The Function of Off-Gaze in Human-Robot Interaction

Sascha Hinte and Manja Lohse

***Abstract*— When and how do users interrupt the interaction with a robot and turn to the experimenter? Usually it is assumed that experimenters affect the interaction negatively and should ideally not be present at all. However, in interaction situations with autonomous systems and inexperienced users this is often not possible for safety reasons. Thus, the participants indeed at times switch their focus of attention from the robot to the experimenter. Instead of seeing this as something purely negative, we argue that answering the questions of when, why and how this happens actually bears important information about the state of the interaction and the users' understanding of it. Therefore, we analyzed a study conducted in a home tour scenario with this respect and indeed discovered certain situations when the users turned away from the robot and towards the experimenter.**

I. Introduction

RESEARCH in human-robot interaction (HRI) usually exclusively focuses on the interaction between the robot and the users. However, in HRI user studies, in addition to the robot and the users, often the experimenter is present for safety reasons. Especially if the robot is mobile, leaving the user alone with it would often be highly unethical. Moreover, the users might even act more natural and self-assured if the experimenter is present. However, the experimenter certainly influences the interaction (see Section II). While these effects are generally seen as something negative, we claim that they could actually contribute to interaction analysis in a valuable way.

When analyzing data from a user study (see Section III), we noticed participants' gazes directed to the experimenter which we are going to call off-gaze. These off-gazes interrupt the flow of the HRI because the user turns away from the robot (see Figure 1). Therefore, we wondered in which situations off-gazes occurred and what we could learn from them about the state of the interaction between participant and robot. These are the questions that we want to address in the following. We believe that answering them bears important information that exceeds the information which we can obtain from analyzing the interaction between the user and the robot.

In Section II we introduce some related work. Section III describes the user study in which the data was acquired and the data analysis process. Section IV summarizes the results. The paper concludes with a discussion of the findings.

Manuscript received March 17, 2011.

S. Hinte is with the Research Institute for Cognition and Robotics (CoR-Lab) and the Applied Informatics Group, Bielefeld University, Bielefeld, 33615 Germany (e-mail: sascha.hinte@uni-bielefeld.de).

M. Lohse is with the Research Institute for Cognition and Robotics (CoR-Lab), Bielefeld, 33615 Germany (phone: +49 521 10612209; fax: +49 521 1062992; e-mail: mlohse@techfak.uni-bielefeld.de).

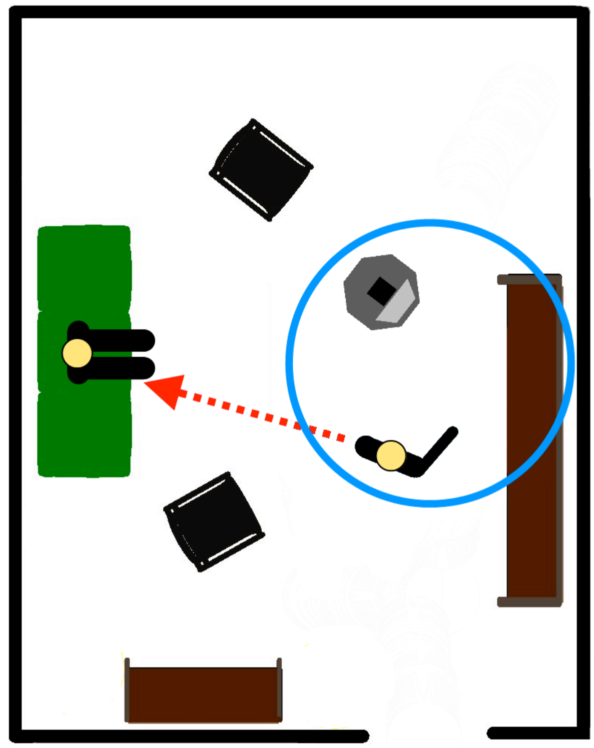

Figure 1. User in interaction with the robot off-gazing at the experimenter

II. Related Work

Feedback is highly important to secure understanding in interaction (see, for example, [1]). Especially in task-oriented interactions, clarification is often necessary in order to ensure mutual understanding and the completion of the task [14]. Hence, if a problem occurs in the interaction with a robot and a competent third person is present, why should the user not ask for help?

