A Drink-Serving Mobile Social Robot Selects who to Interact with Using Gaze

Oskar Palinko[1], Kerstin Fischer[2], Eduardo Ruiz Ramirez[1],
Lotte Damsgaard Nissen[2], Rosalyn M. Langedijk[2]

[1] SDU Robotics, Maersk Mc-Kinney Moller Institute, University of Southern Denmark, Odense, Denmark
[2] Department of Design and Communication, University of Southern Denmark, Sonderborg, Denmark
ospa@mmmi.sdu.dk, kerstin@sdu.dk, err@mmmi.sdu.dk, lniss06@student.sdu.dk, rla@sdu.dk

ABSTRACT

Robots will soon deliver food and beverages in various environments. They will need to communicate their intention efficiently; for example, they should indicate who they are addressing. We conducted a real-world study of a water serving robot at a university cafeteria. The robot was operated in a Wizard-of-Oz manner. It approached and offered water to students having their lunch. Our analysis of the relationship between robot gaze direction and the likelihood that someone takes a drink show that if people do not already have a drink and the interaction is not dominated by an overly enthusiastic user, the robot's gaze behavior is effective in selecting an interaction partner even "in the wild".

CCS CONCEPTS

• Human-centered computing → Empirical studies in HCI; • Computer systems organization → Robotics

KEYWORDS

Human robot interaction; social robotics; gaze tracking; in the wild

ACM Reference format:

Oskar Palinko, Kerstin Fischer, Eduardo Ruiz Ramirez, Lotte Damsgaard Nissen and Rosalyn M. Langedijk. 2020. A Drink-Serving Mobile Social Robot Selects who to Interact with Using Gaze. In *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction (HRI '20 Companion), March 23–26, 2020, Cambridge, United Kingdom.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3371382.3378339

1. Introduction

Robots are slowly making their way from factory floors to our everyday living environments. A potentially large area of application could be in the hospitality industry. For instance, they could be used for food preparation, food serving, luggage service, guiding, room service, or as receptionists. Despite the many potential use cases, their usage is still to take off [1].

HRI '20 Companion, March 23–26, 2020, Cambridge, United Kingdom.

ACM ISBN 978-1-4503-7057-8/20/03.
https://doi.org/10.1145/3371382.3378339

Figure 1. The mobile social robot a) front b) back and c) eyes.

One of the tasks that robots could successfully perform is food and drink serving. Ideally, such robots are socially-aware and equipped with profound human-robot interaction capabilities. They could also be used in health and elderly care, where caregivers could be freed up to focus on more critical tasks.

It is very important that these robots understand their physical and social environments. To navigate safely, they need to know where walls, obstacles and people are located, and they need to express their intentions to humans in a readable way [2]. For example, a robot can communicate with non-verbal signals, like its gaze direction. Eye gaze can be used to select who should talk to the robot next [3]. In controlled HRI-studies, robot gaze has been successfully employed to select an interaction partner (e.g.[4]), but little is known about influencing factors in real-world human-robot interaction scenarios. In this short paper we report on how a drink serving robot selects an addressee by directing its gaze at people in an unstructured environment.

2. Method

Our group has designed and developed a modular mobile social robot, Smooth, which could be used in the above described situation of food and drink serving, see Figure 1. Because of its modularity, the robot can be adapted to many different applications. In the current task, it carries a bin module with a tray on top. The mobile base of the robot has two driving and one caster wheel. It can be teleoperated using a gamepad. Its head is equipped for social interaction: it has microphones, speakers, cameras and two touch screens (front and back of the head). The front touchscreen displayed a pair of simulated eyes, with a white sclera and black pupils on a gray background. The robot was operated in a Wizard-of-Oz manner from a hidden location. The robot driver was responsible for navigation and controlling the robot's gaze: looking left, right or straight. Another wizard was

Figure 2. a) robot approaching students at the university cafeteria, as seen from the Wizards’ point of view, b) people interacting with the robot.

responsible for speech interaction: she listened through the robot’s microphones and produced speech by selecting from a pre-recorded set of utterances on an interface screen. These utterances followed this order: 1.) greeting (e.g. “Hello”, “Hi”, “Hi there”, “Sorry to bother you”, 2.) offering statement (e.g. “Can I offer you some water?”, “How about some water?”), 3.) persuasive sentences if water is not taken (e.g. “Research shows that it is important to drink enough water during the day”), 4.) water-related jokes and 5.) closing utterance (“Have a lovely day”, ”It was nice meeting you”).

