A Multimodal Robot-Driven Meeting Facilitation System for Group Decision-Making Sessions

Ameneh Shamekhi

Khoury College of Computer Science, Northeastern University, Boston, Massachusetts, United States, ameneh@ccs.neu.edu

Timothy W. Bickmore

Khoury College of Computer Science, Northeastern University, Boston, Massachusetts, United States, bickmore@ccs.neu.edu

ABSTRACT

Group meetings are ubiquitous, with millions of meetings held across the world every day. However, meeting quality, group performance, and outcomes are challenged by a variety of dysfunctional behaviors, unproductive social dynamics, and lack of experience in conducting efficient and productive meetings. Previous studies have shown that meeting facilitators can be advantageous in helping groups reach their goals more effectively, but many groups do not have access to human facilitators due to a lack of resources or other barriers. In this paper, we describe the development of a multimodal robotic meeting facilitator that can improve the quality of small group decision-making meetings. This automated group facilitation system uses multimodal sensor inputs (user gaze, speech, prosody, and proxemics), as well as inputs from a tablet application, to intelligently enforce meeting structure, promote time management, balance group participation, and facilitate group decision-making processes. Results of a between-subject study of 20 user groups (N=40) showed that the robot facilitator is accepted by group members, is effective in enforcing meeting structure, and users found it helpful in balancing group participation. We also report design implications derived from the findings of our study.

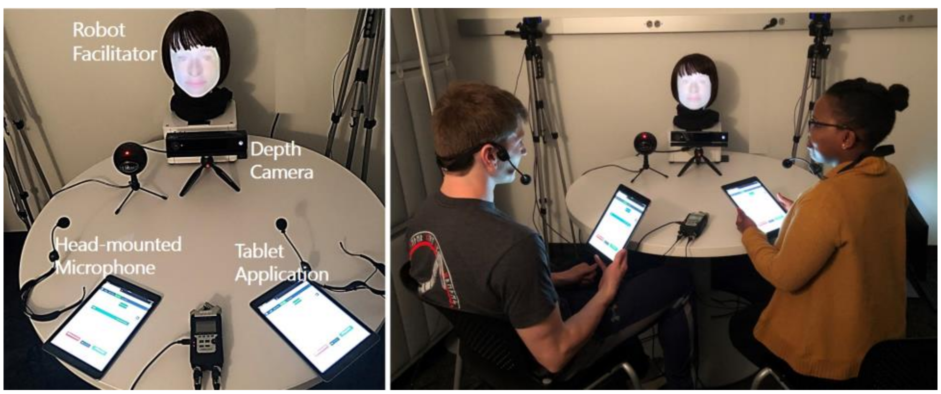

Figure 1: The automated group facilitation system and the study setting; the robot facilitates a small group decision-making session. Participants use a tablet application for some of the decision-making tasks

CCS CONCEPTS

• Human-centered computing~Interactive systems and tools • Human-centered computing~Collaborative and social computing systems and tools

KEYWORDS

Multimodal System, Human-Robot Collaboration, Group Facilitation, Interactive System

1 Introduction

In most organizations, meetings are one of the primary means of communication and coordination. On average, one third of employees' time is spent in meetings, involving many kinds of collaborative decision-making tasks [1, 19]. In group decision-making meetings, employees can view a problem from different perspectives, get exposure to a diverse set of insights and ideas that are not otherwise accessible, and reach consensus on decisions. However, there are many challenges in group decision-making meetings that can lead to suboptimal performance and efficiency. For example, group dynamics such as social influence, conflict, social loafing [24], groupthink [3], and dominance [9] can complicate the decision-making process and impact a group's performance. As a result, there is some debate about whether group decision making necessarily results in a more effective decision-making process and better decisions compared to individual decision making [17].

Professional human facilitators have been shown to be effective at improving group decision-making processes and outcomes. For example, Westley and Waters demonstrated that group facilitators are effective at addressing many of the common problems in meetings, such as not establishing or diverging from a meeting agenda, intragroup conflict, and domination of the discussion by one participant [43]. Viller also demonstrated that a group facilitator can greatly improve meeting quality by enforcing a structure, eliciting equal participation, and managing conflicts [41].

Despite the potential benefits of group facilitators, the high cost of hiring professional consultants or training in-house facilitators prevents many organizations from using them for the myriad meetings held every day. Automated systems may be able to perform many of the functions of a group facilitator at a fraction of the cost, and can be available whenever needed, for all meetings regardless of size or importance. We posit that a physically embodied conversational agent (ECA), with the ability of recognizing and displaying human-like interactional social cues, can provide effective meeting facilitation that requires minimal user training. Although significant research has been conducted on the design and development of ECAs that interact with single users, there is much less research on ECAs, virtual or robotic, designed for multiparty interaction. Meeting facilitation provides an ideal problem domain and testbed for multimodal multiparty interaction research that has the potential for significant societal impact.

We hypothesize that several capabilities of a conversational *humanoid* robot, such as the ability to interact with human collaborators in natural language and nonverbal conversational modalities, the ability to hold a non-judgmental and neutral point-of-view, and access to information resources, can make a Robot Facilitator an effective meeting facilitator. Moreover, as suggested by previous research [36] a conversational agent with human-like embodiment, functioning as the group facilitator, could improve the users' perception of the facilitator's social behavior, intelligence and authority, and serve as the locus of attention to lessen distraction.

We designed and developed an Automated Group Facilitation (AGF) system in which a robot facilitator uses multimodal sensory inputs to moderate a small group decision-making session. In our system, meeting facilitation functions, such as agenda management, time management, and participation management, are layered on top of more foundational conversational processes, such as turn-taking and grounding, which are necessarily multimodal. Our robotic facilitator uses cues from user gaze, speech, prosody, and proxemics, in addition to task status from a tablet application, to coordinate a decision-making meeting.

In this paper, we describe a prototype multimodal robot meeting facilitator, that performs the following basic functions: agenda management, to ensure all essential topics are discussed; time management, to ensure the meeting is conducted in an efficient manner; and participation management, to ensure the opinions of all participants are heard. We evaluate our prototype in the context of a hiring meeting, addressing some of the known challenges in these meetings, including lack of structure and imbalanced participation (dominance). In addition to measuring task outcomes, we are particularly interested in assessing whether participants accept an ECA in the role of a facilitator, and how an ECA impacts the social dynamics of meetings.

The main contributions of this work are: 1) we provide initial evidence confirming the use and acceptability of an anthropomorphic robot in the role of a meeting facilitator; 2) we present a new design for a fully-automated robot-driven group facilitation system; 3) we offer important design insights for the use of robot facilitators in groups. In the rest of this paper, we first review related work, then describe the architecture and the design of the new automated facilitation system. We finally present results of an evaluation study to demonstrate how the system works/is perceived by participants in a laboratory setting.

2 Related Work

2.1 Group Decision-Making Challenges and Group Facilitation

There have been many studies of the social dynamics that unfold in human workgroups, dysfunctional behavior that can lead to poor outcomes, and solutions for holding productive meetings. Latane's "social impact theory" [26] characterizes the complex interactions among different social factors in groups which, together, determine how individuals are impacted by other group members. The theory explains how many factors—such as lack of an agenda, dominant individuals, and "derailers"—can result in dysfunctional meetings, frustration, and poor outcomes. Nunamaker, et al, reviewed work on Group Support Systems, and identified additional factors that could lead to poor meeting outcomes, including a lack of focus, hidden agendas, conflict, distractions, and poor planning [33]. Other studies have demonstrated that not having an agenda or not following it represent major barriers for meeting efficiency [34].

