25th IEEE International Symposium on
Robot and Human Interactive Communication (RO-MAN)
August 26-31, 2016. Columbia University, NY, USA

Analysis of Multi-party Human Interaction towards a Robot Mediator

Hisato Fukuda, Yoshinori Kobayashi, Yoshinori Kuno, *Member, IEEE*, Akiko Yamazaki, Keiko Ikeda and Keiichi Yamazaki

***Abstract*— In this research, we aim to reveal the difference between robot mediated multi-party human interaction and human mediated multi-party human interaction. We do this by examining and comparing the multi-party participants' interactions towards the robot human mediators in regard to the mediator's questions. From this experiment, we found that the participants engaged in interaction among each other more when the questioner was the robot. This paper suggests a concept and a design of robots, which enhances social interaction among humans because of its intervention.**

I. Introduction

In order to provide visitors with unique experiences in museums and have visitors enjoy the museum, we have developed a museum guide robot to provide explanations for exhibits. This robot focuses on how robots can combine visible actions with speech to effectively explain exhibits to multiple visitors. The motions of this robot are designed based on our ethnographical study using sociological methodologies and has been developed to adopt various human behaviors in its movements. This robot will first ask a question related to the exhibits and then selects an answerer by observing the visitors' reactions with its vision system [14]. Fig. 1 shows a Robovie [17] that we used as a robot platform in our experiments, and Fig. 2 shows an example scene from the experiments.

When we pursue observations on human participants' reactions vis-à-vis the museum guide robot, we find that quite often, those who are not appointed a turn by the guide robot speak up, or consult with the one actually appointed by talking or looking at each other. In human interaction, turn-taking behavior is universal and prevails [8]. In particular, the asking of questions is one of the most effective strategies for the current speaker to select the next speaker. The finding that participants may engage in more interaction among themselves when the questioner is a robot is an important difference in the participants' behavior from when they face a human guide. In both cases (i.e., human guide or robot guide) participants do respond and answer to the question given to them, this is common. When they are dealing with the guide robot, however, interaction among the multiparty participants is prone to take place. Some may see increased interaction among participants during the museum guide's talk to be distracting. However, in museum studies, the interactive appreciation method has also gained attention. In this method, a guide or a teacher frequently asks visitors questions in order to include them in a richer museum experience [3]. By posing questions, active interactions are achieved between the guide and each visitor as well as among visitors as well. In school settings, the teacher asks a question in order to encourage participation with his/her pupils. For these reasons, we view this emergence of social interaction among multiple human participants in one setting because of the robot's presence as a positive result.

Figure 1. Robovie.

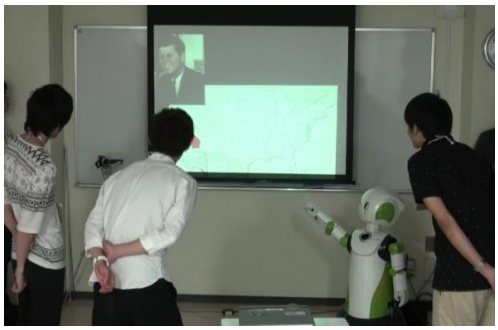

Figure 2. An example scene from the experiments.

Many researchers have developed robots that interact with multiple parties (e.g. a robot teacher for children in a classroom [5], a museum guide robot [9]). In everyday settings, there are various cases where a human interacts with multiple parties. In HRI, many researchers aim to increase richer interactions between the robot and multi-party group [10, 14]. However, in order to develop the robot system for the multi-party case, the researchers have to both examine and compare whether humans respond in the same manner or not towards the robot system as they do to a human. So far, the research has been carried in regards to comparing human responses towards a robot mediator versus a human mediator [6]. Other researchers have also become interested in

The following authors are affiliated with Saitama University (Saitama City, Saitama, 338-8570 Japan). H. Fukuda, Y. Kobayashi and Y. Kuno are with the Graduate School of Science and Engineering; K. Yamazaki is with the Faculty of Liberal Arts. H. Fukuda is also JSPS Research Fellow.