Indeed, it can be assumed that the presence of the experimenter affects the interaction. Firstly, so-called observer-expectancy effects have been reported in the literature. These describe the subconscious influence of the experimenter on the participant. One famous example of the observer-expectancy effect from the 1900s describes Wilhelm von Osten believing that his horse (Clever Hans) was capable of solving mathematical problems. In fact the horse was not but von Osten unknowingly provided it with cues that enabled it to answer the problems correctly [9]. More generally spoken, it has been found that the experimenter's personality and her personal orientation about the experimental outcome influence the participants (see, for example, [11]).

Secondly, subject-expectancy effects have been reported. These effects cause the subjects to act in reaction to what they believe is the desired behavior. Subject-expectancy effects have also been termed demand characteristics [8]. Originally, Rosenberg [10] believed that subjects desire to be evaluated in a positive way by the experimenter. However, later research by Weber and Cook [13] showed that the subjects might indeed take different roles. This can also be found in HRI, for example, when users test the limits of the system instead of completing the task.

Next to these psychological processes, the presence of the experimenter can simply lead to situations in which the participant tries to talk to her. Talking to the experimenter is one kind of off-talk. Oppermann and colleagues [7] have identified several types of off-talk like thinking aloud or conversation with other persons present. They define off-talk as "every utterance that is not directed to the system as a question, a feedback utterance or as an instruction" ([7], p.1). In our data we found very few occurrences of off-talk. In 14 runs, off-talk occurred only six times. However, we encountered that the users often looked at the experimenter. Also Batliner et al. [2] described occurrences of what they called "off-view". In their study, the users had to focus their attention on a screen while solving a task. Every view away from the screen was called off-view. Instances of off-view were found when the users reported their progress to someone else or were interrupted in the process of use. They were connected to direct interruptions which were not intended in our study. However, in previous work, also Lohse [4] and Vollmer et al. [12] have considered off-view in HRI by coding gaze behavior of the users using the categories "at the robot", "at relevant objects" and "somewhere else". The last category is similar to off-view. However, in the work presented here we do not want to focus on off-view but rather on off-gaze. In contrast to view, the term gaze implies an intentional behavior [6] which is important here as we focus on situations in which users intentionally gaze away from the robot and at the experimenter. In these situations the flow of the HRI is interrupted. Therefore, the questions that should be answered in the following are why and when the participants intentionally gaze at the experimenter?

III. User Study

The off-gaze analysis was conducted on data from a home tour user study. Details about the study will be given in this section. We describe the study procedure in order to explain the spatial relations between participants, robot, and experimenter. These details are important to understand the off-gaze behavior and the following analysis.

A. *Home tour scenario*

In the home tour scenario the robot is guided through an apartment (see Figure 2) by the user and learns about rooms and objects. This is necessary because it is not possible to pre-program the robot for every environment. Hence, it has

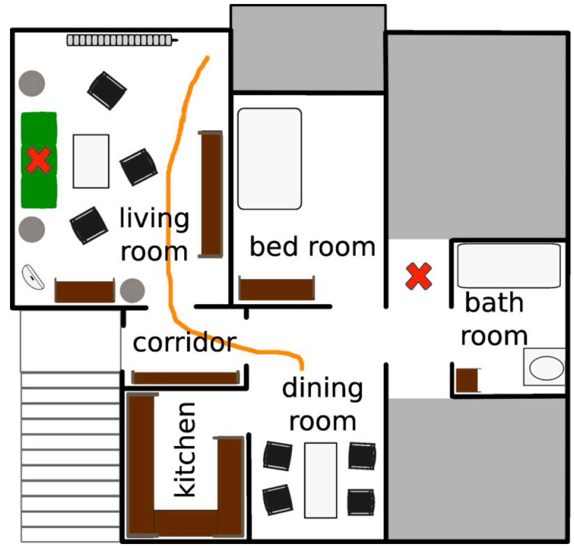

Figure 2. Apartment floor plan, path the robot was guided (orange line), and positions of the experimenter (red crosses)

to learn about the apartment with the help of the user. Once the robot has acquired the knowledge about its surroundings, it can serve as a kind of "butler" providing personal services (for example, laying the table, cleaning rooms). In the current implementation, the learning is the prominent part of the scenario.

