Title	Multiple-Robot Mediated Discussion System to support group discussion *
Author(s)	Ikari, Shogo; Yoshikawa, Yuichiro; Ishiguro, Hiroshi
Citation	29th IEEE International Conference on Robot and Human Interactive Communication, RO-MAN 2020. 2020, p. 495-502
Version Type	AM
URL	https://hdl.handle.net/11094/78295
rights	© 2020 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.
Note	

Multiple-Robot Mediated Discussion System to support group discussion*

Shogo Ikari, Yuichiro Yoshikawa and Hiroshi Ishiguro

Abstract— Deep discussions on topics without definite answers are important for society, but they are also challenging to facilitate. Recently, advances in the technology of using robots to facilitate discussions have been made. In this study, we developed a multiple-robot mediated discussion system (m-RMDS) to support discussions by having multiple robots assert their own points and lead a dialogue in a group of human participants. The robots involved the participants in a discussion through asking them for advice. We implemented the m-RMDS in discussions on difficult topics with no clear answers. A within-subject experiment with 16 groups (N=64) was conducted to evaluate the contribution of the m-RMDS. The participants completed a questionnaire about their discussion skills and their self-confidence. Then, they participated in two discussions, one facilitated by the m-RMDS and one that was unfacilitated. They evaluated and compared both experiences across multiple aspects. The participants with low confidence in conducting a discussion evaluated the discussion with m-RMDS as easier to move forward than the discussion without m-RMDS. Furthermore, they reported that they heard more of others' frank opinions during the facilitated discussion than during the unfacilitated one. In addition, regardless of their confidence level, the participants tended to respond that they would like to use the system again. We also review necessary improvements to the system and suggest future applications.

I. INTRODUCTION

It has been argued that our society has a great need for deeper discussions among people about problems of humanity and ethics that have no clear answers [1]. Paris Declaration for Philosophy by UNESCO [2] posits that free open-ended discussions with no final answers on universal problems of human life and existence are essential for the training for citizenship.

However, it is difficult for people to get involved in discussions on such complex problems. This is due to the level of uncertainty about contributing to a discussion that has much freedom in its process and in the way conclusions are drawn. Therefore, participants are more likely to become silent, fearing the risk of being alienated [3] for making irrelevant statements. At the same time, such discussion creates possibilities for dominance [4], a situation characterized by eloquent speakers talking excessively.

It is known that an intervention of a skilled facilitator can reduce the inequality in the amount of speaking time by settling the arguments and correcting the direction of the

discussion [5]. However, skilled facilitators are not available everywhere, and there is a cost to training them. Even with a skilled facilitator, some people may be concerned about the facilitator's assessment and may be reluctant to speak freely in the presence of a facilitator. One reason is that people with low self-confidence tend to be concerned about the evaluations of others [6]. Furthermore, there is a need for motivational support for those who are not confident in their ability to conduct an argument because they may be less motivated to assert their own opinions as self-confidence and motivation are correlated in various situations [7].

For the reasons stated above, in our study, we employ robots to facilitate a discussion. In the domain of discussion facilitation, some studies have employed robots as facilitators [8] [9]. Since instructions issued by a robot tend to place less pressure on people than instructions given by a human [10], people may be able to speak more freely in the presence of robot facilitators than in the presence of human facilitators.

However, it is still difficult technically to perform speech recognition when people are talking rapidly and in disconnected ways [11]; this can occur especially in group situations, where multiple persons participate in the conversation [12]. In addition, if participants do not have clear opinions on the discussion topic, the facilitator elicits their opinions by providing advice, such as raising related issues or providing examples. Therefore, it is complicated to develop a robot facilitator that is capable of performing autonomous group facilitation. Moreover, it is not easy for the current dialog system to process conversations because of the limited capability of speech recognition to understand the flow of arguments and generate effective statements to guide the discussion.

Given these difficulties, Arimoto et al. proposed a method to facilitate natural conversation without speech recognition by involving a human in a scenario-based dialogue between multiple robots [13]. In this study, we extend this idea by involving humans in a discussion where robots exchange opinions. To provide motivational support for participants with low self-confidence, we give human participants a position of advisors to the robots, which assert their opinions.

Specifically, we develop a structured method to automatically generate a scenario in which robots exchange opinions with each other based on provided sentences, including counterarguments and responses on a given topic. Also, we design an interaction to collect statements to be uttered by robots; human participants play a role of advisors for the robots in these interactions, which are conducted directly before the robot-mediated discussion with other participants. Based on

this, dialogue among robots might be more immersive for human participants, because i) robots' statements contain advice given by the participants, and ii) through giving advice, social relationships are built between the human participants and robots. By implementing these mechanisms, we will be able to build a facilitation system that makes it easier to involve people with low self-confidence, characterized by lack of opinion and lack of motivation, in discussion situations.

In the rest of this paper, we introduce a prototype of a multiple-robot mediated discussion system (m-RMDS), which is implemented based on the ideas described above. Then, we conduct laboratory experiments to examine the effectiveness of the proposed system.