Several studies have demonstrated that having a *group facilitator* can improve group performance and decision-making effectiveness by mitigating some of these challenges and enforcing a structured decision making process [8, 30]. Previous research on meeting dynamics demonstrated that one of the main tasks of group facilitators is to manage a meeting by keeping discussions on track and ensuring the group follows an agenda [31, 41].

2.2 Technology-Driven Group Facilitation

Viller examined the role and responsibilities of a group facilitator in computer supported cooperative work (CSCW) and discussed the extent to which this role should be automated in the group setting. He described how a facilitator's role changes over a meeting lifecycle, from a more central role in the beginning, stepping back and only providing high-level interventions in the middle stage, and again taking on a more directive role at the end, integrating the ideas and wrapping up the meeting. Viller identified five general meeting problems that can be addressed by a group facilitator, including: lack of structure, domination, conflict, and lack of engagement [41].

A number of Group Decision Support Systems (GDSSs) have also been designed to support or even replace some of a human facilitator's roles. GDSSs in general can support face-to-face and remote group meetings and different types of GDSSs are designed to remove communication barriers, enforce structured decision making techniques, and use formalized rules to systematically direct the timing or content of discussion [12].

To tackle a lack of leadership and promote structured collaboration in distributed groups, Farnham et al. customized a chat system to enforce the structure of group communication and the group decision-making process. They conducted an evaluation user study of the system in a hiring decision making scenario, and demonstrated that participants who used the system were more likely to reach consensus, made higher quality decisions, and showed a greater recall of discussions [15] .

Prior work in the CSCW community has also investigated how technology can enhance group collaboration and improve unequal participation in meetings. For example, Meeting Mediator is a real-time portable system that provides visualized feedback of members' verbal participation on a mobile display to enhance overall group collaboration [25]. Furthermore, much research has been done to design smart meeting rooms both to provide real-time assistance and to capture human behavior in groups [11, 42]. Smart meeting rooms are usually equipped with different sensors, multiple cameras and microphones to capture details of a meeting and detect contextual information. In recent work by Bhattacharya et al., a model is proposed to automatically process multimodal sensors outputs, and detect human verbal and non-verbal behaviors and understand group dynamics [5]. The work to date in this area indicates the great potential of employing intelligent systems for group facilitation. However, to the best of our knowledge none of the work on GDSSs has leveraged multimodal ECAs as virtual meeting facilitators.

2.3 Robots and Conversational Agents in Groups

Advances in robotics technology have enabled a growing number of robots designed to work in human teams [16], with examples spanning from invasive telesurgeries [14] to space exploration missions [13]. Several researchers have studied how the presence of robots in teams can affect human group dynamics and team performance. Jung and his colleagues have explored the use of robots to moderate team conflicts through a Wizard-of-Oz study in which a robot attempts to repair interpersonal violations in 3-member teams [23]. Jung also introduced the concept of "affective grounding", as one of the essential prerequisites for coordinating human-robot interaction [22]. Matsuyama conducted a series of studies investigating a robot-based facilitation framework designed to balance participant engagement in a group conversation. Results of an evaluation study indicated that the procedural approach to manage the floor was more accepted and improved feeling of groupness [28]. Robots have been also used as moderator and team member in the game context to study the effects of different roles that a robot can play in game teams [7, 25]. Other studies investigated effects of a conversational agent's embodiment in the group setting and found physical embodiment of robots can improve group engagement, attentiveness, and social perception of the agent [2, 36]. Tennent et al. studied effects of a non-humanoid robot promoting participant engagement in a group. They evaluated a peripheral robotic object (microphone) that shows non-verbal implicit actions to follow and encourage participation in a group conversation [40]. Their intervention resulted in increased participant engagement.

There is also a growing body of literature on ECAs designed for multiparty interaction [6]. For example, Bohus and Horvitz conducted a series of studies and proposed a model to represent an ECA's turn-taking behaviors by gaze, gestures and speech [6]. Such a model can smooth an ECA's speech flow in a multiparty conversation. Skantze et al., also explored how a robot can manage turn-taking and attention when multiple players are working with a shared display [38]. In other studies, non-verbal cues such as gaze and respiration were used to predict the next speaker and facilitate turn-taking management in a multiparty interaction [20].

3 Automated Group Facilitation (AGF) Robot; System and Architecture

We designed the Automated Group Facilitation (AGF) system to support collaborative activities in group meetings by providing interventions to mitigate some of the known challenges such as diverging from the agenda and meeting domination. The AGF system is *multiparty*, interacting with all meeting participants, and is *multimodal*, using diverse inputs and outputs to support natural conversational interaction with group members. To further facilitate natural interaction, the system is personified with a humanoid robot that serves as a group facilitator (Figure 1) and is able to conduct a quasi-natural conversation. We chose to use a physical robot as the group facilitator, since several studies have demonstrated that physical robots are more engaging than virtual (screen-based) agents [2], the robot's physicality improves user's perception of its presence in the team [36], and a physical robot can use its gaze to more accurately display attentiveness and feedback to speakers and to indicate a desired addressee [2].

The AGF system uses three channels of sensory inputs to enable the robot to see group members (via 3D depth camera), hear their voice (microphone), and understand the decision-making task status (via events from a hiring task web application used by participants). Multimodal outputs of the system include: the synthetic speech of the robot, conversational nonverbal behavior (lip synchronization, communicative facial displays, gaze orientation), and actions such as updates to the task application.

The main goals of the system are to improve overall meeting structure and quality and increase engagement in and satisfaction with the collaborative activity by providing just-in-time instructions. To achieve this, the system is designed to provide two different types of support during a group decision-making session: social facilitation and decision-making and meeting facilitation.

Social Facilitation: One of the most important tasks of a group facilitator is to catalyze the social interaction among members of a workgroup [35]. Social facilitation tasks of a group facilitator include: greeting everyone and introducing herself and the members of the group; "setting the stage" by orienting group members to the task and process; providing “ice-breakers” to acquaint group members with each other and begin building trust and rapport; discussing desired meeting outcomes; and reviewing the agenda. For group members to accept the robot in this social role, it needs to engage them in natural, multimodal, face-to-face conversation. In order to build rapport and trust as a social facilitator, it needs to display understanding and empathy using conversational backchannels [4] and active listening behavior [27].

Decision-Making and Meeting Facilitation: In order to alleviate some of the common challenges of group decision-making meetings, such as inefficient and unstructured discussion and imbalanced participation, the AGF system is designed to provide five types of facilitation intervention:

- *Enforce meeting structure:* To avoid off-topic discussions and deviation from the meeting agenda, the system enforces the meeting structure by offering step-by-step instructions and monitoring the decision-making process. The system needs to know the context and the meeting stage to offer timely feedback, and guide the meeting flow by asking targeted questions.
- *Encourage Balanced Participation:* Dealing with people who tend to dominate the conversation in a meeting is an important function of a group facilitator [8]. To ensure balanced participation, and to create a safe space for all participants, the system detects when some participants are significantly less vocal during a meeting and engages them in the conversation.
- *Content-based Feedback:* A group facilitator also provides feedback related to the task being discussed by the group. Hence the AGF system needs to recognize some speech content to provide interventions based on the topic of participants’ discussion. For example, the system should track aspects of the task that need to be discussed and remind the group if they forget to discuss them.
- *Time Management:* Group facilitators keep meetings on track by watching the time and agenda. The AGF system should also track the time spent on each task in the meeting to both remind team members of the time, if needed, and offer comments based on the time and the stage of the meeting.

