=5.10, SE=0.022) than the fake groups (M=3.39, SE=0.024). The post-hoc test on Groups also revealed significant differences, which we attribute to our specific choice of conversational groups as before.

We also found that the interaction between Agent and Type led to significant differences for the In Group results, F[7, 453] = 4.06 (p<0.0002). A Tukey HSD test indicated that the ratings for all the agents in the real groups were significantly higher than for them in the fake groups. For the fake groups, the ratings for the Turtlebot (M= 3.70, SE= 0.073) and Jackal (M= 3.66, SE= 0.070) were significantly higher than for the Male agent (M= 3.06, SE= 0.062), as shown in Fig. 2 (middle). The results partially support H3 for the fake groups, but provide no evidence for the real groups.

Social Distancing. A REML analysis on perceptions of the agents being "*uncomfortably too close to a human to socially engage*" provided partial support for H4. Agent (F[7, 453]=3.93, p=0.0004), Type (F[1, 453]=85.05, p<0.0001), Group (F[1, 453]=85.05, p<0.0001), and Gender (F[1. 453]=7.51, p<0.0064) all had significant effects on these perceptions. A Tukey HSD test on Agent showed that the ratings were significantly higher for Jackal (M= 2.89, SE= 0.053) than the Tall Fetch (M= 2.40, SE= 0.054), Kuri (M= 2.30, SE= 0.050), the Male agent (M= 2.23, SE= 0.051), and the Female agent (M= 2.18, SE= 0.048). These results are shown in Fig. 2 (right). We were surprised by the small Turtlebot having ratings comparable to Pepper. Perhaps this could be explained by the fact that some people had a hard time imagining interacting with the Turtlebot. For example, one participant indicated that "*The size of it [Turtlebot] would be an issue. Having to look towards the ground would be problematic.*"

In regard to Type, ratings for the Real groups (M=2.81, SE=0.022) were significantly higher than for Fake groups (M=2.07, SE=0.028). This difference can be explained by the Fake groups often having agents away from a group. Lastly, in terms of participant Gender, the post-hoc test suggested that females (M=2.34, SE=0.026) gave significantly lower ratings than males (M=2.55, SE=0.026).

5 LIMITATIONS & FUTURE WORK

Our work focused on evaluating the perception of specific agents in group formations originally established by humans. Expanding the set of agents and using human-robot interaction groups instead are interesting directions for future work. Also, the visual stimuli that we used for our study was static. The images of social environments omitted details about the motion of the agents. In the future, we would like to extend this type of evaluations to interactive HRI simulations, like [22], in which participants can observe the motion of the robots as well. We would also like to expand our collected data to broader categories of gender, as we had to omit the two non-binary responses we received from our analysis due to the extremely small sample size. Importantly, further experiments are needed to validate our results in in-person human-robot interactions.

6 CONCLUSION

We explored the intersection between spatial positioning and embodiment in perceptions of human-robot conversational groups with two or more interactants. Our findings provide concrete evidence that robot embodiment can influence perceptions of spatial positioning in these groups. Further, our findings suggest that an important factor to consider in this regard is whether robot embodiment leads to a discernible orientation for the agent. Taken together, this means that it is important to consider robot embodiment when investigating spatial patterns of behavior in HRI. Also, researchers should carefully consider the assumption that human interaction data is a valid source of examples for creating perception models to reason about human-robot spatial behavior, as well as for implementing decision-making algorithms for robots to adapt to spatial formations. While people might reason about the spatial behavior of robots in a similar way to how they reason about the behavior of humans in many cases [7, 9], robot embodiment might alter these perceptions in some cases, resulting in different expectations.

7 ACKNOWLEDGEMENTS

This work was supported by the National Science Foundation (NSF), Grant No. (IIS-1924802). The findings and conclusions in this paper are those of the authors and do not necessarily reflect the views of the NSF. Thanks also to X. Ruiz for modeling the 3D environment.

REFERENCES

[1] Dan Bohus, Sean Andrist, and Eric Horvitz. 2017. A study in scene shaping: Adjusting F-formations in the wild. In *Proceedings of the 2017 AAAI Fall Symposium: Natural Communication for Human-Robot Collaboration*.

[2] Elwys De Stefani and Lorenza Mondada. 2014. Reorganizing mobile formations: When "guided" participants initiate reorientations in guided tours. *Space and Culture* 17, 2 (2014), 157–175.

[3] Eric Deng, Bilge Mutlu, and Maja Mataric. 2019. Embodiment in socially interactive robots. *arXiv preprint arXiv:1912.00312* (2019).

[4] Mar Gonzalez-Franco, Eyal Ofek, Ye Pan, Angus Antley, Anthony Steed, Bernhard Spanlang, Antonella Maselli, Domna Banakou, Nuria Pelechano, Sergio Orts-Escolano, et al. 2020. The Rocketbox library and the utility of freely available rigged avatars for procedural animation of virtual humans and embodiment. *Frontiers in Virtual Reality* 1 (2020), 20.

[5] Edward Twitchell Hall. 1966. *The hidden dimension*. Vol. 609. Garden City, NY: Doubleday.

[6] Hooman Hedayati, Annika Muehlbradt, Daniel J Szafir, and Sean Andrist. 2020. REFORM: Recognizing F-formations for Social Robots. *arXiv preprint arXiv:2008.07668* (2020).

[7] Helge Hüttenrauch, Kerstin Severinson Eklundh, Anders Green, and Elin A Topp. 2006. Investigating spatial relationships in human-robot interaction. In *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 5052–5059.

[8] Adam Kendon. 1990. *Conducting interaction: Patterns of behavior in focused encounters*. Vol. 7. CUP Archive.