II. RELATED WORK

A. Conversational Robots to Ease Barriers in Verbal Communication

There have been various studies on robots that support human communication. Telecommunication using a robot as an avatar, which is expected to induce more self-disclosure, was applied between the elderly and their families [14], between children with autism spectrum disorder and their caregivers [15], and among inter-generational community [16]. Shimaya et al. showed that a robot teleoperated by students in a classroom setting encouraged the students to talk to their teacher [17]. Robots have been used to mediate human social interaction by participating in it as a third agent and have been applied in various communication settings between two persons, for example, between elementary school students [18] and between adults [19]. It has been shown that a bystander android robot that nodded its head relaxed the communication between a doctor and a patient in an examination room [19]. Birmingham et al. showed that a robot can improve trust within groups of university students by asking one member questions in front of the group and encouraging the member to disclose [20]. Traeger et al. showed that groups with a robot making vulnerable statements have more conversation and the members speak more equally [21].

Although the robots in those previous work contributed to conversation facilitation, which is an important factor of a discussion, their role was not to promote discussions in groups.

B. Robot-Mediated Group Discussion

Group discussion situations have been studied in some research, where robots show non-verbal behaviors to enhance active listening of the participants [22] or designate the next speaker [9]. Shamekhi et al. developed a humanoid robot, whose face was drawn on a monitor display mounted on a robot trunk, that provided interventions required to facilitate a discussion, regarding a specific decision-making task. It suggested points to be discussed in a predefined order and provided time management, balancing the amount of participants' speech and encouraging them to participate in the discussion [8]. All of these studies introduced a single

robot to the group, and the robots did not participate in the discussion with their opinions or points of view.

C. Application of Multiple Conversational Agents

Researchers have been interested in using multiple robots or agents to effectively provide information to humans. Such studies include using multiple agents to capture people's attention at a train station [23], provide a product recommendation [24], and deliver useful information in the cockpit of a car [25]. These studies aimed to efficiently provide information by showing conversation between artificial agents and were not designed to engage people actively in a conversation.

Furthermore, some studies have shown that a conversation system with multiple robots can prevent the disruption of dialogue by having the robots continue a conversation to maintain the flow [13] [26]. These studies were focused on continuing the dialogue in a natural way using robots. Kumazaki et al. have also been interested in changing the awareness of communication using multiple robots. Specifically, they reported that children with autism spectrum disorders learned to understand the importance of nonverbal communication through letting them create a program for multiple humanoid robots to talk to each other [27]. Leite et al. employ two robots in interactive storytelling for elementary school students [28]. However, none of these studies employed multiple robots in a situation where multiple persons argue with each other.

III. MULTIPLE-ROBOT MEDIATED DISCUSSION SYSTEM

The prototype m-RMDS is implemented in groups of four people. In this system, four robots join in the discussion, each as a peer of a human participant. They are given IDs of A, B, C, and D. In the discussion using m-RMDS, there are two types of interaction patterns between the robots and humans. The first type is a participant individually advising their peer robot, and the second one is a discussion led by scenario-based dialogue among multiple robots.

A. Advising the Peer Robot

In the advice-based interaction, participants give advice to the robots by typing in advice interface (Fig. 1). Each participant is asked by their peer robot for advice on a specific topic using sentences presented in Table I.
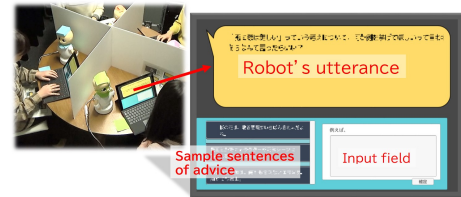

Fig. 1. Advice Interface

When the participants advise the robots, each robot expresses its predetermined opinion on the discussion theme.

Type of advice	What the robots say
example(EX)	"When I argue that XX[1], and they ask me to give an example because they don't know what I mean, what do you think I should say?"
ground of argument(GA)	"What do you say when someone asks you to give a reason why you argue XX? Even if it's not what you really think, I'd like your advice."
counter-argument(CA)	"I want to provide a counterargument, so please tell me how you would counter the opinion XX."
reassertion(RA)	"What do you think I should say in response to that counterargument? I'd like your advice, even if it's not what you really think."

[1]XX represents the predetermined opinion of the robot on a given discussion theme.

For example, if "death" is the theme, it may have an opinion such as "the moment of dying is beautiful" or "the moment of dying is miserable." Then, it asks for human help to successfully assert that opinion. Specifically, it asks an example of a situation that supports the opinion, grounds of argument of the opinion, a counterargument to the opinion, and reassertion of the opinion. The robot collects the sentences provided by the participant; these sentences are used to compose the dialogue among the robots in the group discussion on the given theme.

When the robots ask for an example or a ground of argument, sample sentences are displayed on the tablet in the advice interface. The robots' opinions and sample sentences were prepared in advance by our research collaborators. All participants give advice to their respective peer robots, as shown in Fig. 1. Advice given by the participants to the robots is recorded in the database and used as material for generating the dialogue script for conversation among the robots described below.

B. Discussion Led by Scenario-based Dialogue of Multiple Robots

1) Automatic Scenario Editor: The templates of the scenarios for the dialogue among four robots are shown in Table II. During the dialogue, the opinions of robots, on which participants have given advice, are discussed one by one.