In the home tour scenario, users are likely to be novices or advanced beginners, meaning that they have little or no experience interacting with a personal service robot for the home [3]. In fact, in the user studies presented in the following, most of the trials were first-contact situations, such that the users had never interacted with a social robot before. As the robot is intended for long-term interaction, one challenge of the scenario is that the users must be enabled to handle the robot easily and to learn about it over time. This is important to know because in realistic situations the users would not be able to turn to the experimenter for help. Thus, it is necessary to find out when and why this happens in order to avoid such situations in future interactions.

B. *Procedure*

On arriving in the robot apartment the participants were welcomed and introduced to the study using a written description of the scenario. They answered a questionnaire on demographic data and their experience interacting with robots. Thereafter, the participants were guided into the living room where the robot was waiting ready for operation. The users received a short training, during which they were allowed to ask questions to the experimenter. The answers to the questions were kept as similar as possible trying not to give biasing information to the participants. Most questions in fact were turned off by motivating the participants to try the robot out themselves and by ensuring them that they could not make any mistakes.

After the tutorial session, the participants carried out the main task which consisted of the following steps:

- guide the robot from the living room to the dining room via the corridor
- show and label the living room and the dining room
- show the bookshelf in the living room and the floor lamp in the dining room
- ask the robot to go back to one of the rooms or objects that had been learned previously and follow it

During the interaction in the living room the experimenter was sitting on the sofa (X on the left in Figure 1) which was located opposite the bookshelf that the participants showed to the robot. Thus, the experimenter was as much out of sight as possible. When the participants guided the robot to the other rooms, the experimenter only entered the corridor when the participants had proceeded to the dining room and then quickly crossed the dining room in order to stand around the corner in front of the bath room (X on the right in Figure 1). The experimenter (being the same person throughout the study), tried to keep this behavior as similar as possible in all trials and did not initiate interaction with the participant. She also tried to avoid gazing at the participants.

After the interaction, the participants answered a second questionnaire which included items on liking of the robot, attributions made towards the robot, and usability of the robot.

C. *Participants*

14 participants (7 male, 7 female) took part in the study. Their age ranged from 18 to 54 years (average 38.9 years, SD=11.83). As mentioned above, their experience with robots was very limited. All participants were German native speakers and interacted with BIRON in German. They were recruited at a public event of the university.

D. *Data*

Throughout the study, videos were recorded from different angles with a hand-held camera and two stationary cameras. The handheld camera allowed us to be close enough to the participants most of the time in order to identify their gaze direction. If gaze direction could not be determined, for example, because the participant just left the room, these passages were not coded and not included in the analysis. All coding was done using the annotation software ELAN[1]. First of all, the gaze direction of the users was annotated. We differentiated gaze at the robot, relevant objects, somewhere else and off-gaze.

Altogether, 790 gazes were annotated. 433 of these were directed at the robot, 106 at objects, and 169 somewhere else. The off-gaze category included 82 annotations which corresponded to a mean of roughly 6 instances per user and 10% of all gazes. These occurred in 94 minutes and 24 seconds of video material which equals an average of 6 minutes and 44 seconds for each person. Roughly one off-gaze per minute was annotated. These numbers show that in relation to the overall number of gaze annotations, the off-gaze is quite frequent and should really be taken into account. However, it should also be mentioned that four people did not show any off-gaze behaviors at all even though their interactions with the robot did not differ from other participants'. This result is probably related to the role that these people took in the interaction (see Section II). It can only be hypothesized that they knew about the experimenter effects and consciously avoided gazing at the experimenter. Another explanation could be that these people simply did not expect the experimenter to help them. In the future, the participants should be asked why they looked at the experimenter or not. Unfortunately this was not done in the study presented here.

However, for the analysis we coded further data from the videos such as the speech of the users segmented into meaningful utterances and movements of human and robot. The robot's utterances were retrieved from the log files. This information was required in order to answer the question of what caused the off-gaze behavior. All annotations together were analyzed and described qualitatively. Situations that were similar to each other were grouped. The resulting categories are introduced in the following.

IV. RESULTS

As mentioned above, the off-gaze behaviors were grouped which resulted in eight categories classified by the causes of the situations (see Table 1).