To provide these facilitations effectively, a facilitator should also be perceived as an authorized and intelligent member to enforce participants’ attendance and adherence to her instructions. As suggested by Cassell, the computer agent’s embodiment helps users to better locate its intelligence and power and improves perceived trust [10]. Using a conversational robot with a human-like face and nonverbal behaviors such as gaze and facial expression as a group facilitator allows participants to treat it as an intelligent and authorized social entity with a persona.

Another important feature of the group decision-making setting is the involvement of multiple humans in the interaction, so the Robot Facilitator must support multi-party conversations. In order to engage in any natural interaction with group members, the robot needs to have certain minimal conversational competencies. To conduct multiparty interaction with all group members, the robot should be designed to receive inputs from multiple sources in order to detect the speaker and addressee of any utterance. It also needs to understand at least some user speech content to provide the right action or response via speech recognition. The robot also needs minimal conversational turn-taking management capabilities, tracking which participant is speaking and indicating attentive behavior by gazing at the speaker, as well as indicating her intention for a specific user to speak by gazing toward them.

3.1 System Architecture

To satisfy the requirements mentioned above, we use a 3-layer system architecture to model group interactions, recognize participant behavior, and provide proper intervention by the robot. This framework adapts the Smart Meeting Rooms’ architecture [42] and integrates it for use with a conversational robot. The facilitator robot is a humanoid head developed by Furhat robotics (Figure 1). It has an animated face, back-projected on a translucent mask, mounted on a two degree-of-freedom mechanical neck that allows the head to turn and nod [2].

In this architecture, audio data and tablet application inputs are used to track the meeting stage, calculate participation time, and provide proper interventions based on the meeting stage, time, and content. The audio input is passed to an automated speech recognition (ASR) module to be used by the dialogue manager. The AGF architecture uses the inputs from the tablet application in addition to the speech input to better detect the current topic, the meeting stage and the decision status. Visual data from the camera is processed to make the robot aware of the participants' location, and head positions and generate adequate non-verbal and verbal listening behavior. The components are controlled in different software modules, and we use an event-driven architecture to handle messaging between modules and the dialogue manager. Events from the camera and microphones are transferred in an internal framework (described in section 5) and events from the application are stored in a local database and passed to the dialog manager through the network. Figure 2 shows the architecture.

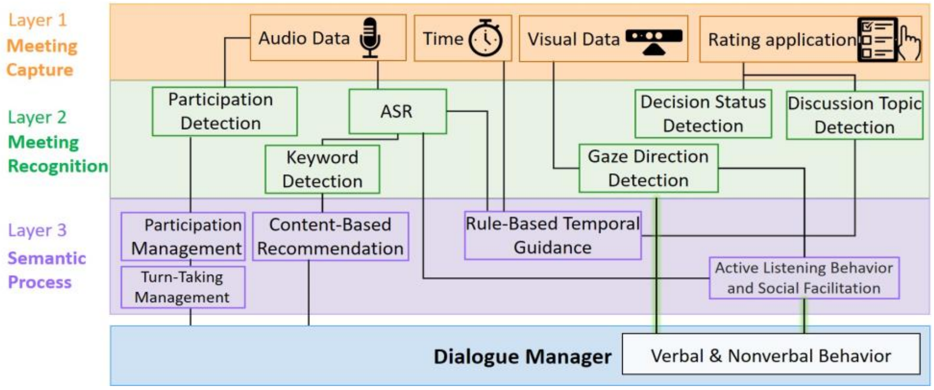

Figure 2. The AGF System Architecture for Decision-Making Meetings

4 Task Domain: Hiring Meeting

Although our system architecture and methods can be used for many kinds of group decision-making meetings, we grounded our development in a single task domain: a hiring meeting. The hiring decision meeting is a very common scenario for group decisions in the workplace, and previous work used similar tasks to study group decision processes [15]. This experimental task and procedure are adopted from previous work [36], in which the authors explored the effects of embodiment of a virtual facilitator in a group setting. For this task, a group of two participants are asked to review and discuss six resumes for a sales manager position and select the one best candidate for an interview. We created fictional resumes based on real resumes of applicants for the sales manager positions. Resumes were designed so that each candidate had unique strengths and weaknesses, and did not have an obvious rank ordering among them for the position to be filled. Participants are given 30 minutes to review, discuss and select the best candidate. A humanoid robot serves as the facilitator for the group decision-making session. We designed a standard agenda [29] to structure our experimental hiring meeting (Figure 3). The steps are as follows:

1. **Greeting and Introduction:** The facilitator initiates the conversation by introducing herself, inviting the participants to introduce themselves, and greeting them with their names. The facilitator then leads several "ice-breaker" exercises to prepare the group for the task. For example, she asks each participant to tell her and the other participant about their favorite part of their work. This conversation acts as a social catalyzer to build rapport and trust among group members.
2. **Agenda Description and Resume Review:** The facilitator then describes the goal of the meeting and the agenda, and asks the group to review resumes individually for a few minutes and provide a preliminary ranking using the tablet application. She also describes how the tablet application interface should be used to rank the resumes. She lets participants know that she will keep track of the time to remind them as the time runs out. Participants can negotiate time extensions if they require more.
3. **Criteria Discussion:** When participants are ready, the facilitator suggests that the group discuss the hiring criteria. She monitors the conversation to ensure the group discusses the main criteria. If participants indicate they are ready to move on but have forgotten to discuss one or more criteria (e.g., education or work experience), the facilitator prompts them (e.g. *"What do you think about their education?"*). During the discussion the facilitator displays active listening behaviors and confirmations. She also monitors vocal participation and prompts less active participants by name during breaks in the conversation.

4. **Main Discussion and Elimination of Unfavored Candidates:** The facilitator invites the group to start discussing the hiring candidates, by going through all the resumes and eliminating the unqualified ones. When participants converge on a decision (by indicating on the tablet application they are keeping or eliminating a particular candidate), the facilitator confirms that decision. If participants skip to another candidate too soon without a decision, the facilitator asks them whether they want to make a decision first. If they cannot decide about a candidate after a few minutes, the facilitator suggests that the group discuss the next candidate. After all candidates are discussed, if too few have been are eliminated, the facilitator suggests that the group discuss and eliminate more before proceeding. Once the hiring candidates have been narrowed down sufficiently, the facilitator summarizes the remaining candidates, and moves on to the next agenda item. As in the previous task step, the facilitator also monitors the time and vocal participation.
5. **Selecting the Best Candidate:** When the number of hiring candidates has been sufficiently narrowed, the facilitator asks participants to discuss the pros and cons of the remaining candidates to select the best one. If participants have difficulty (pausing for too long), the facilitator provides suggestions to facilitate the decision-making (e.g. to review the criteria mentioned in the job description, or suggest the group reflect on their initial voting). The facilitator again monitors time and vocal participation.
6. **Decision on the Final Candidate:** When the meeting time is up or the group reaches a final consensus decision, the facilitator summarizes the session and confirms the final decision with each participant.
7. **Meeting Wrap Up:** Finally, the facilitator thanks the group for participating and asks them to use the tablet application to rank the candidates individually based on their discussion.