The four robots are divided into two groups: Robot A and Robot B argue a certain opinion (that is represented as XX in Table II) and Robot C and Robot D an opposing one (that is represented as YY in Table II). In the experiment, the names of robots A, B, C, and D were set to Blue, Green, Pink, and Orange, respectively.

Consider a case where Participant A, who is an advisor to Robot A, has typed "I think it's ugly when you're dying because people can't do anything before they die" as a counterargument and "The last glow of dying is beautiful. For example, the cherry blossoms in the fall" as a reassertion. In such case, Robot C will refute Robot A by saying "I think it's ugly when you're dying because people can't do anything before they die", and after that Robot A will reassert that "The last glow of dying is beautiful. For example, the cherry

blossoms in the fall." Here, since the counterargument and reassertion have been typed by the same person, the dialogue between the robots should be coherent.

Once the robot presents all the arguments given in the advice-giving stage, it will ask for additional explanation from the participant who advised it in Counter argument and reassertion Phase. This part will be repeated for each robot. At this point, a discussion interface placed at the center of the m-RMDS (Fig. 2) shows the last two sentences uttered by the robots and a button for the participant to declare the end of their speech. When the button is pressed by the participant, the robots will resume their dialogue with each other.
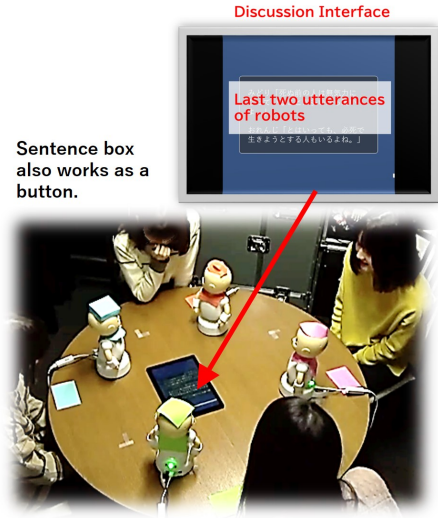

Fig. 2. Discussion Interface

2) Robot Behaviors: m-RMDS employs a tabletop communication robot, CommU[1]. CommU can open and close its mouth while emitting sounds through a speaker embedded in its chest; this gives the impression that it is giving a speech. It has 14 degrees of freedom, allowing it to move every part of its body. In m-RMDS, the robot produces head and hand gestures when it speaks. When a robot is speaking, the other robots look at it. In addition, when a robot asks a participant for additional explanation during the discussion, it looks at the participant.

3) Implementation Environments: The advice interface and discussion interface are running with Unity[2]. Transmission Control Protocol /Internet Protocol (TCP/IP) is used for advice interface, discussion interface, and communication of robots with each other. In the *advice-giving* stage, the advice interface running on a tablet computer (Microsoft Surface) sends commands to each corresponding robot through TCP/IP. In the *discussion* stage, the discussion interface running on the tablet computer sends commands to all robots.

[1]https://www.vstone.co.jp/products/commu/index.html
[2]https://unity.com/

TABLE II

SCENARIO TEMPLATE

Phase	Scenario
Raise argument	Robot C: "Today's theme is 'death'. Let's talk." Robot A: "Can I give my opinion? I think XX[1]." Robot D: "I don't agree." Robot C: "I don't agree, either. I rather think YY[1]." Robot B: "I am for XX. "
Share examples and grounds of arguments	Robot C: "I don't understand at all. So, give me an example of such a case." Robot A: "For example, EX-A[2]." Robot B: "Yes. In addition, GA-A." Robot C: "But EX-C." Robot D: "Robot C is right. And GA-D"
Counter argument and reassertion[3]	Robot C: "I think what Robot D says is more convincing. In addtion, CA-A." Robot A: "Oh, no way. RA-A." Robot D: "I am unclear on what you are saying. [Name_pA][4], what was Robot A trying to convey?" Robot A: "Could you help me, [Name_pA]?" # All robots look at Participant A. Participant A is asked to give an explanatory statement. Robot D: "OK I understood. But I cannot be on the same page as Robot A."
Discussion between human participants	Robot C: "To sum up the arguments so far, the point is that we're at odds." Robot A: "I feel like it doesn't make any sense to keep talking like this." Robot B: "Let's hear what the humans think." Robot D: "Well then, humans, please discuss the theme until we stop you." # The robots become silent.
Conclusion	# When there are two minutes remaining, one of the robots gives a statement to the floor. Robot B: "Now that we have two minutes left, "I hope you guys can come to a conclusion." # The robot becomes silent again.
End of the discussion	# After two minutes, the robots start to talk again. Robot A: "Thank you, everyone." Robot B: "Thanks to humans we've learned a lot." Robot D: "Let's end today's discussion for now. We'll have to think about it again." # The end.

[1]XX and YY represent the predetermined opinions of the robot on a given discussion theme which contradict with each other. [2]The advice given by the participant to the robot is represented in the abbreviations shown in Table I and the advisor's ID, such as EX-A for advice on example given by Participant A and CA-C for advice on counterargument given by Participant C. [3]This part will be repeated 4 times: Robot C refuting Robot A, Robot B refuting Robot C, Robot D refuting Robot B, and Robot A refuting Robot D. [4][Name_pA] is replaced by the name of Participant A the first time, and later in the discussion by the other participants.

