Small groups in human-robot interaction

Manja Lohse

CoR-Lab, Bielefeld University
Universitaetsstrasse 25, 33615 Bielefeld, Germany
mlohse@techfak.uni-bielefeld.de

Abstract. Most research in human-robot interaction so far has focused on one robot and one person. Such situations have been evaluated in order to enable the robot to act appropriately. However, often interaction occurs in groups. Groups, even if they are small, require novel robot abilities such as detecting who is talking to whom or how to behave. The paper focuses on small group research and relates it to human-robot interaction by introducing very first results of an experimental user study with 80 users (40 pairs) that was conducted to determine the perception of the robot in a small group object-teaching scenario.

Keywords: human-robot interaction, small groups, user study

1 Introduction and related work

Groups of people work together in organizations, they meet for parties and sports events, or form a family. Hence, a large portion of human behavior occurs in groups and systematic research on small (human) groups by social psychologists has started as early as the 1930's [5]. Many definitions have been proposed for the term *group*, emphasizing characteristics such as interdependence, communication, structure, and shared identity of the participants [5]. Levine and Moreland [5] rather agree with the view of "groupiness" as a social dimension. Thus, every group has a certain amount of groupiness that depends on the number of members and their relationship with each other.

In the work presented here we focus on small groups consisting of two people and one robot. The question for human-robot interaction (HRI) with respect to such small groups is how the robot fits into them? How does it have to behave in order to be helpful for achieving goals and satisfying users' needs? And what abilities does it need to deal with group situations? As a first step towards answering these questions, an experimental Wizard of Oz (WoZ) user study with 80 participants (40 pairs) was conducted. The task of the participant pairs in the study was to teach objects to a robot. The robot behaviors were manipulated regarding the robot addressing the users in singular or plural (one at a time or both at the same time) and the robot recognizing the first object that was taught or not. The goal of the manipulations was to test two assumptions of how robot behavior could influence small group interaction and in-group relations.

In this work, the in-group relations are considered based on different factors. One important variable that has been found to determine group relations is closeness between the participants. Therefore, we used parts of the relationship closeness inventory (RCI) [3]. The RCI measures relationship types (friends, family, etc.), the time that people spend together, the activities they do together, and influence on thoughts, feelings, and behavior of the other person. However, the RCI focuses on North American college students [2] and is not suitable for comparison of different populations of people or robots (e.g., we cannot ask users how often they go to the bar with a robot). Therefore, we also measured closeness between the participants and between the participants and the robot with the Inclusion of Other in the Self Scale (IOS scale, [2]) which is a simple pictorial scale that can be employed to measure robot-human relations. We were further interested in the participants' willingness to collaborate. That is why we used the WTC-Trait [1] which is conceived as active communication involvement with others in the process of decision-making. As assertiveness is one important factor of the WTC-Trait, we additionally used the Rathus Assertiveness Schedule (RAS) [6]. Given the page limit, WTC and RAS shall not be discussed here.

2 Experimental user study

As mentioned above, we conducted an experimental WoZ user study with the robot BIRON [7]. Next to generally collecting ratings of the robot in a group setting, we tested the two manipulations described in the introduction. Moreover, all participants completed a positive and a negative trial that differed in the percentage of objects that the robot learned correctly (50% vs. 25%) and incorrectly (15% vs. 40%). This procedure was chosen based on previous work [4]. All conditions were translated into scripts of verbal robot behaviors. Additionally to the utterances in the scripts, the participants were greeted and asked for their names and the robot said good-bye. In the singular condition only one of the participants was asked for his/her name and also greeted with his/her name whereas in the plural condition both were. Moreover, the scripts were adapted as in German there are different forms of "you" depending on the number. In all scripts, 4 out of 20 utterances (20%) were manipulated. The variation between the conditions recognizes first object/does not recognize first object consisted of exchanging the first and the second utterance (success and failure).

Ten pairs of participants were assigned to all four conditions which equals an overall number of 80 participants (50 female, 30 male). 17 pairs were female/female, 16 female/male, and 7 male/male. Most participants were students. Their age ranged between 20 and 57 years (mean age=25.46 years, sd=5.90). They came from all kinds of disciplines (sociology, linguistics, law, etc.) and had some experience with computers (mean=2.98 (all means on a scale of 0 to 4), sd=.80) but hardly any with programming (mean=.86, sd=1.18) and robots (mean=.51, sd=.87). Their mean IOS rating was 4.71 (sd=1.73) (on a scale from 0 to 7). More than half the participants (45) felt very close to each other (IOS score between 5 and 7). The IOS scale ratings were correlated with the time people had known each other prior to the user study (r=.277*), with the time they spend together each week (mean=25.54 hours, sd=39.32; r=.479**), and with the number of days they see each other in a normal week (mean=3.65 days, sd=2.16; r=.551**). In other words, the closer the participants felt to each other, the longer they had known each other and the more time they spent together. The mean IOS rating of the robot before the interaction was 1.36 (sd=0.73). This value is significantly lower than the IOS between the participants (t(158)=15.948, p<0.0001**). Hence, the participants knew each other much better than the robot.