A. Yamazaki is with the School of Media Science, Tokyo University of Technology (Hachioji City, Tokyo, 192-0982 Japan).

K. Ikeda is with the Division of International Affairs, Kansai University (Suita City, Osaka, 564-8680 Japan).

The first author's e-mail address is fukuda@cv.ics.saitama-u.ac.jp

comparing the responses of humans between multiple parties and the robot system [4, 16].

In this research, we compare how multi-party human participants interact with the robot system versus a human mediator in regard to their responses and interactions. These responses and interactions are facilitated through the answering of questions in a quiz scenario. We want to clarify whether the robot would really increase multi-party human interactions and what contributions from the robot make it so by comparing against a human mediator asking the same questions as the robot. Moreover, in these experiments, we also consider design choices for communication robots that enhance interaction among humans. This paper explores not only the issue of human vs. robots in a quiz scenario but also the design of robots for HRI.

II. RELATED WORK

There are some studies that spotlight the question in the field of conversation analysis. For instance, before asking a question, a speaker may ask the addressee a preparatory question (hereafter, pre-question) [12], which functions as a check of the recipient's epistemic status. Moreover, according to a series of papers examining question-response sequences in ten different languages [1], what is common among them is that the speaker typically gazes towards the addressee(s) when asking a question to a multi-party audience. Moreover, studies of interaction in various languages reveal similarities in the ways recipients display understanding or epistemic status. For instance, upon hearing new information, recipients may utter a change-of-state token such as 'Oh' [2].

Some robot systems based on these findings have been developed. For instance, in our previous work, we showed that the strategies which use a pre-question and a question effectively encourage the human responses among participants even when the questioner is a robot [14]. Some researchers have reported that the robot's gaze increases the human-human interaction among them [10, 13, 16]. We also explored human responses to a robot, and showed the person addressed the robot's gaze frequently and nodded to display a response when the robot said a specific keyword or finished speaking in experiments [15]. In this paper, we investigate the properties of robot mediated multi-party human interaction considering the change of epistemic status in a question, by comparing against a human mediator.

III. PRELIMINARY STUDIES

We have done two pilot studies based on the studies explained above. The first pilot study was done in order to investigate the changes of audience reactions by various types of questions. The second study was done in order to find out whether the same audience reactions would occur when the questioner is a human agent. In this paper, we will analyze our experiments based on observing responses and interactions among the audience. In our analysis, "response" means a reaction that is not shared among the audience, such as changes in facial expression and nodding. "Interaction" means an action that is shared among the audience, such as looking at each other and laughing at each other.

For the first pilot study, a total of 90 students from Saitama University (male: 60, female: 30) participated. They were sorted into groups of three, and a Robovie-R Ver.3 [17] (height: 108cm, hereafter abbreviated as Robovie (Fig. 1)) provided some guidance concerning the projected items on the screen in front of them and then would ask the participants a question. During its guidance and question session, Robovie also made use of pointing gestures and body language such as looking at the participants one by one. These kinds of motions were designed based on our ethnographical study using sociological methodologies and were developed to adopt various human behaviors in Robovie's movements. The robot's behaviors during its guidance talk in our experiment are very human-like, adopting various non-verbal modalities such as directing its gaze toward the participants and waving its hands, during its talk just as the human agents would [15]. Robovie asked six questions that were formatted with different structures, and at the end of each question, it appointed at a single person from among the three participants to answer the question. The participants were told prior to the experiment that they would need to answer the question when appointed by Robovie. In this pilot study, we embedded one question type in which participants' epistemic status would change (e.g., changing from "I know the answer" to "I do not know the answer"). The details of this type of question will be discussed further in the main experiment section of this paper. Observing the participants' response to this particular question, we have learned that they display such epistemic status changes through body language and verbalized words. They displayed their status and shared this among the group by interacting with each other through eye contact, smiling, or exchanging words briefly. We did not observe the same behavior for the other question types in the study.