IV. EXPERIMENTAL EVALUATION

A. Participants and Conditions

Sixteen groups of four undergraduate and graduate students (N=64; age:21.5 ± 2.1; 32 female and 32 male), recruited through an agency, participated in the experiment. The participants in each group met for the first time during the experiment, and the ratio of males to females in the individual groups varied.

The experiment was a within-subjects design with two conditions: a discussion with m-RMDS (m-RMDS condition) and a discussion without m-RMDS (control condition). Each group discussed the two themes of "death" and "empathy".

TABLE III

POST-DISCUSSION QUESTIONNAIRE

	Items
Overall satisfaction	Q1 I enjoyed the discussion. Q2 I am satisfied with the discussion. Q3 The discussion deepened.
Sense of involvement	Q4 I felt a strong sense of involvement in the discussion.
Individual performance	Q5 I contributed to the process of drawing conclusions in the group. Q6 I contributed to the integration of different opinions.
Group performance	Q7 The conclusions reached by the group reflected many of the different opinions expressed during the discussion.
Self-disclosure	Q8 I shared my true feelings with the other group members.
Understanding of others	Q9 I heard frank opinions from the other group members.
Group closeness	Q10 The Inclusion of Other in the Self Scale[1] [31]

[1]To measure intragroup closeness, we used the IOS scale [33], a single-item pictorial measure how close the respondent feels with another person or group.

The assignments of themes and experimental conditions were counterbalanced.

In the m-RMDS condition, participants were randomly assigned an ID and became advisors to the robot with the same ID. Robot A and Robot B asserted that "a moment of dying is beautiful" and Robot C and Robot D that "a moment of dying is miserable" on the theme of "death". On the theme of "empathy" Robot A and Robot B asserted that "empathy makes people grow" and Robot C and Robot D that "empathy makes people stupid." In the control condition, each participant was given a claim corresponding to their ID in advance, and was asked to express their thoughts regarding the assigned claim in the early stage of discussion.

B. Questionnaire

1) Preliminary questionnaire: In order to measure participants' self-evaluation of their discussion skills, we used the sub-item of the *self-confidence in thinking skills* factor in the Japanese version of the Critical Thinking Disposition Inventory [29]. These items were selected and translated into Japanese from the "self-confidence" subscale of the California Critical Thinking Disposition Inventory [30]. The pre-questionnaire included other items measuring *inquisitiveness*, *objectivity*, and *carefulness*; for these, items with loadings of 0.6 or more in Hirayama et al. [29] were chosen.

2) Post-discussion questionnaire: To measure the subjective evaluation of the discussion from various perspectives, participants completed a post-discussion questionnaire. The questionnaire comprises items shown in Table III, and they were answered on a seven-point Likert scale.

3) Post-experiment questionnaire: The post-experiment questionnaire (Table IV) included items comparing the m-RMDS and control conditions from various perspectives and items asking the participant's preference of a discussion with or without m-RMDS in various situations. The questionnaire

Items [1]	
Qc.1	Which discussion did you enjoy more?
Qc.2	Which discussion moved forward more easily?
Qc.3	In which argument did you say more of what you wanted?
Qc.4	In which argument was it easier to listen to others' opinions?
Qc.5	If you had an opportunity to have another discussion with the same members, which condition would you like to participate in?
Situ.1	Which condition would you prefer in discussions in your class?
Situ.2	—- in meetings of your club?
Situ.3	—- with your friends?
Situ.4	—- with people you do not like?
Situ.5	—- with your family?
Situ.6	—- with people you meet for the first time?
Qa	Would you like to use the discussion system again?(7-Likert)
Qf.1	What are the flaws of the discussion system?(free description)
Qf.2	In what situations would you want to use the discussion system? (free description)

[1]Qc. 1 to 5, and Situ. 1 to 6 are forced choice items. Qa, Qf. 1, and Qf.2 refer only to m-RMDS condition.

included questions about the participant's feelings regarding the discussion led by m-RMDS, and whether they would like to have another discussion with m-RMDS and in what situations.

C. Procedure

Each participant took one minute to introduce themselves to the other members of the group, and then all members filled out the *preliminary questionnaire.* This was followed by a first round of discussion. Then they filled out the *post-discussion questionnaire* and participated in a second discussion. Afterwards, they responded to the same *post-discussion questionnaire* as after the first discussion. Finally, they filled out the *post-experiment questionnaire.*

The discussions proceeded as follows. In m-RMDS condition, the participants had 10 minutes to give advice to the robot before the discussion. At this time, the participants were not able to communicate with each other. Afterwards, they participated in a 20-minute discussion led by the robots. After the exchange of opinions by the robots, the participants took the lead in the discussion. Two minutes before the end, the robot asked the group to summarize their opinions. As instructed in advance, after the discussion time elapsed, each participant wrote down the ideas they had arrived at through the discussion. This process was included to provide motivation for the discussion.

In the control condition, before the discussion, participants spent 10 minutes filling out a worksheet consisting of the same items that the robot would seek advice on in the case of m-RMDS condition. During this time, the participants were not able to communicate with each other. They could refer to a handout with the sample sentences that were displayed as choices in the m-RMDS advice interface. During the discussion, participants were able to refer to the worksheet they had completed but not to the handout with sample opinions. In the control condition, the discussion time was also 20 minutes. Two minutes before the end, the experimenter

told the group to summarize their opinions. As instructed in advance, the participants individually wrote down the ideas they had arrived at through the discussion.

