Perception of Humanoid Social Mediators in Two-Person Dialogs

Yasir Tahir
Institute for Media Innovation
Nanyang Technological
University (NTU),Singapore
yasir001@e.ntu.edu.sg

Umer Rasheed
School of EEE
Nanyang Technological
University (NTU),Singapore
umer1@e.ntu.edu.sg

Shoko Dauwels
CIRCQL
Nanyang Technological
University (NTU),Singapore
sdauwels@ntu.edu.sg

Justin Dauwels
School of EEE
Nanyang Technological
University (NTU),Singapore
jdauwels@ntu.edu.sg

ABSTRACT

This paper presents a humanoid robot (Nao) that provides real-time sociofeedback to participants taking part in two-person dialogs. The sociofeedback system quantifies speech mannerism and social behavior of participants in an ongoing conversation, determines whether feedback is required, and delivers feedback through Nao. For example, Nao alarms the speaker(s) when the voice is too high or too low, or when the conversation is not proceeding well due to disagreements or numerous interruptions. In this study, participants are asked to participate in two-person conversations with the Nao robot as mediator. They then assess the received sociofeedback with respect to various aspects, including its content, appropriateness, and timing. Participants also evaluate their overall perception of Nao as social mediator via the Godspeed questionnaire. The results indicate that Nao can be effectively used to provide sociofeedback in discussions. The results of this research may be used in developing natural human-robot interaction for social mediators in a multi-party dialog system.

Categories and Subject Descriptors

H.5.2 [**Information Interfaces and Presentation**]: Miscellaneous; D.2.8 [**Robotics**]: Commercial Robots and Applications; J.4 [**Computer Applications**]: Social and Behavioural Sciences

General Terms

Human Factors, Design, Experimentation

Keywords

Real-Time Sociofeedback, Humanoids, Human Robot Interaction, Social Mediator.

1. INTRODUCTION

One of the key objectives of research and development in robotics is to come up with various robots than can assist humans in everyday domestic environments. Nowadays, robots are increasingly being viewed as social entities to be integrated in our daily lives. Socially interactive robots are used to communicate, express and perceive emotions, maintain social relationships, interpret natural cues, and develop social competencies [1, 2]. Prominent application scenarios for such robots are manifold and span from shopping robots [3] and tour guides [4] to home assistance and care [5, 6] etc.

With increasing demand of robots for domestic environments, research on human-robot interaction (HRI) has gained more and more importance. In order to enhance human-robot interaction, the need for integration of social intelligence in such robots has become a necessity [7, 8, 9]. Socially intelligent robots can effectively engage with humans and maintain a natural interaction with them over extended periods of time.

Understanding of human behavior is a necessary requirement for allowing a robot to behave in a socially intelligent manner [10]. If a robot can understand the behavior of humans with whom it is interacting, then it can respond accordingly. HRI in multi-party dialogs [11] can be greatly improved if the robots are able to interpret the human behavior to some extent. Human behavior involves various patterns of actions and activities, attitudes, affective states, social signals, semantic descriptions and contextual properties [12]. A promising approach for human behavior understanding is to apply pattern recognition and automatically deduce various aspects of human behavior from different kinds of recordings and measurements, e.g., audio and video recordings [13].

In [14], we presented a novel approach towards comprehensive real-time analysis of speech mannerism and social behavior. We performed non-verbal speech analysis to analyze human behavior. Non-verbal speech metrics are a direct manifestation of human behavior, and play a vital role for the meetings to be pleasant, productive, and efficient [15].

By considering these low-level speech metrics, we quantified speech mannerisms and sociometrics, i.e., interest, agreement and dominance of the speakers. We collected a diverse speech corpus of two-person face-to-face conversations; it allowed us to train machine learning algorithms for reliable 5-level classification of the sociometrics with speech metrics as input features. The combined metrics for speech mannerism and social behavior provided a clear picture of human behavior in dialogs. Such information could be used to provide appropriate feedback via Nao in real-time.

In [16], we conducted a preliminary user study to investigate how sociofeedback could be provided via a humanoid robot, Nao. It is widely accepted that the combination of modalities and capabilities improves human-robot interaction. We adopted a very simple way to evaluate this by systematically varying each modality. We conducted a study by providing users with sociofeedback in open-loop conditions i.e. the participants of the survey needed to assess six basic feedback messages delivered by Nao, without actually participating in a conversation. The participants were then asked to assess sociofeedback messages with only audio, and later a combination of audio and gestures. The user study confirmed the hypothesis that combining two modalities, i.e. audio and gestures, helps the participants to identify the sociofeedback messages in a much better way.

In this paper, we wish to extend our work and integrate our sociofeedback system with Nao to investigate the perception of humanoids as social mediators in two-person dialogs. This paper presents the following contributions and novelties:

- We integrate a real-time sociofeedback system that analyzes nonverbal speech metrics to assess the social states of participants in a two-person conversation with a humanoid robot, Nao. The robot uses this information and effectively provides an appropriate feedback in real-time. Currently, we limit ourselves to six social states only.