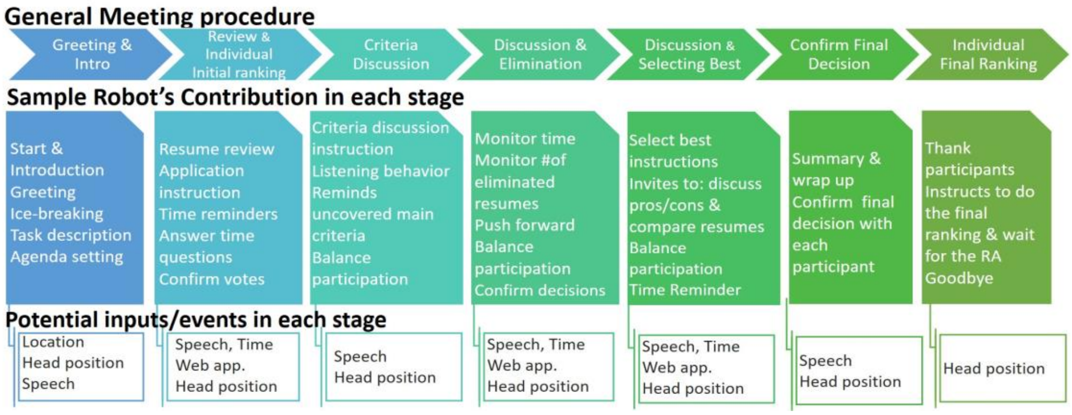

Figure 3. Group decision-making procedure. Examples of robot's contribution and multimodal inputs in each stage are shown in the bottom.

5 Dialogue Management

The dialogue manager of the AGF system is designed to accomplish five main goals to facilitate the group decision-making process: 1) Handle multiparty interaction; 2) Enforce structure for the group decision-making meeting; 3) Ensure content coverage by providing content-based recommendations; 4) Balance group participation by nudging the less active participants; 5) Track and manage the time by conveying rule-based temporal guidance.

We used IrisTK, an open source dialogue system framework [39] to manage the conversation flow in the multiparty interaction with the Furhat robot, using a state machine-based dialogue manager. Using IrisTK we designed an event-driven rule-based dialogue manager that uses information from the input modules and creates appropriate responses. In section 5.1 and 5.2, we describe more details on the input signals, the output actions the system provides to achieve the goals mentioned above.

5.1 Input Modules

Four main input modules are used to capture information about the meeting.

5.1.1 Visual input from depth camera

A Kinect camera is used to capture visual and depth information. Inputs from the Kinect enable the system to sense user presence, identify user location, and estimate their gaze direction using head orientation as a proxy. This module provides the number of users, sends an event when a user enters or exits the scene, and detects whether a user is attending to (gazing at) the robot or not.

5.1.2 Voice input from microphone

User speech is captured by head-mounted microphones and recognized and converted to text using Google's cloud-based automatic speech recognizer (ASR). The speech signal is also processed to provide three additional types of information: 1) silence vs. speech intervals for each user; 2) who the current speaker is; and 3) user vocal participation from the accumulated speech time of each participant in each meeting stage. To understand the output of the ASR module, and to handle the semantic interpretation of the speech input, we used the W3C Speech Recognition Grammar Specification (SRGS)[1] and Semantic Interpretation for Speech Recognition (SISR)[2] tools integrated into the IrisTK framework. The recognized semantic intent of the speech input is used to drive robot behavior in each dialogue state throughout the interaction. For example, the system uses that to understand general speech inputs such as answers to yes-no questions, repeat requests, queries including the robot's name, questions and thanking utterances.

5.1.3 Decision status from the hiring task tablet application

We designed and developed a custom tablet application to support the group decision-making tasks. The application is used for individual and group activities during the group-decision making activity. It provides some of the basic features of group decision support systems such as voting and sharing the decision status. The application facilitates the communication and sharing of opinions by providing a discussion interface that reflects the group decision at each stage of the decision-making process. Participants use the tablet application to review and rank candidates individually, and to share ideas during the group discussion.

The application has three views for the three stages of the meeting: 1) Individual resume review and initial ranking (Fig.4.a) 2) Group discussion, elimination and selecting the best candidate (Fig.4.b and c), and 3) Individual final ranking (Fig.4.d). During the group discussion phase, most of the participant actions are shared with other participants to facilitate communication and sharing of ideas. The application logs some contextual data (e.g., the meeting stage, the active candidate, and the group decision status on each candidate) and sends signals to the dialogue manager on occurrence of each event. The dialogue manager obtains data such as the user names, candidate's status, and the number of discussed and eliminated candidates from the application. Other input signals are received from the application at the following moments during the session: 1) when initial ranking is submitted during the individual review phase, 2) when the active candidate is updated during the group discussion, 3) when the decision status on each candidate is updated (kept or eliminated), 4) when the best candidate is selected.

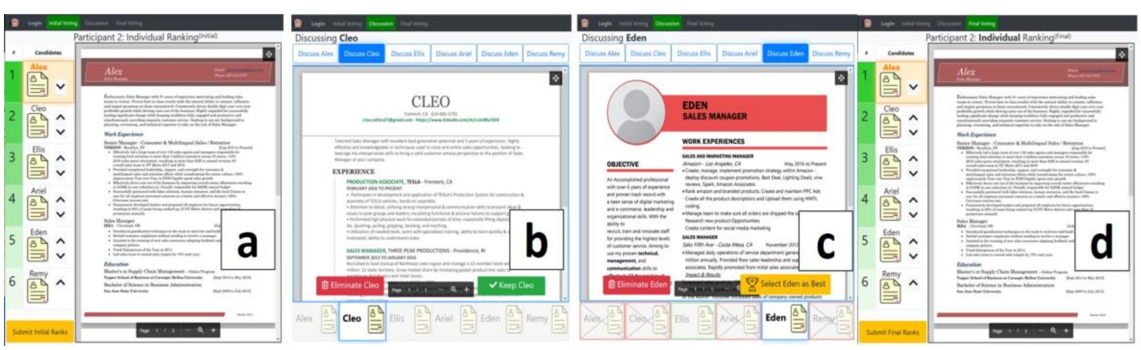

Figure 4. The hiring task application; a) Initial individual ranking interface; b) Group discussion interface; c) Group discussion interface after eliminating four candidates; d) Final individual ranking interface

[1] https://www.w3.org/TR/speech-grammar/
[2] https://www.w3.org/TR/semantic-interpretation/

5.1.4 Time

In the dialogue manager, timer events are used by the dialogue manager to initiate timing signals at several occasions during the interaction. Examples include: during the resume review phase to send time reminders, and during the discussion phase to track the time spent on each stage and on each resume.

5.2 Output Actions

In this section, we describe the outputs generated by the dialogue manager based on inputs presented in the last section. The three different types of outputs are: non-verbal, verbal, and application actions.