V. RESULTS

A. Critical Thinking Disposition of Participants

The mean and standard deviation (M(SD)) values of the participants on the Critical Thinking Disposition subscales were 2.95(.91) in *self-confidence in thinking skills*, 4.22(.51) in *inquisitiveness*, 4.69(.51) in *objectivity*, and 3.42(.81) in *carefulness.*

In the following analysis, participants with *self-confidence in thinking skills* scores higher than 3.0 are defined as high-confidence participants, and ones with *self-confidence in thinking skills* scores lower than 3.0 are defined as low-confidence participants. Of the 64 participants, 29 were high-confidence and 29 were low-confidence participants. Six participants whose scores were 3.0 were excluded from both groups.

Since the purpose of this experiment is to make within-subject comparisons, the main or simple main effects of confidence will not be discussed.

B. Subjective Evaluation of Discussion

A mixed-factor analysis of variance for confidence(CF) and experimental conditions(EC) was performed on the questionnaire on satisfaction with the discussion Fig. 3. Multiple comparisons were performed using the Holm method. Four participants with missing responses were excluded from the analysis.
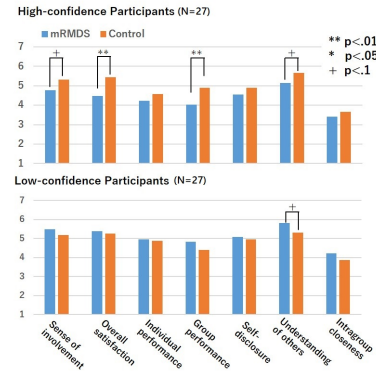

Fig. 3. Subjective evaluation by high- and low-confidence participants

For *overall satisfaction*(OS), there was a significant trend of interaction($F(1,53) = 3.98, p < .1$). A significant trend of simple main effect for EC was found in the high confidence participants(high-CF)($F(1,53) = 3.25, p < .1$) but not in low confidence participants(low-CF)($F(1,53) = 1.04, p > .1$). For *sense of involvement*(SI), there was a significant interaction($F(1,53) = 5.85. p < .05$). A significant simple main effect for EC was found in high-CF($F(1,53) = 8.87, p < .01$) but not in low-CF($F(1,53) = 0.20, p > .1$). For *individual performance*(IP), there was no significant interaction($F(1,53) = 1.28, p > .1$) and a main effect of EC

was not significant($F(1,53) = 0.56, p > .1$). For *group performance*(GP), there was a significant interaction($F(1,53) = 6.02, p < .05$). A significant trend of simple main effect for EC was found in high-CF($F(1,53) = 4.11, p < .1$) but not in low-CF($F(1,53) = 1.35, p > .1$). For *self-disclosure*(SD), there was no significant interaction($F(1,53) = 1.08, p > .1$). Also, no main effect was significant($F(1,53) = 0.70, p > .1$ for CF; $F(1,53) = 0.29, p > .1$ for EC). For *understanding of others*(UO), there was a significant interaction($F(1,53) = 7.58, p < .01$). A significant trend of simple main effect for EC was found in high-CF($F(1,53) = 3.93, p < .1$) and in low-CF($F(1,53) = 3.65, p < .1$). For *intragroup closeness*(IC), there was a significant interaction($F(1,53) = 4.09, p < .05$). However, no simple main effect for EC was significant in high-CF($F(1,53) = 1.45, p > .1$) no low-CF($F(1,53) = 2.75, p > .1$).

C. Intention to Use

One-sample t test found that both the high-CF and low-CF groups responded to the question "Would you like to use the robot discussion system again?" with ratings significantly higher than the chance level (high-CF: $t(28) = 2.83, p < .01$ low-CF: $t(28) = 3.36, p < .01$). Mean(SD) values of their scores were $4.68(1.27)$ and $5.03(1.63)$, respectively.

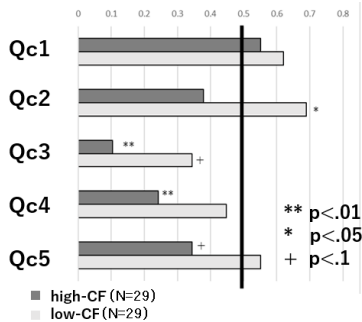### D. Participant evaluation of m-RMDS over various aspects

Fig. 4. Ratio of participants who rated the discussion in the m-RMDS condition as better than the discussion in control condition over various aspects

Chi-square test found following significance (Fig. 4). There were significantly less participants who chose the m-RMDS condition for Qc.3($3/29, \chi^2(1) = 18.2, p < .01$) and Qc.4($7/29, \chi^2(1) = 7.75, p < .01$) than those who chose the control condition. In addition, there was a significant trend that less high-CF participants chose m-RMDS condition for Qc.5($10/29, \chi^2(1) = 2.79, p < .1$) than the control condition. No significant trend in Qc.1($16/29, \chi^2(1) = 0.31, p > .1$) and Qc.2($11/29, \chi^2(1) = 1.68, p > .1$) was observed among the high-CF participants. In the low-CF group, there were significantly more participants who chose m-RMDS condition for Qc.2($20/29, \chi^2(1) = 4.17, p < .05$). Also, there was a significant trend that less low-CF participants chose m-RMDS condition for Qc.3($10/29, \chi^2(1) = 2.79, p < .1$). No significant trend in Qc.1($18/29, \chi^2(1) = 1.68, p > .1$), Qc.4($13/29, \chi^2(1) = 0.31, p > .1$), and

Qc.5($16/29, \chi^2(1) = 0.31, p > .1$) was observed among the low-CF participants.