[9] Hideaki Kuzuoka, Yuya Suzuki, Jun Yamashita, and Keiichi Yamazaki. 2010. Reconfiguring spatial formation arrangement by robot body orientation. In *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 285–292.

[10] Rui Li, Marc van Almkerk, Sanne van Waveren, Elizabeth Carter, and Iolanda Leite. 2019. Comparing human-robot proxemics between virtual reality and the real world. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 431–439.

[11] Paul Marshall, Yvonne Rogers, and Nadia Pantidi. 2011. Using F-Formations to Analyse Spatial Patterns of Interaction in Physical Environments. In *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work* (Hangzhou, China) *(CSCW '11)*. Association for Computing Machinery, New York, NY, USA, 445–454. https://doi.org/10.1145/1958824.1958893

[12] Ross Mead and Maja J Mataric. 2012. A probabilistic framework for autonomous proxemic control in situated and mobile human-robot interaction. In *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*. 193–194.

[13] Ross Mead and Maja J Matarić. 2017. Autonomous human–robot proxemics: socially aware navigation based on interaction potential. *Autonomous Robots* 41, 5 (2017), 1189–1201.

[14] Jonathan Mumm and Bilge Mutlu. 2011. Human-Robot Proxemics: Physical and Psychological Distancing in Human-Robot Interaction. In *Proceedings of the 6th International Conference on Human-Robot Interaction* (Lausanne, Switzerland) *(HRI '11)*. Association for Computing Machinery, New York, NY, USA, 331–338. https://doi.org/10.1145/1957656.1957786

[15] Sarah Sebo, Brett Stoll, Brian Scassellati, and Malte F Jung. 2020. Robots in Groups and Teams: A Literature Review. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW2 (2020), 1–36.

[16] Francesco Setti, Chris Russell, Chiara Bassetti, and Marco Cristani. 2015. F-formation detection: Individuating free-standing conversational groups in images. *PloS one* 10, 5 (2015), e0123783.

[17] Chao Shi, Michihiro Shimada, Takayuki Kanda, Hiroshi Ishiguro, and Norihiro Hagita. 2011. Spatial formation model for initiating conversation. *Proceedings of robotics: Science and systems VII* (2011), 305–313.

[18] Nan M Sussman and Howard M Rosenfeld. 1982. Influence of culture, language, and sex on conversational distance. *Journal of Personality and Social Psychology* 42, 1 (1982), 66.

[19] Mason Swofford, John Peruzzi, Nathan Tsoi, Sydney Thompson, Roberto Martín-Martín, Silvio Savarese, and Marynel Vázquez. 2020. Improving Social Awareness Through DANTE: Deep Affinity Network for Clustering Conversational Interactants. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW1 (2020), 1–23.

[20] Leila Takayama and Caroline Pantofaru. 2009. Influences on proxemic behaviors in human-robot interaction. In *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 5495–5502.

[21] Lili Tong, Audrey Serna, Simon Pageaud, Sébastien George, and Aurélien Tabard. 2016. It's Not How You Stand, It's How You Move: F-Formations and Collaboration Dynamics in a Mobile Learning Game. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Florence, Italy) *(MobileHCI '16)*. Association for Computing Machinery, New York, NY, USA, 318–329. https://doi.org/10.1145/2935334.2935343

[22] Nathan Tsoi, Mohamed Hussein, Jeacy Espinoza, Xavier Ruiz, and Marynel Vázquez. 2020. SEAN: Social Environment for Autonomous Navigation. In *Proceedings of the 8th International Conference on Human-Agent Interaction*. 281–283.

[23] Sebastiano Vascon, Eyasu Zemene Mequanint, Marco Cristani, Hayley Hung, Marcello Pelillo, and Vittorio Murino. 2014. A game-theoretic probabilistic approach for detecting conversational groups. In *Asian conference on computer vision*. Springer, 658–675.

[24] Marynel Vázquez, Elizabeth J Carter, Braden McDorman, Jodi Forlizzi, Aaron Steinfeld, and Scott E Hudson. 2017. Towards robot autonomy in group conversations: Understanding the effects of body orientation and gaze. In *2017 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 42–52.

[25] Marynel Vázquez, Aaron Steinfeld, and Scott E Hudson. 2016. Maintaining awareness of the focus of attention of a conversation: A robot-centric reinforcement learning approach. In *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 36–43.

[26] Michael L Walters, Kerstin Dautenhahn, René Te Boekhorst, Kheng Lee Koay, Dag Sverre Syrdal, and Chrystopher L Nehaniv. 2009. An empirical framework for human-robot proxemics. *Procs of new frontiers in human-robot interaction* (2009).

[27] Michael L Walters, Dag S Syrdal, Kerstin Dautenhahn, René Te Boekhorst, and Kheng Lee Koay. 2008. Avoiding the uncanny valley: robot appearance, personality and consistency of behavior in an attention-seeking home scenario for a robot companion. *Autonomous Robots* 24, 2 (2008), 159–178.

[28] Mohammad Abu Yousuf, Yoshinori Kobayashi, Yoshinori Kuno, Akiko Yamazaki, and Keiichi Yamazaki. 2012. Development of a mobile museum guide robot that can configure spatial formation with visitors. In *International Conference on Intelligent Computing*. Springer, 423–432.

[29] Gloria Zen, Bruno Lepri, Elisa Ricci, and Oswald Lanz. 2010. Space Speaks: Towards Socially and Personality Aware Visual Surveillance. In *Proceedings of the 1st ACM International Workshop on Multimodal Pervasive Video Analysis* (Firenze, Italy) *(MPVA '10)*. Association for Computing Machinery, New York, NY, USA, 37–42. https://doi.org/10.1145/1878039.1878048