5.2.1 Non-verbal behaviors

The robot facilitator performs a range of human-like non-verbal behavior to show active listening behavior and to manage the turn-taking in multiparty interaction. When the group members enter and when they are conversing, the robot gazes towards the active speaker, smiles, nods her head, and uses confirmation filler words (aha, I see, that makes sense, etc.). Across all stages of the meeting, the robot handles multiparty interaction by attending to each participant when she asks them a question, or when they are speaking. The robot looks at the group (i.e., shifts gaze among participants) when she is making a general statement.

5.2.2 Verbal behaviors

Using the processed information from the input modules the dialogue manager decides what the robot facilitator should say based on the meeting stage, user speech input, time, or inputs from the tablet application. The verbal outputs are categorized as follows:

1. Robot-Initiated Utterances: These are the instructions, comments, and questions that the robot utters based on the dialogue state, the meeting stage, and/or the decision status. Examples include: asking each participant to introduce themselves at the start of the meeting, and providing the task description and goal to be accomplished at the beginning of each stage.
2. User Speech-Initiated Utterances: Participants are told that they can interact with the robot at any time in a way similar to how they interact with humans. The robot uses SRGS and SISR to interpret participants utterances and respond to them when they call her name, greet her, ask a question about the time, etc., at any point during the meeting.
3. Tablet Application-Initiated Utterances: These commands are triggered as users interact with the application on their tablets. For example, when a participant submits a candidate ranking on the application, the robot thanks them and check the status with the other participant.
4. Time-based Utterances: Using the internal time of the system, the robot provides time reminders at certain points during the session.

In all dialogue states, if the robot does not hear users speech input (the voice is low), or if she does not understand the input (e.g. ASR output does not match expected semantic intents) in response to a query, she informs participants that she did not understand what they said and repeats her question. Also, in cases that the system receives multiple speech inputs at the same time, the robot handles the responses separately.

5.2.3 Actions

The dialogue manager may also output an action instead of verbal or non-verbal behavior. For example, the dialogue manager may extend the time allowed for a given decision-making task (upon user request), move forward to the next task, update the tablet application interface, or calculate participants' contribution time to nudge less active participants.

5.3 Meeting Facilitation Services

The AGF system adopts the output actions to achieve the four Decision-Making and Meeting Facilitation functions described in Section 3.

Enforce meeting structure: The system tracks the meeting stage and enforces the meeting structure by providing certain services such as offering step-by-step instructions to move forward to next stage, asking targeted questions, acknowledging actions and asking confirmation questions, monitoring the decision-making process and suggesting corrective actions (e.g. try to eliminate more options), providing feedback, decision summaries and wrapping up the sessions.

Ensure content coverage: At some stages during the meeting, the system listens to the discussion and provides comments and instructions based on keywords detected in the content. For example, the system monitors the conversation when participants are discussing hiring criteria, and reminds them to further discuss any of the main criteria that have not been adequately covered.

Encourage and Balance Participation: At the end of each stage of the decision-making process (e.g. criteria discussion or each resume discussion), the system references the tracked vocal participation for each user as their contribution to the discussion. If an imbalance is discovered (e.g. >10 seconds difference in speech contribution), the robot nudges the less active participant by asking what s/he thinks, or if s/he wants to add anything else.

Manage the Time: The robot provides some comments and feedback based on the time and the stage of the meeting, as ascertained from the tablet application input. The system also tracks the time spent on each stage, and is able to extend the time upon request.

6 System Evaluation

We conducted a randomized between-subjects experiment to evaluate the effectiveness of the automated group facilitation system used by teams of two participants in the hiring decision-making task described in Section 4. We compared the full system (the robot, tablet application, and supporting multimodal system) to a control condition in which participants performed the same hiring decision-making task without the facilitation robot (they still had access to the same tablet application for individual tasks and group decision-making). In both conditions, each participant was given a job description printed on paper, and a tablet computer (iPad) with the hiring application running. Participants used the application for their individual review and ranking of resumes, as well as to their group discussions. In the ROBOT condition, the robot facilitator provided instruction for each stage of the meeting (see Section 4). In the CONTROL condition participants were provided with written instructions for the meeting. The written instruction includes the same agenda that the robot uses in the ROBOT condition, as well as time management instructions. Figure 5 shows the setup of the experiment. All sessions were video recorded using four synchronized cameras for subsequent analysis.

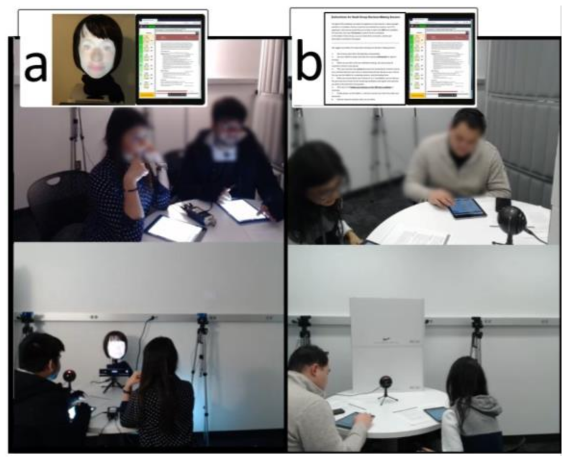

Figure 5. Experimental setup for the evaluation study. a) robot condition; b) no-robot condition

The tablet application was used in both conditions to view the resumes of hiring candidates, and to record the results of individual candidate ranking and display group decisions (Figure 4). In the ROBOT condition the interface was automatically updated based on the meeting stage and intervention actions taken by the AGF system, while in the CONTROL condition navigation instructions were shown as alert messages, and participants could use a navigation bar to move to different views.

Participants: Participants were recruited by posting advertisements around our institution and on online job posting sites. Participants were required to be 21 or older, be able to read and speak English, and have some professional work experience involving participation in professional meetings. Participants were randomly matched to form two-member groups. To minimize the effect of gender on group performance, we ensured that the ratio of mixed-gender/same-gender groups were equal in the two conditions. The study was approved by our institution's IRB, and participants were compensated for their time.

Measures: Rapport with the facilitation robot was assessed by asking a set of questions suggested by [32]. Attitudes towards the robot were assessed in the ROBOT condition using the bond subscale of Working Alliance Inventory [18], and single item questions about the robot (e.g. the robot's knowledgeability, authority, and friendliness), and perceptions of the its effectiveness in managing participation and disagreements in the group (Table 1). Feeling towards their team and the decision-making process was assessed using a 5-item 7-point Likert scale (items: '*The members of this team get along well together*', '*I am pleased with the way my partner and I participated in the decision-making task*', '*I am very satisfied with working in this team*', '*I felt my partner and I understood each other*', '*I found we reached a decision efficiently*'). Perceptions of task conflict were measured using a three-item scale introduced and validated by Jehn et al., [21] (items: *'How frequently did people on your team disagree regarding the work being done?', 'How frequently were there conflicts about ideas on your team?', 'How frequently were there differences of opinion on your team?'*). Satisfaction with the overall system was assessed via a 9-item Likert scale questionnaire that included five questions to measure general usability (sample items include: *"I am satisfied with the interaction with the facilitation system", "I found it easy to interact with the facilitation system."*) as well as four questions assessing decision-making facilitation usability of the system (sample items include: *"The facilitation system was able to manage the discussion and interaction during the meeting."*). Perception of group performance was measured with a single item scale (1-100 .