The most common cause for off-gaze was a *lack of response* by the robot (24 cases, 29% of the cases). The users were waiting for the robot to do something in order for the interaction to continue. Lack of response was coded

Table 1. Causes of off-gaze behaviors, description, and number of situations belonging to each group (#)

Cause	**Description**	**#**
Lack of response	the robot does not respond	24
Problem understanding	the robot has not understood what the user just said and asks for repetition	16
Uncertainty about task	the user is not sure which room or object to show next	14
Uncertainty about utterance	the user is not sure if the system understood the instructions	4
Problem moving	the robot is moving very slowly or in an unusual way	9
Success	the robot has learned the name of an object or room correctly	3
Joint attention	the robot looks at the experimenter and the user follows its gaze	3
Anticipation	the user announces by gaze that she and the robot will move into the direction of the experimenter	4
	Other	5
	Total	82

[1] http://www.lat-mpi.eu/tools/elan/

whenever the waiting phase was longer than 5 seconds. The off-gaze behaviors occurred after a mean lack of response of 7.5 seconds (standard deviation 5.6).

For comparison, all lack of response situations in which no off-gaze occurred were coded. Altogether 53 annotations were made which equals roughly twice the number of situations in which the participants actually looked at the experimenter. In other words, in every third lack of response situation the participants gazed at the experimenter. If they did not, they usually tried to get a reaction from the robot by repeating their utterance or saying something different. Only if this did not work, they gazed at the experimenter. This result is in line with previous findings about how people change their discursive behavior in HRI [5]. In this previous study it was shown that a lack or delay of robot reactions often led the users to try out different behaviors.

The second cause for off-gaze was *speech understanding problems* by the robot (16 cases, 20% of all cases). The users learned about these situations because the robot asked them to repeat their utterance, said that it had not understood, or gave an inappropriate answer. Example dialogs could look like this:

1 Robot: You want me to show you something, right?
User: Yes.
Robot: I am sorry, I did not understand what you just said.
User: *off-gaze*

2 User: This is the shelve.
Robot: This is the living room then.
User: *off-gaze*
No. This is the shelve.

Instead of simply repeating the utterance the users looked at the experimenter as if they were not sure why the robot did not understand them.

Again, for comparison all speech understanding problems were coded that did not cause off-gaze behaviors. Out of an overall 581 robot utterances, 155 signaled speech understanding problems. This shows that the robot quite frequently (in almost 30 percent of the utterances) showed the user that it had not understood her. This is not surprising given the current state of the art of performance of speech recognition systems in noisy environments with inexperienced users. However, it also means that the users actually gazed at the experimenter in only 9% of the speech understanding problems.

The third cause for off-gaze was the users' *uncertainty about the task* (14 cases, 17% of all cases). Before the main task, the users had been instructed to teach the robot the names of several objects and rooms while moving through the apartment. After successfully teaching a name, they sometimes were uncertain which room or object to show or where to go next. In fact, this behavior was not directly connected to the HRI but the participants rather used off-gaze to request information from the experimenter that the robot could not provide. However, the uncertainty about the task interrupted the flow of the interaction. Instances in which the users addressed the experimenter were easily assigned to this group based on the content of the question or obvious gestures. This was not as easily possible for cases without off-gaze that needed to be identified for comparison. Therefore, criteria were defined to identify these cases. The first criterion (as for off-gaze situations) was the occurrence right after the end of the last task. The second criterion was the fact that the users did not initiate the next task immediately after having finished the previous one. Finally, the third criterion was a corresponding facial expression and/or gesture which could clearly be identified as uncertainty about what to do next (for example, questioning facial expression and pointing in a certain direction meaning something like "should I guide the robot to this location?"). With these criterions, eight sequences were identified that did not co-occur with an off-gaze. Thus, when obviously uncertain about the task, the participants more often asked the experimenter than they tried to determine what they needed to do next without asking.

Next to uncertainty about the task also *uncertainty about utterances* (4 cases, 5% of all cases) occurred. These cases were rare, because usually the users sticked to the utterances which they had learned in the training session before the main task. However, if the robot did not understand them, they tried out other utterances. In this case, looking at the experimenter probably meant something like "Does the robot understand this utterance?". As these cases were rare, we do not want to analyze them further.