- We conducted a user study with 20 participants (17 males, 3 females). Each participant received sociofeedback via Nao for all six social states. Participants were then asked to evaluate several aspects of sociofeedback e.g. whether they *agree* with the feedback, whether they *like* the feedback, whether they feel they received the feedback *timely*. We asked them to provide assessment of each feedback they received during the conversation.

- We also investigated the overall experience of users about Nao. The participants were asked to rate the anthropomorphism, animacy, likability, perceived intelligence and perceived safety by means of Godspeed questionnaire [17].

The paper is structured as follows. In Section 2, we describe related work. In Section 3, we present a brief overview of the sociofeedback system, where the Nao robot provides feedback to participants of two-person dialogs. In Section 4, we explain how we designed experiments to participants assess the sociofeedback delivered by Nao robot. In Section 5, we present our results, and in Section 6, we offer concluding remarks and suggest several topics for future research.

2. RELATED WORK

In this section, we briefly discuss related work on socially aware robotic systems, their applications and relevant user studies to assess human-robot interaction. In the recent past, many social robots have been designed for real world interactions e.g. Kismet [18], Mel [19], Pearl [20], Robovie [21], and Robota [22], and Paro [23]. Nowadays, social robots are successfully helping children in their social, emotional and communication deficits. They create interesting, appealing, and meaningful interplay situations that compel children to interact with them. One of the emerging applications of social robotics is the therapy of children with autism [24, 25, 26]. The roles and benefits of socially aware robots for therapy of children with autism is reviewed in [27]. Similarly, social robots are being actively deployed in nursing homes for assistance of the elderly. Those studies typically investigate what different social functions the device can play in the living environment of the elderly, as well as how social functions can facilitate actual usage of the device [23].

Apart from that, many application-centric social robots are being deployed in domestic environments where the goal is to interact with humans as naturally as possible. HCI Institute has developed an advisor robot that traces people's mental mode from a robot's physical attributes [28]. Similarly, iCAT has been deployed in [29] to investigate dynamic multi-party social interaction with a robot agent. CALO meeting assistant [30] provides for distributed meeting capture, annotation, automatic transcription and semantic analysis of multiparty meetings. Similarly, Furhat [31] has been designed to facilitate multimodal multiparty human-machine dialogues.

In order to validate the performance of human-robot interaction, many user studies have been conducted to assess how humans perceive robots in their specific roles. Such studies rate the human-robot interaction with respect to likability, perceived safety, anthropomorphism, animacy etc. For example, The work presented in [32] investigated how humans perceive affect from robot motion. The work in [33] was carried out to prove that humans do perceive different affect by observing different motions of the robot. Similarly, in [34] studies have been carried out to see if humans can identify emotions expressed by a humanoid using gestures. In [35] Nao narrated a three minute story to a group of participants. The study investigated the effect of gazing and gestures on the persuasion of the robot. In [36], experiments were carried out to understand whether a robot can effectively modify its speech according to the speaker's behavior.

In this paper, our focus is on a specific application. We are using a humanoid, Nao, as social mediator to facilitate dyadic conversations. Therefore, we conducted a study to investigate, in detail, different aspects of human-robot interaction when Nao provides a real-time sociofeedback to participants in a two-person face-to-face dialog.

3. SYSTEM OVERVIEW

This section is structured as follows. First, we explain the hardware setup for audio recording of conversations. Next, we briefly describe the extraction of nonverbal speech cues. Then, we explain how we infer social states from those cues. Finally, we illustrate how Nao uses social states to provide
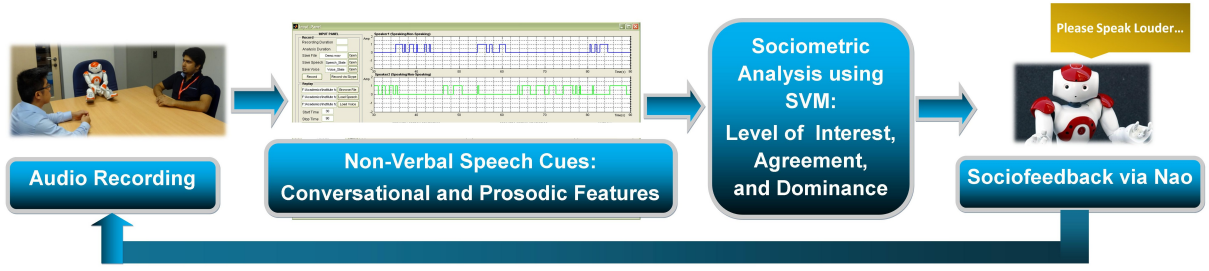
Figure 1: System Overview. The system records audio data, computes several conversational and prosodic features, and from those features, determines levels of interest, agreement, and dominance via support vector machines (SVM).

real-time sociofeedback. The overall system is illustrated in Fig.1.

3.1 Sensing and Recording

We adopt easy-to-use portable equipment for recording conversations; it consists of lapel microphones for each of the two speakers and an audio H4N recorder that allows multiple microphones to be interfaced with the laptop. The audio data is recorded in brief consecutive segments as a 2-channel audio .wave file and sent to the laptop (2GHz dual-core processor and 2GB RAM) running MATLAB.