Confidence in the hiring decision was assessed via self-report within the tablet application (1=not at all confident, 5=very confident), measured immediately after participants performed their individual rank-ordering, and after they reached a final hiring decision as a group. **Task outcome** was also assessed by analyzing initial and final candidate rankings. Metrics of **meeting structure and time management** were derived from analysis of session video recordings, coding whether groups performed each of the task stages, and how much time they spent on each stage (Table 2).

Procedure: Prior to the study session, a research assistant described the decision-making task, obtained informed consent, and administered sociodemographic questionnaires. Participants were given with the tablet application. The research assistant then showed them the tablet application, explained the interface, and how to use it for different tasks. Participants were told how much time they had for the task and were guided to the room with the facilitation robot to start the experiment.

In the ROBOT condition, the research assistant first introduced the robot and explained how she will see and listen to the participants. They were instructed to wear head-mounted microphones and seated, and the research assistant left the room after starting the program.

For all participants, the research assistant monitored the session from another room and came back to end the session when either 30-minutes had elapsed, or the group reached a consensus hiring decision. Participants were asked to fill out post-session questionnaires, and a semi-structured interview was conducted with both team members.

6.1 Quantitative Results

A total of 40 individuals (20 teams) completed the study. 60% of participants were male, aged 24.2 (SD=1.8). Most (95%) participants did not know each other before the experiment, and most (85%) were university students. Among the 20 groups, 11 (N=22) were randomly assigned to the ROBOT condition, and 9 (N=18) assigned to the CONTROL condition.

Participants in the ROBOT condition rated the robot significantly higher than neutral on *rapport*, $t(21)=10.08$, $p<.001$, and *Working Alliance*, $t(21)=8.08, p<.001$. Participants also rated the robot's knowledge, authority, friendless and her abilities in meeting facilitation significantly higher than neutral (Table 1).

Table 1. Self-Report ratings of the robot facilitator (Sarah) (N = 22), single sample Wilcoxon signed ranked test demonstrates all ratings were significantly greater that neutral

Items **7-point Likert scales (1= strongly disagree, 4=neutral, 7= strongly agree)**	**Robot Mean (SD)**	***p-value***
I found the facilitation robot **knowledgeable**.	5.09 (1.7)	.014
I felt the facilitation robot was **powerful and confident**.	5.45 (1.4)	.001
I felt the facilitation robot was **friendly and warm** toward me.	6.36 (0.7)	.000

The facilitation system assured that my partner and I have **equal chance** to express our ideas.	5.86 (1.4)	.000
The facilitation system allowed **sufficient discussion**.	5.27 (1.6)	.003
The facilitation system **encouraged** participation	6.27 (1.03)	.000
The facilitation robot could effectively **manage the disagreements** between my partner and I.	4.95 (1.4)	.016
The facilitation robot **attempted to manage the conflicts** raised in the meeting.	4.81 (1.4)	.021

An independent sample t-test on responses to the team and decision-making satisfaction questionnaire (α=.88) indicated that participants in the ROBOT condition (Mean (SD)=6.4(.56)) were relatively more satisfied with their team and decision-making activity, compared to participants in the CONTROL condition (Mean(SD)=5.8(1.2)), t(38)=1.8, p=.06.

We also found groups in the ROBOT condition (Mean (SD) =1.8 (1)) reported less task conflicts than those in the CONTROL (Mean (SD) =2.5 (1.2)), and independent sample t-test showed marginal significance (t(38)=-1.69, p=.09.

We did not find any significant difference between the ROBOT condition (Mean(SD)= 5.68(1.05)) and the CONTROL condition (Mean(SD)= 5.4(1.48)) evaluating the system usability regarding decision-making facilitation as reflected in a 4-item composite measure (α =.81). However, participants in the CONTROL condition rated the general usability of the overall facilitation system (via a 5-item scale (α =.95)), higher (Mean (SD)=6.08 (1.1)) compared to those in the ROBOT condition, (Mean(SD)=4.99 (1.6)), t(38)=-2.42,p=.02. There were no significant differences between the two conditions on perceived performance and dominance.

There was no significant difference in the time (minutes) taken to reach a decision (ROBOT: Mean(SD) = 25.5(5.8), CONTROL: Mean(SD) = 24.7(5.2)), nor differences in the final selected candidate between the two conditions. However, we did see a significantly greater increase in decision confidence in the ROBOT condition (Mean(SD) =1.32(.9)) compared to the CONTROL condition (Mean(SD)=.56(.5)), t(38)=2.94, p=.03, with change measured immediately after participants performed their individual rank-ordering, and again after they reached a final hiring decision as a group.

Analysis of session videos yielded several significant differences in group behavior between conditions for each stage (Table 2).

Greeting: among the nine groups in the CONTROL condition, three groups did not have any forms of greeting and three groups had a very short (~10 seconds) introduction. All participants in the ROBOT condition greeted each other and the robot.

Criteria discussion: None of the groups in the CONTROL condition discussed hiring criteria before they started discussing candidates, while all groups in the ROBOT condition did this.

Final decision: In the ROBOT condition, after participants discussed the pros and cons of the final candidates, they select the best candidate, and then the robot confirms the consensus decision with each participant. In the CONTROL condition, the session was usually ended by one of the participants selecting the best candidate; only three groups (33%) reviewed and confirmed the best candidate together.

Participants in the CONTROL condition spent significantly more time reviewing resumes on their own before beginning discussion, t(17)=-4.27, p=.001. We also found a common pattern in the CONTROL condition in which one participant finished their review and waited in silence until the second participant finished his/her ranking. This caused the individual review stage to be significantly longer in the CONTROL condition (Mean (SD)=109.3(83.4)) compared to the ROBOT condition (Mean (SD)=24.8(27.8)) in which the robot confirms when one participant submits their individual ranking and asks the other participant whether they are finished or not, t(17)=-3.02, p=.008. During the Discussion stage, all groups in the ROBOT condition reviewed resumes in order (as shown in the application), compared to only 33% of those in the CONTROL group.

Table 2. The percentage of groups that performed each suggested meeting stage, and the amount of time spent on each stage, in each condition.

Meeting Stage	Percentage of groups who performed this stage			Time spent by groups on each stage (in seconds)		
	ROBOT (N=10)	CONTROL (N=9)	chi-square p-value	ROBOT (N=10) Mean (SD)	CONTROL (N=9) Mean (SD)	ttest p-value
Greeting	100%	67%	.03	152 (21)	21 (43)	.000
Individual review and ranking (Initial)*	100%	100%	1	307(56)	639(239)	.001
Criteria Discussion	100%	0	.000	172(38)	0	.000
Group Discussion	100%	100%	1	833(226)	598(302)	.06
Select Best Discussion	100%	89%	.82	91(85)	87(109)	.94
Finalize and Confirmation	90%	11%	.001	25(11)	1(2)	.000

* This is total time that the group spent on resume review and initial ranking phase. If participants finished their review in different times we considered the longer time here, as the group could only move forward when both participants finished their ranking.

6.2 Qualitative Results

From our qualitative analysis we derived themes related to the advantages and disadvantages of having a robot facilitator for small group decision-making meetings.