For the second pilot study, 30 students from Saitama University (male:14, female: 16) participated. Similar to the first study, they were sorted into 10 groups of three and the same questioning style experiment took place, except that the questioner was a human agent instead of the robot. In the second pilot study, the human questioner asked four questions and we embedded two questions, which like in the first pilot study, had an epistemic change. The participants were also instructed to answer when they were appointed by the questioner. The human questioner was directed to behave similarly to Robovie, in other words s/he was told not to spontaneously respond to any reactions by the participants, and s/he also used body language in the same manner as Robovie. For the cases in which the human guides were involved in our research, we asked them to resemble the timing, major movements, and ways they carried themselves to how our robot was programmed to do so as a control against the human mediator's personal mannerisms. The human agents in our experiment are therefore taking the same track of actions programmed for the robot. By having the two different "agents" (human or robot) doing the same track of actions, we can argue that the participants would have different reactions due to the type of the agents presented to them. In this second study, we did not observe any cases where the participants interacted with each other through eye contact, brief conversations, or smiling, regardless of question types (0 cases out of the 10 groups, for the total of 40 tokens of questions, 20 tokens of those questions had an epistemic change.).

TABLE 1. The number of groups that responded and interacted with each other out of the 20 groups in our study.

	Human		Robovie	
	Response	Interaction	Response	Interaction
Q2a	3	0	13	3
Q3a	5	1	6	2
Q3b	2	0	7	5
Q3c	9	1	13	6

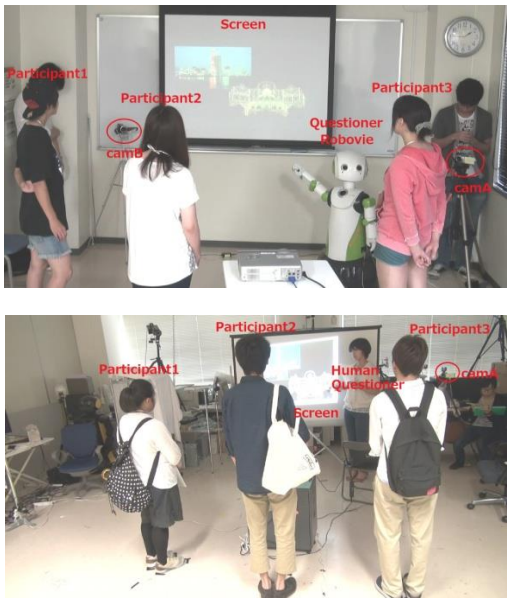

Figure 3. Experiments by a robot questioner (up), and a human questioner (down).

Based on the results here, we infer that a robot agent and questions with epistemic changes embedded in them, facilitates an environment where the participants would interact more among themselves. For the main experiment, we set up an experiment with two settings (human agent and Robovie as the questioner) with the same set of questions.

IV. MAIN EXPERIMENT: COMPARISON BETWEEN HUMAN-HUMAN INTERACTION AND HUMAN-ROBOT INTERACTION

120 students from Saitama University (male: 78, female: 42) participated in the main experiment. 40 groups of three were formed, and 20 groups faced the human agent as the questioner, and the other 20 groups had Robovie. The gender distribution consisted of male: 40, female: 20 for the human agent groups, and male: 38, female: 22 for the Robovie groups. Figure 3 shows examples of scene of experiments.

The standing positions of Robovie (Figure 1) and the human agent were kept the same. Their behaviors during the experiment were made as identical as possible. Since we used a recorded female voice for the robot's voice, three female agents played the questioner role and rotated for every three groups. The below are the questions used in the main experiment. Three different types of questions were designed. At the end of each question, the questioner would appoint a single participant, and the participants were randomly instructed beforehand to answer when they were appointed.

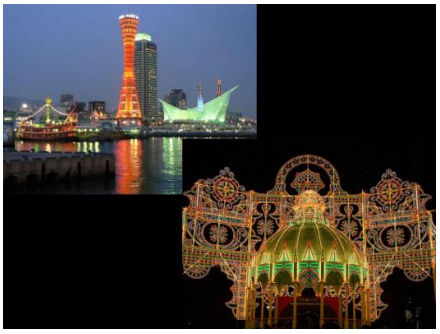

Figure 4. Image projected on screen for Q2.

Figure 5. Image projected on screen for Q3.