E. Usefulness of m-RMDS in various situations

There were significantly more people who chose m-RMDS condition for Situ.3 ($56/64, \chi^2(1) = 36.0, p < .01$) and Situ.5 ($46/64, \chi^2(1) = 12.3, p < .01$), and less people who chose m-RMDS condition for Situ.2 ($6/64, \chi^2(1) = 42.3, p < .01$) and Situ.4 ($8/64, \chi^2(1) = 36.0, p < .01$), and no significance for Situ.1 ($29/64, \chi^2(1) = 0.56, p > .1$).

In addition, according to the answers to the free-description question that asked when they would like to use m-RMDS, the situations they described can be categorized as follows: "when the atmosphere of the discussion is tense" (11 participants), "when the discussion topic is abstract or difficult to argue" (9 participants), "when many people are participating and it is difficult to speak" (5 participants), and "when disagreements tend to arise" (5 participants).

VI. DISCUSSION

A. Participants with Low Self-Confidence

Subjective evaluation shows that participants with low confidence felt that other group members expressed frank opinions more in the m-RMDS condition than in the control condition. This implies that m-RMDS enabled them to understand other participants' opinions more deeply. Furthermore, through the post-experiment questionnaire, it was found that the participants with low confidence felt that it was easier to move the discussion forward with the m-RMDS than without it. These results indicate that the benefits of m-RMDS were likely provided to the low-confidence participants as originally envisioned.

In an unfacilitated discussion, participants find their pace through alternately taking turns to speak. People with low confidence tend to not actively engage in the turn-taking [32]. This makes it difficult for them to find their pace in the discussion and to feel that they can move the discussion forward easily. On the other hand, in m-RMDS condition, the robots designated the next speaker, and asked the speaker to add explanation to the preceding dialogue of the robots. Therefore, for the participants to contribute, they did not need to determine what topic to be focused on and when to speak. In this way, m-RMDS helped the participants with low confidence reach their own pace in the discussion. This process may have made the low-confidence participants feel that the discussion was easier to move forward.

Moreover, people tend to maintain consistency in speech and behavior, and when they fail to do so, they experience dissonance [33]. However, those with low self-confidence are more likely to lose their consistency by changing their opinions due to the influence of others [34]. Therefore, in normal discussions, low-confidence participants may face dissonance and experience stress as a result. On the other hand, with m-RMDS, the discussion proceeded based on the opinions expressed by the robots, and therefore participants do not have to feel responsible for them. Then, low-confidence participants could avoid facing dissonance more

in m-RMDS condition than in the control condition. This may have led to them to conclude that it is easier to move the discussion forward with m-RMDS.

Furthermore, the low-confidence participants rated m-RMDS condition more positively in terms of being able to hear frank opinions of others, and this suggests that they are likely to regard responses to robots' questions as the frank opinions of other participants. Birmingham et al. have shown that when each group member responds to a question from a robot, trust within the group increases [20]. Utterances prompted by a robot in front of the group might make the speaker seem more honest and as a result recognized as trustworthy by the other participants. Therefore, the result of the increased experience of frankness in the current experiment encourages us to consider extending the research of m-RMDS to support low-confidence participants to establish trusting relationships.

Compared with the positive evaluations discussed above, the low-confidence participants tended to be less satisfied with their opportunities to say what they wished to say in the m-RMDS condition. This may be because m-RMDS did not allow a participant to speak while the robot was speaking. In fact, one of the comments indicating the inadequacies of the system was ″I wish I could talk a little more.″ In this regard, if the m-RMDS is modified to allow people to interrupt and express their opinions during the robots' speeches, it is possible to reduce the participants' experience that opportunities for speech are restricted.

B. Participants with High Self-Confidence

For the high-confidence participants, discussions with m-RMDS were rated more negatively than those without m-RMDS. However, answers to the question ″Would you like to use the system again?″ showed a significantly higher rate of affirmative answers than chance level, and about half of them found the m-RMDS condition more enjoyable than the control condition. Therefore, it is worth determining the aspects that disappointed the high-confidence participants and that could be improved. In discussion with m-RMDS, the participants could speak only when they were designated to do so. This could be a source of stress in discussions with m-RMDS, as opportunities to speak can be experienced as limited. Therefore, modifying the system to allow a person to interject their opinion while the robots are speaking should be important in future work.

C. Suggested Improvements

Both high- and low- confidence participants pointed out some inadequacies of the system: i) the robots had poor facial expressions; ii) their speech was sometimes difficult to understand because of the lack of inflection; and iii) the speed of their speech was too slow. If the dialogue between robots becomes more expressive and its tempo becomes more natural by improving the points above, it is expected that the participants' feel more comfortable during robots' speech.