Overall most of our participants found the robot facilitator "helpful", "friendly", "interactive" and "well designed". All groups said that they would use and recommend using this system for group decision-making meetings, but some groups indicated it would be especially helpful in large and diverse groups when there are more likely to be disagreements, domination, and off-topic conversation. Several groups also mentioned that to them the robot "*is not only a robot*" and they felt "*there is a third person sitting*" with them, who unlike humans is very "*unbiased*".

6.2.1 The Robot's Role in Balancing the Participation

Out of 11 groups in the ROBOT condition, 10 groups proactively talked about how the robot "*ensured equal participation from both [participants] and gave [them] equal opportunity to speak up*". Many people appreciated this feature by comparing this experience with their past group meeting experiences. For example, group 8 discussed it like this: "P1: *Group discussions mostly are somewhat chaotic. [but] here we were given a fairly good chance to speak about each and every resume and to state what we like and what we do not like …. The opportunity given here was fairer than what generally happens in group discussion.* P2: *and the usual respect that we need to use space for each person is ignored, but Sarah*[3] *maintains that well.*". Some groups also found this feature helpful for resolving conflicts in the teams. For example, P1 in group 13 said: "*She gave Equal opportunities to both of us to speak, that kind of resolved conflict itself because everyone feels their voice is being heard."*

6.2.2 The Robot's Role in Organizing the Meeting and Time Management

Our participants found the robot helpful as a meeting moderator who follows a protocol step by step and guides the team. Participants mentioned how they found the robot's guidance helpful in remaining on track: "[G15-P1] *if we could not eliminate more than two people she was encouraging that we go through it again and come to a conclusion. So this part was really good.* [G15-P2]: *she was trying us to engage in conversation and have discussion, and she was guiding the whole process.".* One group compared their past experiences in group meetings and mentioned how the robot made the experience different: *"Sarah would be more efficient when we want to ensure that we are not digressing from the agenda. Because from our past experience we easily identify people who usually have the habit of not speaking on the topic. So Sarah can be good … because she keeps everyone on track.*

[3] Sarah is the name of the robot.

I would definitely recommend Sarah to be used for that.". Moreover, all groups reported on the usefulness of the robot in keeping track of the time, however some people desired less frequent reminders even when they confirmed they would have spent more time on the task without the time reminders.

Several groups mentioned the slow interactional response of the robot and delayed reaction as the worst thing about the system. "*the only problem with this one was it was taking a lot of time because of the delay. Otherwise the experience was really good."* [G19-P1]. Two groups indicated they would have preferred the robot to be more involved in the discussion and "*give her opinion especially when [we] are stuck*". Some groups also stated issues with the voice recognition and that sometimes they had to repeat things for the robot to understand. They said they did not find it annoying but time consuming.

7 Discussion and Conclusion

Individuals who participated in hiring meetings facilitated by the robot felt positively about the robot and experience, based on self-reported attitudes (Working Alliance, assessments of the robot's knowledgeability and friendliness). Participants in the ROBOT condition were also more satisfied with their team and the decision-making activity and expressed a significantly greater increase in decision confidence compared to those in the CONTROL condition.

The robot had a definite impact on group behavior, with significant differences in time spent in various meeting activities, adherence to the agenda, and whether they performed certain recommended task steps at all. In post-session interviews, all groups said they would use the robot facilitation system if they could and would recommend it to others. Most (10 of 11) groups volunteered that the robot helped balance participation in the meeting. Participants also felt the robot was effective at keeping the meeting structured and efficient.

Participants did rate the tablet application alone (CONTROL) significantly higher on overall satisfaction compared to the full facilitation system with the robot and tablet (ROBOT). There are several possible reasons for this. First, participants may have had very different frames of reference for these evaluations: tablet applications are likely very familiar to most participants and participants in the CONTROL group may have interpreted the overall satisfaction question as pertaining only to the design of the tablet application GUI, whereas participants in the ROBOT group were focused more on evaluation of their experience with the robot, possibly comparing it to human-human interaction as a point of reference. Second, the robot did have issues with response latency and speech recognition errors, which may have influenced overall satisfaction ratings. Finally, it may be that simply because the robot facilitator was effective at directing group members, they enjoyed the experience less even though it resulted in a more efficient and productive meeting, and we may have seen similar low ratings for a human facilitator.

7.1 Limitations

There are several important limitations to our evaluation study, beyond the small number of participants and the biased convenience sample of mostly college students who participated. First, we used the minimal group size of two participants: although the system and architecture are readily scalable, it remains to be seen how the robot facilitator will fare in larger groups. Our participants were mostly strangers to each other in an artificially-constructed setting, so were likely on their "best behavior". Real meetings, in which participants have histories of conflict and many agendas, are much more likely to have serious imbalances in participation that would be a better test for the robot. Our participants were also peers, and our architecture may need to be adapted for meetings in which hierarchy (e.g., the presence of a manager) influences the social dynamics. Last but not least, the presence of the robot in the robot condition can result in different group sizes compared to the control condition. Although we controlled the robot's influence on the decision making by limiting it's involvement to be only a process facilitator, the presence of another conversational partner in the room may affect the other interactions.

7.2 Future Work

There are many important directions for future work that are enabled by our system. Handling a wider range of social dynamics, including various kinds of conflict among participants, is an important area of future research. Simply identifying and classifying conflict is an important first step, but the development and evaluation of intervention strategies are equally important. Extending our framework for a wider range of meeting types, with heterogeneous participants (managers, customers, etc.), represents a large number of opportunities to study and facilitate a number of interesting and important social dynamics.

REFERENCES

[1] 55 Million: A Fresh Look at the Number, Effectiveness, and Cost of Meetings in the U.S.: *https://blog.lucidmeetings.com/blog/fresh-look-number-effectiveness-cost-meetings-in-us*. Accessed: 2019-05-13.

[2] Al Moubayed, S. et al. 2012. Furhat: A Back-Projected Human-Like Robot Head for Multiparty Human-Machine Interaction. *Cognitive Behavioural Systems* (2012), 114–130.

[3] Baron, R.S. 2005. So Right It's Wrong: Groupthink and the Ubiquitous Nature of Polarized Group Decision Making. *Advances in experimental social psychology, Vol. 37*. Elsevier Academic Press. 219–253.

[4] Bevacqua, E. et al. 2010. Multimodal Backchannels for Embodied Conversational Agents. *Intelligent Virtual Agents* (2010), 194–200.

[5] Bhattacharya, I. et al. 2018. A Multimodal-Sensor-Enabled Room for Unobtrusive Group Meeting Analysis. *Proceedings of the 20th ACM International Conference on Multimodal Interaction* (New York, NY, USA, 2018), 347–355.

[6] Bohus, D. and Horvitz, E. 2010. Facilitating Multiparty Dialog with Gaze, Gesture, and Speech. *International Conference on Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction* (New York, NY, USA, 2010), 5:1–5:8.

[7] Bohus, D. and Horvitz, E. 2011. Multiparty Turn Taking in Situated Dialog: Study, Lessons, and Directions. *Proceedings of the SIGDIAL 2011 Conference* (Stroudsburg, PA, USA, 2011), 98–109.

[8] Bostrom, R.P. et al. Group Facilitation and Group Support Systems. 23.