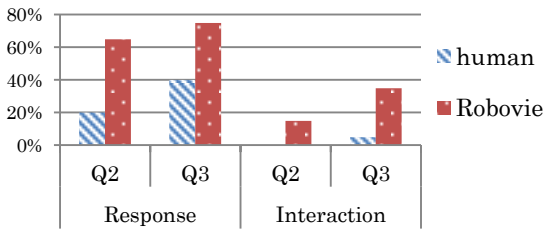

Figure 6. Comparison between human and Robovie.

Question 1 projects an image of a float for an American festival. The float's top roof is made of vegetable products of the particular region, and it is designed to take the form of the national flag of the United States. The question asks what the design is. Question 2 projects some images from Kobe city in Japan (Figure 4), and it asks the name of the prefecture where the city is located. These questions refer to the fact that the name of the city (Kobe) is better known than its prefecture. Question 3 asks who the former California governor is while projecting the picture of J.F. Kennedy (his niece was married to the governor, Arnold Schwarzenegger, Figure 5). In Japan, he is better known as an actor than a politician. Among these three questions, the last two have a change of epistemic state (the answer for Question 1 is detectable by observing the projected image carefully) and we have investigated the results. Specifically, the questions were the following:

Figure 7. human interaction at Q3b.

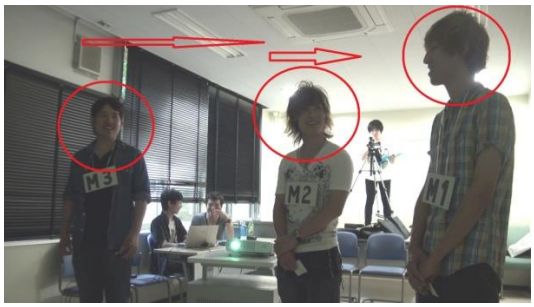

Figure 8. human interaction at Q3c.

Q2: This well-known port in Japan has some famous tourist sites such as Mt. Rokko and Arima hot springs. This city was heavily damaged in the Great Hanshin/Awaji Earthquake. Every year in December, the city holds a Luminaria memorial event to remember the dead and pray for the restoration of the city. What is the name of the prefecture that this city is located in? ←Q2a Please answer. (The questioner turns to face and point at one of the participants.)

Q3: The man on the left of this photo is a very famous president of the United States, isn't he? ←Q3a A person who is related to this man served as a governor of the state shown on this map. Do you know the name of the governor? ←Q3b He is also a famous Hollywood film star. ←Q3c Please answer. (The questioner turns to face and points one of the participants.)

Figure 6 indicates the ratio of (i) the participants' responses during the questioner's talk and (ii) the particiants' interactions among themselves during the experiment. In this experiment, we confirmed that Robovie generated more responses from the participants than the human questioner (Response: Q2. 20%vs.65% ($p=0.005<0.05$, Fisher's exact test), Q3. 40%vs. 75% ($p=0.027<0.05$), Interaction: Q2. 0%vs.15% ($p=0.115$), Q3. 5%vs. 35% ($p=0.022<0.05$)).

Table 1 shows the observations during questioning with epistemic change. For Q2, all cases of participants' interaction occurred at Q2a (above). There were cases when some interaction also occurred prior to position Q2a, but all participants did react at this point. For Q3, responses occurred at Q3a. At this point, the question checks with the participants if they knew Kennedy, thus we observed their nodding or verbalizing of their agreement. Similar to the findings from the previous study [15], the person addressed the questioner's eye gaze frequently and nodded to display a response. In the Robovie sets of the experiments, there were 2 examples (out of 20 cases) in which the participants indicated that they did not know the answer, and showed one example of interaction among participants. At the Q3b, the answer was relatively hard to generate, and at Q3c, a hint was provided. At this point, the participants' epistemic state changed from "not knowing" to "knowing" the answer. At Q3b, there were 2 cases with responses in the human questioner sets, and 7 cases in the Robovie sets. With interaction among the participants, there were no cases in the human questioner sets, and 5 cases in the Robovie case. Although there were no interactions among the participants, there were two cases in which the participants said, "I don't know." or simply laughed by themselves for the human sets. For the Robovie sets, there were 7 such cases. For Q3c, there were 9 cases with responses in the human sets and 13 cases in the Robovie sets. There was 1 case with interaction among participants in the human sets whereas there were 6 cases in the Robovie sets. The types of responses observed for the Robovie cases were either (i) eye contact, (ii) brief conversation, and (iii) laughing together. For the human case (1 case), we observed that the participants laughed together when they were not able to come up with an answer.