Also *problems with the movements of the robot* led to off-gaze (9 cases, 11% of all cases). It sometimes happened that the robot moved very slowly or in an unusual way (for example, it turned left and right instead of moving forward because it got stuck in the carpet). In these cases, the users were not sure whether the robot would actually go to where it should go as they also repeatedly stated in the interviews after the interaction. In another 17 cases the robots' problems moving did not lead to off-gaze. However, the analysis did not reveal any differences between these sequences that might have led to off-gaze or not.

Besides all the problems, in some cases (3 cases, 4% of all cases), also *success* caused off-gazes. In situations when the dialog had been problematic and the robot finally learned the name of the room or object the users shared their happiness about this by looking at the experimenter. Again, we coded similar situations without off-gaze. Eight similar situations without off-gaze could be identified. Thus, the low overall number of cases as well as the relation to cases without off-gaze indicates that this off-gaze behavior is not very common. Furthermore, analyzing it would not help the HRI much as these cases do not point to problems in the interaction that need to be fixed.

In another 3 cases (4% of all cases) the participants followed the gaze of the robot that accidentally looked at the experimenter with its pan-tilt camera. We called these *joint* *attention* as the users and the robot jointly focused their attention on the experimenter. These cases will not be analyzed in more depth as the off-gaze was rather initiated by the robot and not by the participant. Thus, the users did not have an intention of their own when off-gazing.

The last group is entitled *anticipation* (4 cases, 5% of the cases). It describes cases in which the users instructed the robot to show them an object located in another room and the experimenter stood in the way (even though she tried to avoid this). Thus, the off-gaze informed the experimenter that the robot was coming into her direction. It was not primarily information about the interaction with the robot but rather information about the experimenter disrupting the interaction flow. These cases were not intended and should ideally be avoided in future studies.

V. Discussion

The previous section has described different causes for off-gaze. In the following, we will further discuss what the off-gaze tells us about the interaction in our study and about HRI in general. We want to focus on the most frequent causes of off-gaze (lack of response, problem understanding, uncertainty about task, and problem moving) which made for 77% of all cases in the current study.

Looking at these groups, it can be found that they occurred at different points of time in the interaction. Lack of response occurred during the tasks at times when the system log files did not show any system output. Problem understanding also occurred during the tasks soon after an utterance of the robot. Uncertainty about tasks occurred after the previous task had been finished and before a new one had been initiated. Finally, problem moving occurred during the guide or show task when the system log files showed that the robot was driving / changings its position. Based on this differentiation, the off-gaze indicates the current state of the interaction.

In the case of lack of response, the off-gaze signals that the robot is really much too slow and the users urgently need information about its current state and about actions they need to take. This finding was underlined by the questionnaire ratings of the robot's speed (mean rating of .93 (SD=.70) on a scale 0 (slow) to 4 (fast)). Obviously the robot itself interrupted the flow of the interaction by being so slow. In order to avoid this, the robot either has to become much faster or at least it employs strategies to signal that it is still running well but needs some time to compute an appropriate response. Such strategies could be an animation on a display or a blinking light.

Strategies to handle the situations are obviously already present for the problems understanding group. In only 9% of these cases the users turned to the experimenter for help with an off-gaze behavior. Most often this happened if the robot either repeatedly did not understand simple utterances like yes or no or if its answer to a users' utterance was really out off place. For HRI in general these cases teach us how often individual people are ready to repeat utterances and to accept that the robot has still not understood. The goal should be to understand the users' utterances within these limits. Moreover, the system needs to recognize users' reactions that show that an utterance was not appropriate. If the system recognized such reactions it could ask back if and why its own utterance was inappropriate and use this information in order to improve its own perception of this and future situations.

The same is true for the problem moving situations. The robot could exploit the users' reactions to determine whether its movement is appropriate or not. In one-third off the cases where it was not, the users actually produced an off-gaze behavior.

In contrast to these three groups, in the case of uncertainty about the task, the robot did not do anything wrong. Thus, this case also has little influence on the HRI. It can be assumed that it does not appear once the users themselves decide what tasks the robot should complete. For now, however, the analysis of these cases shows that the experimenter was actually accepted as an expert on the task that the user needed to complete in the study. If necessary, the robot could be enabled to also be an expert on this by proposing a task which it would like to do next.