The robot was deployed at a university dining hall with an open atrium, see Figure 2. It was approaching students on the ground floor while they were eating their lunch and offered them glasses of filtered water, for free. The two wizards were located on a higher floor of the atrium with clear lines of sight towards the robot while operating it inconspicuously. The students were informed of the presence of the robot with posters on large panels at the entrances to the area. After the interaction, they were asked to sign consent forms, which all except one did.

3. Analysis

We operated the robot at the cafeteria on two consecutive workdays during lunchtime for about an hour. We elicited a total of 58 interactions with students and staff. In 34 of these interactions, at least one person took a glass of water, while in 24 cases, no water was taken. Out of the 58 total, 26 interactions were with individuals, which are of less interest since here speaker selection is trivial (the interactions will however be used in further analyses later). The remaining 32 interactions concern groups of two to five participants (17 cases with two people, 11 cases with three, 2 cases with four and 2 cases with five people). Furthermore, within the group interactions we identified 13 situations in which one user *dominates* the interaction (for example, by coming close to the robot, leaning towards the robot or being eager to speak), see Table 1. In these situations, the robot’s gaze played little role as it was always the dominant individual who replied first. In this category, the person who the robot looked at spoke first in only five cases, while the dominant person spoke seven times. In contrast, in more *democratic* groups, group members acted on a similar level of engagement before the interaction began. In this group, gaze selection was most useful. Out of the total 12 democratic interactions, in 9 cases the person whom the robot looked at spoke first, while in only 3 this was not the case. Out of the 9 successful interactions, one was with a 5-member group, four were with 3-member groups and four with pairs of students.

Table 1. Number of interactions with two or more people split by category.

Dominant	person gazed at responds	5
	dominant person responds	7
Democratic	person gazed at responds	9
	someone else responds	3
Too close		8
Total (without individuals)		32

In these interactions, mutual gaze [5] worked successfully to select the conversation partner. In further 8 cases, group members were sitting too close to each other for robot gaze to be discernable.

In 41% of the 58 cases, people did not take any water. These were situations where students already had water or other beverages with them, or they simply didn’t feel like interacting with the robot. When they already had beverages, the success rate of water offering was 31%, while it was 69% when they had no beverages.

4. Discussion and Future Work

In this paper, we report on a part of a larger real-world study in which we tested a service robot offering drinks at a cafeteria. Concerning gaze selection, we found that group dynamics and physical closeness had a big influence on who will speak. For the *dominant* category, robot gaze was helpful in only 38% of the cases, while for *democratic* groups it was 75%. Thus, the effect of robot gaze is influenced by how enthusiastic some group members are about talking to the robot, as well as by their actual need for water. Regarding the correlation between gaze and the robot’s success rate, the amount of data gathered is too small for in-depth conclusions; however, it suffices to inform future studies. The real-world environment also taught us about the benefits and drawbacks of our setup. It seems that the fact that the serving tray was on the back of the robot (technical limitation) did not help the study, and thus the cup holding system needs to be redesigned. We also concluded that even though real-world studies are coveted in HRI, the outcomes are sometimes too complex to be able to draw statistical conclusions. Therefore, we plan to continue with both real-world and controlled studies with invited participants in controlled environments.

ACKNOWLEDGEMENT

This work was supported by the SMOOTH project (6158-00009B) of the Innovation Fund Denmark.

REFERENCES

[1] S. H. Ivanov, C. Webster, and K. Berezina, “Adoption of robots and service automation by tourism and hospitality companies,” *Rev. Tur. Desenvolv.*, vol. 27, no. 28, pp. 1501–1517, 2017.

[2] A. D. Dragan, K. C. T. Lee, and S. S. Srinivasa, “Legibility and predictability of robot motion,” in *ACM/IEEE International Conference on Human-Robot Interaction*, 2013.

[3] 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,” *Hum. Factors*, vol. 2, no. 1, pp. 61–68, 2009.

[4] Y. Matsuyama, I. Akiba, S. Fujie, and T. Kobayashi, “Four-participant group conversation: A facilitation robot controlling engagement density as the fourth participant,” *Comput. Speech Lang.*, 2015.

[5] M. Argyle and M. Cook, “Gaze and mutual gaze.,” 1976.