In this study, we learned that change of epistemic status did induce the participants' responses in both the human and Robovie sets. In this study, we learned that change of epistemic status did induce the participants' responses in both the human and Robovie sets. However, the types of responses differed in the two contexts. We observed nodding and verbal responses to show agreement in the case of the human questioner and Robovie seeking a confirmation from the participant. Indications to each other that they know or do not know the answer was rarely observed in the human sets, whereas more cases were found in the Robovie sets.

Let us illustrate two specific cases of interaction among the participants at this point. In Figure 7, we see that F2 turns to F1 looking puzzled when the questioner asked, "Do you know the name of the governor?" F1 turns the gaze to F2 in reply. In Figure 8, when the questioner asked, "Do you know the name of the governor?" M3 turned to M1 and M2 and asked, "Do you know the answer?" Later on when the questioner says, "He is also a famous Hollywood film star," M1 again indicates with laughter to the other participants that he knows the answer. As we see here, the participants apply various indications to share their own epistemic state among the group. Interaction occurs among the participants in order to do so.

V. DISCUSSION

From this main experiment, we confirmed that the participants do behave similarly both in the human and Robovie questioner sets when the questioner seeks a confirmation from them. As we reported in [13], this suggests to us that adopting human-like performance in the robot is an effective method. In addition to this, we also found that the participants engaged in interaction among each other more when the questioner was the robot. As shown in the results, the exchange among the participants took place during the questioner's guidance talk and also right after the question was given to them. This phenomenon can be seen as a kind of violation of the turn-taking system for conversational exchange [8] since the participants were neglecting the questioner as the current speaker.

The reactions and interactions among the human participants increased and were well facilitated, most likely because the human participants regarded the robot as not being a part of their social interactions (for example they did not engage in regular turn-exchange system with the robot in their interactions, as shown in the paper). Some reasons that the humans think so can be considered. One of the reasons is the difference of ability between the robot and human mediators. Since the robot is programmed to perform as a guide, it does not respond to any of the participants' actions. In the experiment, participants understand that the robot does not observe their actions. Therefore, it is likely that they think the robot is excluded from their interactions. The relationships between participants would be also factor. Although the participants we used in experiments were classmates, we did not consider whether the participants in the different groups were friends or not. In order to investigate this point, we need to conduct further experiments.

However, interaction among the human participants is in fact to be appreciated in many contexts such as museum or classroom settings, where active communication among those present is expected. A modern trend in the museum business now is to shift the way human guide speak so that the visitors' spontaneous communication is encouraged.

As the robot design becomes closer to that of human agents, it has been pointed out before that human participants reduce their social responses towards the robot [11]. There are also discussions regarding to what extent human agents should go out their way to smooth out the interaction with the robot. As studies on care robots for autistic patients point out, the design of the robot is better off not resembling humans too much [7]. In these studies, the researchers observed how the appearance of robots affect human attitudes. While, in this research, we studied how the design of robots including its actions affect multiple humans simultaneously and found that interactions among the multiple humans were more frequent in the case of the robot questioner. Based on this finding, we conclude that it is useful to employ robots for enhancing interaction between humans for education and exhibitions. On the other hand, if we could determine the cause for increase interaction, we could better design robots that enhance human interaction among each other. We need to consider various robot designs in conjunction with prescribed actions for this purpose. As our future work, in order to reveal what aspects of robots enhance human interaction, we plan to perform experiments in various scenarios (e.g. robot design and relationship between participants).

VI. CONCLUSION

This paper suggests a design for a robot, which enhances social interaction among the human participants because of its intervention. Based on the findings from this study and further exploration, we can suggest a robot system that detects the level of the human agents' mutual understanding and encourages their engagement further by naturally intervening into their interaction. There needs to be more experimental studies to investigate various designs of the robot, level of intervention, and potential types of social interactions that take place in such a situation. Our group is planning on future studies as a next step in these veins of development. We believe that findings from such studies will bring valuable knowledge to many HCI and HRI scholars of multi-party interaction.