[9] Callaway, M.R. et al. 1985. Effects of dominance on group decision making: Toward a stress-reduction explanation of groupthink. *Journal of Personality and Social Psychology*. 49, 4 (1985), 949–952. DOI:https://doi.org/10.1037/0022-3514.49.4.949.

[10] Cassell, J. et al. 2000. *Embodied Conversational Agents*. MIT Press.

[11] Chen, H. et al. 2004. Intelligent agents meet the semantic Web in smart spaces. *IEEE Internet Computing*. 8, 6 (Nov. 2004), 69–79. DOI:https://doi.org/10.1109/MIC.2004.66.

[12] DeSanctis, G. and Brent Gallupe, R. 1987. A Foundation For The Study Of Group Decision Support Systems. *Management Science*. 33, (May 1987), 589–609. DOI:https://doi.org/10.1287/mnsc.33.5.589.

[13] Diftler, M.A. et al. 2011. Robonaut 2 - The first humanoid robot in space. *2011 IEEE International Conference on Robotics and Automation* (May 2011), 2178–2183.

[14] Duysburgh, P. et al. 2014. Exploring the Social-technological Gap in Telesurgery: Collaboration Within Distributed or Teams. *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing* (New York, NY, USA, 2014), 1537–1548.

[15] Farnham, S. et al. 2000. Structured Online Interactions: Improving the Decision-making of Small Discussion Groups. *Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work* (New York, NY, USA, 2000), 299–308.

[16] Groom, V. and Nass, C. 2007. Can robots be teammates? Benchmarks in human-robot teams. *Interaction Studies*. 8, (Oct. 2007), 483–500. DOI:https://doi.org/10.1075/is.8.3.10gro.

[17] Hill, G.W. 1982. Group versus individual performance: Are N + 1 heads better than one? (1982).

[18] Horvath, A.O. and Greenberg, L.S. 1989. Development and validation of the Working Alliance Inventory. *Journal of Counseling Psychology*. 36, 2 (1989), 223–233. DOI:https://doi.org/10.1037/0022-0167.36.2.223.

[19] How Much Time Do We Spend in Meetings? (Hint: It's Scary): 2014. *https://www.themuse.com/advice/how-much-time-do-we-spend-in-meetings-hint-its-scary*. Accessed: 2019-05-13.

[20] Ishii, R. et al. 2015. Multimodal Fusion Using Respiration and Gaze for Predicting Next Speaker in Multi-Party Meetings. *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction* (New York, NY, USA, 2015), 99–106.

[21] Jehn, K.A. and Mannix, E.A. 2001. The Dynamic Nature of Conflict: A Longitudinal Study of Intragroup Conflict and Group Performance. *The Academy of Management Journal*. 44, 2 (2001), 238–251. DOI:https://doi.org/10.2307/3069453.

[22] Jung, M.F. 2017. Affective Grounding in Human-Robot Interaction. *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction* (New York, NY, USA, 2017), 263–273.

[23] Jung, M.F. et al. 2015. Using Robots to Moderate Team Conflict: The Case of Repairing Violations. *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction* (New York, NY, USA, 2015), 229–236.

[24] Karau, S.J. and Williams, K.D. 1993. Social loafing: A meta-analytic review and theoretical integration. *Journal of Personality and Social Psychology*. 65, 4 (1993), 681–706. DOI:https://doi.org/10.1037/0022-3514.65.4.681.

[25] Kim, T. et al. 2008. Meeting Mediator: Enhancing Group Collaborationusing Sociometric Feedback. *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work* (New York, NY, USA, 2008), 457–466.

[26] Latané, B. 1981. The psychology of social impact. *American Psychologist*. 36, 4 (1981), 343–356. DOI:https://doi.org/10.1037/0003-066X.36.4.343.

[27] Lee, T. et al. 2007. Building an Interactive Caring Agent for Students in Computer-based Learning Environments. *Seventh IEEE International Conference on Advanced Learning Technologies (ICALT 2007)* (Jul. 2007), 300–304.

[28] Matsuyama, Y. et al. 2015. Four-participant group conversation: A facilitation robot controlling engagement density as the fourth participant. *Computer Speech & Language*. 33, (2015), 1–24. DOI:https://doi.org/10.1016/j.csl.2014.12.001.

[29] McMillan, S.S. et al. 2016. How to use the nominal group and Delphi techniques. *International Journal of Clinical Pharmacy*. 38, (2016), 655–662. DOI:https://doi.org/10.1007/s11096-016-0257-x.

[30] Miranda, S.M. and Bostrom, R.P. 1997. Meeting facilitation: process versus content interventions. *Proceedings of the Thirtieth Hawaii International Conference on System Sciences* (Jan. 1997), 124–133 vol.2.

[31] Niederman, F. and Volkema, R.J. 1999. The Effects of Facilitator Characteristics on Meeting Preparation, Set Up, and Implementation. *Small Group Research*. 30, 3 (Jun. 1999), 330–360. DOI:https://doi.org/10.1177/104649649903000304.

[32] Novick, D. and Gris, I. 2014. Building Rapport between Human and ECA: A Pilot Study. *Human-Computer Interaction. Advanced Interaction Modalities and Techniques* (2014), 472–480.

[33] Nunamaker, J.F. et al. 1996. Lessons from a Dozen Years of Group Support Systems Research: A Discussion of Lab and Field Findings. *J. of Management Information Systems*. 13, (1996), 163–207. DOI:https://doi.org/10.1080/07421222.1996.11518138.

[34] Romano, N.C. and Nunamaker, J.F. 2001. Meeting analysis: findings from research and practice. *Proceedings of the 34th Annual Hawaii International Conference on System Sciences* (Jan. 2001), 13.

[35] Schuman, S.P. 1996. The Role of Facilitation in Collaborative Groups. *Creating Collaborative Advantage*. SAGE Publications Ltd. 127–140.

[36] Shamekhi, A. et al. 2018. Face Value? Exploring the Effects of Embodiment for a Group Facilitation Agent. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2018), 391:1–391:13.

[37] Short, E. and Mataric, M.J. 2017. Robot moderation of a collaborative game: Towards socially assistive robotics in group interactions. *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)* (Aug. 2017), 385–390.

[38] Skantze, G. et al. 2015. Exploring Turn-taking Cues in Multi-party Human-robot Discussions About Objects. *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction* (New York, NY, USA, 2015), 67–74.

[39] Skantze, G. and Al Moubayed, S. 2012. IrisTK: A statechart-based toolkit for multi-party face-to-face interaction. (Sep. 2012).

[40] Tennent, H. et al. 2019. Micbot: A Peripheral Robotic Object to Shape Conversational Dynamics and Team Performance. *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (Mar. 2019), 133–142.
[41] Viller, S. 1991. The Group Facilitator: A CSCW Perspective. *Proceedings of the Second European Conference on Computer-Supported Cooperative Work ECSCW '91*. L. Bannon et al., eds. Springer Netherlands. 81–95.
[42] Waibel, A. et al. 2003. SMaRT: the Smart Meeting Room Task at ISL. *2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03).* (Apr. 2003), IV–752.
[43] Westley, F. and Waters, J.A. 1988. Group Facilitation Skills for Managers. *Management Education and Development*. 19, 2 (Jun. 1988), 134–143. DOI:https://doi.org/10.1177/135050768801900207.