D. Future Applications

After making the improvements described above, the system developed in this study is expected to be useful in supporting face-to-face discussions. In particular, many of the subjects reported that they would like to use the system when having a discussion with a person they had met for the first time or a person they disliked. This suggests that the system could be useful for discussions among people who have a shallow relationship with each other. The analysis of the subjects' free descriptions further showed that many people would like to use it in discussions dealing with difficult issues or in situations, in which they were hesitant to speak. Therefore, it could be a useful device for conducting discussion-based workshops, such as philosophical dialogues [35], which require in-depth discussion of unanswered questions in groups, whose members meet for the first time. Moreover, the system could also be utilized in cross-generational discussion, which is prone to difficulties in moving forward [36].

E. Limitations

In the current experiment, we did not control for the number of high- and low-confidence members within a group. Therefore, it is not clear to what extent the reported effects of m-RMDS depend on the ratio of high- and low-confidence members in a group.

In addition, the participants were recruited from a population of college students, potentially with a higher intellectual ability compared with other populations. Therefore, it is necessary to test whether consistent results would be obtained with participants sampled from more diverse populations.

Since this was a laboratory experiment and the subjects were motivated to participate in the discussion, it is not clear whether m-RMDS would be able to engage people who are not externally motivated in a discussion. Field experiments should be conducted to determine this.

VII. CONCLUSION

In this study, we proposed a novel framework of discussion facilitation that employs multiple robots to lead a discussion along scenario-based dialogue among them. We developed the system to support discussion especially for facilitating discussions on complicated topics without clear answers. We investigated how the system affected the participants' evaluation of the discussion through a controlled laboratory experiment. We found that the multiple-robot system was useful for people with low confidence in their discussion ability. For participants with high confidence, satisfaction with the discussion was reduced when the robots were leading it; however, they reported that they would like to use it again.

The findings obtained through this study can contribute to the development of technologies that will enable robots to facilitate difficult discussions among humans more effectively.

REFERENCES

[1] D. Bohm. (2004) *On Dialogue.* Psychology Press.

[2] UNESCO. (2005). Report by the Director-General on an intersectoral strategy on philosophy. *UNESCO. Executive Board, 171st, 2005.*

[3] E. Morrison & F. Milliken. (2000). Organizational Silence: A Barrier to Change and Development in a Pluralistic World. *Academy of Management Review.* 25, pp. 706-725.

[4] M. R. Callaway, R. G. Marriott, & J. K. Esser. (1985). Effects of dominance on group decision making: Toward a stress-reduction explanation of groupthink. *Journal of Personality and Social Psychology.* 49(4), pp. 949–952.

[5] F. Westley, F. & J. A. Waters. (1988). Group Facilitation Skills for Managers. *Management Education and Development.* 19(2), pp. 134–143.

[6] M. Church & C. Truss, & W. Velicer. (1980). Structure of the Janis-Field Feelings of Inadequacy Scale. *Perceptual and motor skills.* 50, pp. 935-9.

[7] R. Bénabou, & J. Tirole. (2002). Self-Confidence and Personal Motivation, *The Quarterly Journal of Economics, Volume.* 117(3), pp. 871–915,

[8] A. Shamekhi, & T. Bickmore. (2019). A Multimodal Robot-Driven Meeting Facilitation System for Group Decision-Making Sessions. *21st ACM International Conference on Multimodal Interaction (ICMI 2019)*

[9] Y. Matsuyama, I. Akiba, S. Fujie, & T. Kobayashi. (2015). Four-participant group conversation: A facilitation robot controlling engagement density as the fourth participant. *Computer Speech & Language.* 33(1), pp. 1-24.

[10] K. Suzuki, & M. Kanoh. (2017). Investigating Effectiveness of an Expression Education Support Robot That Nods and Gives Hints. *Journal of Advanced Computational Intelligence and Intelligent Informatics.* 21(3), pp. 483-495

[11] A. Lee, & T. Kawahara. (2009). Recent Development of Open-Source Speech Recognition Engine Julius. *em Proceedings of the 2009 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference.*

[12] M. Shiomi, T. Kanda, I. Howley, K. Hayashi, & N. Hagita. (2015). Can a Social Robot Stimulate Science Curiosity in Classrooms?. *Int. J. Soc. Robot.* 7(5), pp. 641-652

[13] T. Arimoto, Y. Yoshikawa, & H. Ishiguro. (2018). Multiple-Robot Conversational Patterns for Concealing Incoherent Responses. *Int. J. Soc. Robot* 10, 583–593.

[14] Y. Noguchi, H. Kamide, & F. Tanaka. (2018). Effects on the Self-disclosure of Elderly People by Using a Robot Which Intermediates Remote Communication. *Proceedings of the 27th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN 2018)*, pp. 612-617.

[15] J. Shimaya, Y. Yoshikawa, H. Kumazaki, Y. Matsumoto, M. Masutomo, & H. Ishiguro. (2019). Communication support via a tele-operated robot for easier talking: Case/laboratory study of individuals with/without autism spectrum disorder, *Int. J. Soc. Robot*, 11(1), pp. 171-184,

[16] S. Josh, S. Šabanović. (2019). Robots for inter-generational interactions: implications for nonfamilial community settings. *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp. 478-486.