VI. Conclusion and future work

The paper has answered the questions of when and why the users gazed at the experimenter in a home tour study. It has indeed been found that off-gaze behavior bears a lot of information about the interaction. Altogether the analysis has shown that we can obtain information about the state of the HRI by looking beyond it and by also taking the interaction between the user and the experimenter into account. In scenarios like the one presented here, in which the experimenter needs to be present for security reasons, such an approach can be explored instead of seeing the presence of the experimenter as something merely negative.

For future work, eye-tracking systems could be used to better capture the gazing direction of the user because annotating it by hand from the videos was actually quite laborious. Moreover, it will be interesting to research how off-gaze behavior changes in situations in which multiple users interact with the robot at the same time. The presence of other users might change the results as the other users might be looked at instead of the experimenter. This would actually be positive because it means that the users themselves can handle the interaction without the help of an expert which should be the goal of HRI design.

Acknowledgments

We thank the reviewers for the helpful comments on the earlier version of this paper. Furthermore, we thank the participants for their time and effort. Finally, our thanks go to Frederic Siepmann for his help in conducting the study.

REFERENCES

[1] J. Allwood, J. Nivre, Joakim, E. Ahlsen, On the semantics and pragmatics of linguistic feedback. *Journal of Semantics*, 9 (1), 1993, pp. 1–26.

[2] A. Batliner, C. Hacker and E. Nöth, (2006). To Talk or not to Talk with a Computer: On-Talk vs. Off-Talk. In How People Talk to Computers, Robots, and Other Artificial Interaction Partners, SFB/TR 8 Report No. 010-09/2006, pp. 79-100.

[3] J. T. Hackos and J. C. Redish, *User and task analysis for interface design,* Wiley and Sons: New York, 1998.

[4] M. Lohse, "Social, functional, and problem-related tasks in HRI - a comparative analysis of body orientation and gaze", Proc. *2nd AISB Workshop on New Frontiers in Human-Robot Interaction*, Leicester, UK, 2010.

[5] M. Lohse, K. J. Rohlfing, B. Wrede and G. Sagerer. "Try something else!" — When users change their discursive behavior in human-robot interaction, in Proc. of *International Conference on Robotics and Automation*, 2008.

[6] Merriam Webster Dictionary, http://www.merriam-webster.com/dictionary/gaze. accessed March 14, 2011.

[7] D. Oppermann, F. Schiel, S. Steininger, and N. Beringer, Off-Talk - a Problem for Human-Machine-Interaction? In Proc. Eurospeech01, Aalborg, 2001, pp. 2197-2200.

[8] M. T. Orne, Demand characteristics and the concept of quasi-controls. In R. Rosenthal & R. L. Rosnow (Eds.), Artifacts in behavioral research. New York: Academic Press, 1969.

[9] O. Pfungst, Clever Hans (The horse of Mr. von Osten): A contribution to experimental animal and human psychology (Trans. C. L. Rahn). New York: Henry Holt, 1911.

[10] M. J. Rosenberg, The conditions and consequences of evaluation apprehension. In R. Rosenthal and R. L. Rosnow (Eds.), *Artifacts in behavioral research*. New York: Academic Press, 1969.

[11] R. Rosenthal, *Experimenter effects in behavioral research*. New York: Appleton-Century-Crofts, 1966.

[12] A. - L. Vollmer, K. Lohan, K. Fischer, Y. Nagai, K. Pitsch, J. Fritsch, K. Rohlfing, and B. Wrede, "People Modify Their Tutoring Behavior in Robot-Directed Interaction for Action Learning", *International Conference on Development and Learning*, vol. 8, Shanghai, China, 2009.

[13] S. J. Weber and T. D. Cook, Subject effects in laboratory research: An examination of subject roles, demand characteristics, and valid inference, *Psychological Bulletin*, Volume 77, Issue 4, April 1972, pp. 273-295.

[14] B. Wrede, S. Kopp., K. Rohlfing, M. Lohse and C. Muhl, Appropriate feedback in asymmetric interactions. In *Journal of Pragmatics*, 42, 2010, pp. 2366-2384.