3 Results

The users were neither very excited about BIRON's behavior nor deeply disappointed (mean=2.10 (all values on a scale of 0 to 4), sd=0.92). The mean rating of fun was significantly higher (mean=2.74, sd=0.98; t(79)=5.872, p<0.0001**). At the same time, both values were strongly correlated (r=.479**). Thus, people who were satisfied with the robot's behavior also had more fun. Regarding success of completing the task with BIRON (mean = 2.66, sd=0.81), the participants felt that they completed the task but could have done better. This finding was supported by ratings of how fast the task was completed (mean=1.79, sd=0.74). The time the participants spent with BIRON felt rather short to them (mean=1.69, sd=1.03). Of course, this result is also related to the actual interaction time which was rather short (about 15 minutes).

Not only behavior and fun were strongly correlated. Correlations between all items mentioned in the previous paragraph were between .274* and .479**. The only exception was efficiency and the time spent with the robot (r=-.205). This value was only marginally significant (p=.068). To sum up, if the participants were more satisfied with the robot's behavior they also had more fun, felt more successful and efficient, and felt that they had spent less time with the robot which is an indicator of flow.

Next to the correlations presented above, fun was also found to be correlated with the IOS ratings of the robot after the interaction (r=.221*). This means that the more fun the participants had, the closer they felt to the robot after the interaction. As the IOS of the robot after the interaction was not correlated with any other of the five items, it seems that fun is even more important for the feeling of closeness than robot behavior and task completion.

Nevertheless, we also analyzed: Who got along better with BIRON? (all items on a 5-point scale from clearly me (2) to clearly the other person (-2)) (mean 0.08, sd=0.74); Who talked more with BIRON? (mean=-0.10, sd=0.84); Who taught more objects to BIRON? (mean=-0.06, sd=0.54); Who received more attention from BIRON? (mean=-0.13, sd=0.49); and Whom did BIRON understand better? (mean=-0,06, sd=0.82). An analysis of the frequencies of the items revealed that the majority of the participants felt that the interaction was rather balanced and none of the participants dominated the interaction.

We also correlated these five items with the five general ratings introduced above. The item "Who got along better with BIRON?" was correlated significantly with all of them (satisfaction: r=.303**, fun: r=.237*, success: r=.463**; efficiency: r=.236*) but the time spent with the robot. Thus, if participants felt that they got along better with the robot than the other person, they also were more positive about the interaction as a whole. "Who taught more objects to

BIRON?" was only correlated with success (r=.330**) and efficiency (r=.380**). This result is intuitive as these were the task related measures. "Who received more attention from BIRON?" was positively correlated with fun (r=.223*). This result indicates that the more attention people get, the more fun they have in the interaction. Considering the results of the IOS ratings, this also means that they feel closer to the robot. "Whom did BIRON understand better?" was only correlated with success (r=.255*). Hence, if speech recognition worked better, the participants felt that they had taught more objects to the robot.

Finally it shall be mentioned that none of the experimental manipulations (singular vs. plural and recognizing first object or not) showed any effects. ANOVAS revealed that the groups did not differ regarding the felt closeness towards the robot or the general evaluation. In future research we will try to determine reasons for this.

4 Conclusion

The study has provided many insights of users' perception of an object-teaching task in a group setting. This is the first work we know of in this direction, taking concepts such as closeness, willingness to cooperate, assertiveness, and user experience into account. However, the study also raised new questions. First of all, factors of robot behavior need to be identified that actually influence the perception of the system, are decisive for a positive interaction, strong group identity, and good task performance. Moreover, methods to measure such factors are needed. Also effects of attention and relationship between the participants need to be more closely researched. A first step in this direction will be the analysis of the videos recorded in the study, however, much more work will be needed to develop robots that are meaningful members of groups.

References

1. Anderson, C.M., Martin, M.M.: Willingness to collaborate as a new communication trait: Scale development and a predictive model of communication traits. Central States/Southern States Communication Associations, Lexington (1993)
2. Aron, A., Aron, E.N., Smollan, D.: Inclusion of other in the self scale and the structure of interpersonal closeness. Journal of Personality and Social Psychology 63, 596–621 (1992)
3. Berscheid, E., Snyder, M., Omoto, A.M.: The relationship closeness inventory: Assessing the closeness of interpersonal relationships. Journal of Personality and Social Psychology 57, 792–807 (1989)
4. Lang, C., Hanheide, M., Lohse, M., Wersing, H., Sagerer, G.: Feedback interpretation based on facial expressions in HRI. RO-MAN'2009 (2009)
5. Levine, J.M., Moreland, R.L.: Small groups. Psychology Press (2006)
6. Rathus, S.A.: A 30-item schedule for assessing assertive behavior. Behavior Therapy 4, 398–406 (1973)
7. Wachsmuth, S., Siepmann, F., Schulze, S., Swadzba, A.: ToBI - Team of Bielefeld: The HRI System for RoboCup@Home. Tech. rep., Bielefeld University (2010)