[17] J. Shimaya, Y. Yoshikawa, O. Palinko, N. Jinnai, K. Ogawa, & H. Ishiguro, (accepted). Active Participation in Lectures via a Colloaboratively Controlled Robot, *Int. J. Soc. Robot.*

[18] S. Chandra, P. Alves-Oliveira, S. Lemaignan, P. Sequeira, A. Paiva & P. Dillenbourg, (2016). Children's peer assessment and self-disclosure in the presence of an educational robot, *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, New York, NY, pp. 539-544.

[19] E. Takano, T. Chikaraishi, Y. Matsumoto, Y. Nakamura, H. Ishiguro & K. Sugamoto. (2009). Psychological Effects on Interpersonal Communication by Bystander Android using Motions Based on Human-like Needs. *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, IROS 2009, pp. 3721-3726.

[20] C. Birmingham, Z. Hu, K. Mahajan, E. Reber, & M. Mataric. (2020). Can I Trust You? A User Study of Robot Mediation of a Support Group. *2020 International Conference on Robotics and Automation (ICRA)*

[21] M. Traeger, S. Sebo, M. Jung, B. Scassellati, N. Christakis. (2020). Vulnerable robots positively shape human conversational dynamics in a human–robot team. *Proceedings of the National Academy of Sciences.* 117. 201910402. 10.1073/pnas.1910402117.

[22] H. Tennent, S. Shen & M. Jung. (2019). Micbot: A Peripheral Robotic Object to Shape Conversational Dynamics and Team Performance. *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Daegu, Korea (South)*, pp. 133-142.

[23] D. Sakamoto, K. Hayashi, T. Kanda, M. Shiomi, S. Koizumi, H. Ishiguro, T. Ogasawara, & N. Hagita. (2009). Humanoid robots as a broadcasting communication medium in open public spaces. *Int. J. Soc. Robot.* 1, pp. 157–169.

[24] Y. Todo, R. Nishimura, K. Yamamoto & S. Nakagawa. (2013). Development and evaluation of spoken dialog systems with one or two agents through two domains. *International conference on text, speech and dialogue. Springer, Berlin*, pp. 185–192.

[25] N. Karatas, S. Yoshikawa, & M. Okada, (2016). Namida: Sociable driving agents with multiparty conversation. *Proceedings of the Fourth International Conference on Human Agent Interaction, Singapore, 4–7*, pp. 35–42.

[26] T. Iio, Y. Yoshikawa, M. Chiba, T. Asami, Y. Isoda, & H. Ishiguro, Twin-Robot Dialogue System with Robustness against Speech Recognition Failure in Human-Robot Dialogue with Elderly People, *Applied Sciences.* 10(4), pp. 1522.

[27] H. Kumazaki, T. Muramatsu, Y. Yoshikawa, Y. Matsumoto, H. Ishiguro, T. Sumiyoshi, M. Mimura, & M. Kikuchi. (2019). Comedic experience with two robots aided a child with autism spectrum disorder to realize the importance of nonverbal communication. *Psychiatry and Clinical Neurosciences.* 73, pp. 423-423.

[28] I. Leite, M. McCoy, M. Lohani, D. Ullman, N. Salomons, C. Stokes, S. Rivers, B. Scassellati. (2015). Emotional storytelling in the classroom: Individual versus group interaction between children and robots. *2015 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp. 75-82.

[29] R. Hirayama, T. Kusumi. (2004). Effect of Critical Thinking Disposition on Interpretation of Controversial Issues. *The Japanese Journal of Educational Psychology* 52(2), pp. 186-198. (in Japanese).

[30] M. Salsali, M. Tajvidi, S. Ghiyasvandian. (2013). Critical thinking dispositions of nursing students in Asian and non-Asian countries: a literature review. *Global Journal of Health Science.* 5(6), pp. 172-178.

[31] A. Aron, E. N. Aron, M. Tudor, & G. Nelson. (1991). Close relationships as including other in the self. *Journal of Personality and Social Psychology.* 60(2), pp. 241-253.

[32] P. Manning, & R. George. (1993). Shyness, Self-Confidence, and Social Interaction. *Social Psychology Quarterly.* 56(3), pp. 178-192.

[33] L. Festinger. (1957). [1954]. A Theory of Cognitive Dissonance. California: Stanford University Press.

[34] W. O. Bearden, R. G. Netemeyer, & J. E. Teel. (1990). Further Validation of the Consumer Susceptibility to Interpersonal Influence Scale". *Advances in Consumer Research Volume 17, eds. Marvin E. Goldberg, Gerald Gorn, and Richard W. Pollay, Provo, UT : Association for Consumer Research*, pp. 770-776.

[35] B. Littig. (2004). The Neo-Socratic Dialogue (NSD): a method of teaching the ethics of sustainable development. *Galea C (ed) Teaching business sustainability: vol. 1, from theory to practice.* Greenleaf Publ, pp. 240-252.

[36] E. T. Harada, M. Ikenaga, R. Shimoyama, & T. Sawada. (2018). Conversation with Elderly: What kind of communication lords do younger members bear? *2018 Cognitive Aging Conference.*