ACKNOWLEDGMENT

This work was supported by JSPS KAKENHI Grant Number JP26240038, JP21300316, JP26590086, and JP23252001.

REFERENCES

[1] N. J. Enfield, T. Stivers, and S. C. Levinson, "Question-response sequences in conversation across ten languages," *An introduction. Journal of Pragmatics*, Vol. 42, 10, 2010, pp. 2615-2619.

[2] J. Heritage, "The epistemic engine: Sequence organization and territories of knowledge," *Research on Language & Social Interaction*, Vol. 45, 1, 2012, pp. 30-52.

[3] M. Ishikawa, "Young people's encounters with museum collections: Expanding the range of contexts for art appreciation," *International Journal of Education through Art*, 8(1), 2012, pp. 73-89.

[4] M. Johansson, G. Skantze, and J. Gustafson, "Comparison of human-human and human-robot turn-taking behavior in multiparty situated interaction," In *Proc. of the 2014 workshop on Understanding and Modeling Multiparty, Multimodal Interactions*, 2014, pp. 21-26.

[5] T. Kanda, T. Hirano, D. Eaton, and H. Ishiguro, "Interactive robots as social partners and peer tutors for children: a field trial," *Journal of Human-Computer Interaction*, Vol.19, 1, 2004, pp. 61-84.

[6] T. Kanda, T. Miyashita, T. Osada, Y. Haikawa, and H. Ishiguro, "Analysis of humanoid appearances in human-robot interaction," *IEEE Trans. on Robotics*, Vol. 24, 3, 2008, pp. 725–735.

[7] H. Kozima, C. Nakagawa, and Y. Yasuda, "Interactive Robots for Communication-Care: A Case-Study in Autism Therapy," In *Proc. of RO-MAN2005*, 2005, pp. 341-346.

[8] H. Sacks, E. A. Schegloff, and G. Jefferson, "A Simplest Systematics for the Organization of Turn-Taking for Conversation," *Language*, Vol. 50, 4, 1, 1974, pp. 696-735.

[9] M. Shiomi, T. Kanda, H. Ishiguro, and N. Hagita, "Interactive humanoid robots for a science museum," In *Proc. of HRI'06*, 2006, pp. 305-312.

[10] C. L. Sidner, C. Lee, L. P. Morency, and C. Forlines, "The effect of head-nod recognition in human-robot conversation," In *Proc. of HRI'06*, 2006, pp. 290-296.

[11] M. Strait, L. Vujovic, V. Floerke, M. Scheutz, and H. Urry, "Too Much Humanness for Human-Robot Interaction: Exposure to Highly Humanlike Robots Elicits Aversive Responding in Observers," In *Proc. of CHI '15*, ACM, 2015, pp. 3593-3602.

[12] A. K. Terasaki, "Pre-announcement sequences in conversation," *Conversation Analysis: Studies from the first generation*, Lerner, Gene H. (ed.), John Benjamins Publishing, Amsterdam, Netherlands, 2004, pp. 171-223.

[13] A. Yamazaki, K. Yamazaki, Y. Kuno, M. Burdelski, M. Kawashima, and H. Kuzuoka, "Precision timing in human-robot interaction: coordination of head movement and utterance," In *Proc. CHI'08*, 2008, pp. 131-140.

[14] A. Yamazaki, K. Yamazaki, T. Ohyama, Y. Kobayashi, and Y. Kuno, "A techno-sociological solution for designing a museum guide robot: regarding choosing an appropriate visitor," In *Proc. of HRI'12*, 2012, pp. 309-316.

[15] A.Yamazaki, K. Yamazaki, K. Ikeda, M. Burdelski, M. Fukushima, T. Suzuki, M. Kurihara, Y. Kuno, and Y. Kobayashi, "Interactions between a quiz robot and multiple participants: Focusing on speech, gaze and bodily conduct in Japanese and English speakers," *Interaction Studies*, 14, 3, 2013, pp. 366-389.

[16] Q. Xu, L. Li, and G. Wang, "Designing engagement-aware agents for multiparty conversations," In *Proc. of CHI '13*, 2013, pp. 2233-2242.

[17] http://www.vstone.co.jp/english/index.html