3.2 Extraction of Non-Verbal Cues

We consider two types of low-level speech metrics: conversational and prosody related cues. The conversational cues account for *who* is speaking, *when* and *how* much, while the prosodic cues quantify *how* people talk during their conversations. We compute the following conversational cues: the number of natural turns, speaking percentage, mutual silence percentage, turn duration, natural interjections, speaking interjections, interruptions, failed interruptions, speaking rate and response time. Fig.2 shows an illustration of audio cues. Speaking and non-speaking regions are shown as black and white respectively. We consider the following prosodic cues: amplitude, larynx frequency (F0), formants(F1, F2, F3), and mel-frequency cepstral coefficients (MFCCs); those cues are extracted from 30ms segments at a fixed interval of 10ms. Those cues fluctuate rapidly in time. Therefore, we compute various statistics of those cues over a time period of several seconds, including minimum, maximum, mean and entropy to infer speaking mannerisms.
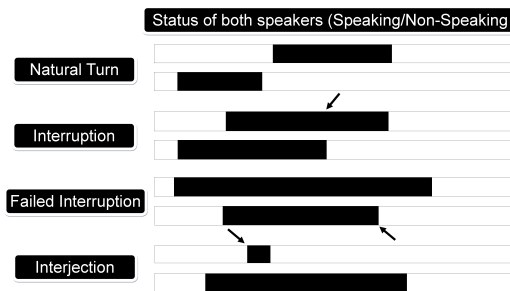
Figure 2: Illustration of turn-taking, interruption, failed interruption, and interjection. Those conversational cues are derived from the binary speaking status (speaking vs. non-speaking).

3.3 Social State Estimation

Once the speech cues are calculated, they are fed to machine learning algorithms such as Support Vector Machines (SVMs) to deduce social state of participants. Speaking mannerisms are quantitatively assessed by low-level speech metrics such as volume, rate, and pitch of speech. The social behavior is quantified by sociometrics including level of interest, agreement, and dominance. Together, they provide a comprehensive picture of the social state of participants in dialogs. On a 2GB dual-core processor with 2GB RAM, It takes approximately 5 seconds to train each SVM, about 5-10 seconds to perform speech detection and compute speech cues from 2-3 min dialogs, and less than a second to perform multi-class classification by SVM, yielding the levels of interest, agreement, and dominance. Therefore, on that computer platform, the total time required for inferring those social indicators from a 2-3 min dialog is about 5-10 seconds, allowing us to perform such analysis in real-time settings with limited delay.

3.4 Feedback Generation via Nao

Once the social state is estimated, appropriate feedback is generated by Nao. The behavior of Nao is programmed via Choregraphe. The Nao robot incorporates inertial sensors, force sensitive resistors, Hall effect sensors, infrared and sonar receivers coupled with its axes that give it 25 degrees of freedom. This multitude of sensors and actuators equips the robot with high level of stability and fluidity in its movements. However, in our experiments we only used very basic movements to simulate gestures along with text to speech generation in order to deliver the audio message. The time taken by Nao to deliver the audio message along with gestures is approximately 3 to 4 seconds. Table 1 provides an overview of the six feedback messages considered in this study.
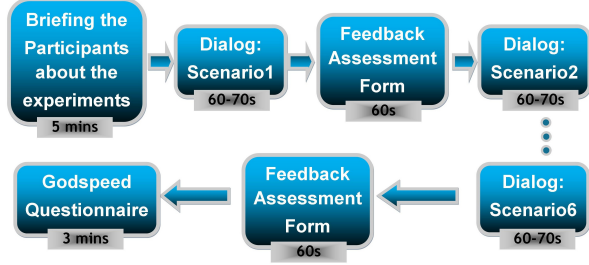
4. EXPERIMENTS

In this section, we first explain the experimental procedure. Next, we discuss how we assessed the user experience.

4.1 Experimental Procedure

The aim of the experiment is to investigate whether Nao can interact as a social mediator in a two-person dialog. We invited participants to have a scenario-based conversation. In each scenario, participants were asked to behave in a certain way, e.g., talk too loud, talk too much, or interrupt frequently. Each experiment contained six scenarios (see Fig. 3). In order to facilitate the scenario-based conversations, we asked the participants to follow scripted conver-

sations. For each conversation, we invited one subject at a time and asked him/her to act as a protagonist of the conversation. The role of the other participant remained neutral and supportive; the second participant was responsible for maintaining the flow of conversation. Each conversation lasted about 60 to 70s, and was analyzed in real-time by the sociofeedback system described in Section 3 (see also Fig. 1). Table 1 illustrates the sociofeedback generated by Nao along with the content of the message for each of the six scenarios. The experiment is conducted as follows:

Figure 3: Different components of the experimental procedure. The experiments lasts about 20 minutes, with estimated duration of each component as indicated.

- First, we setup the recording system properly.

- The two speakers sit about 1.5m apart so that each microphone only records the voice of the respective speaker, and there is no interference from the other speaker.

- We attach the lapel microphones to the speakers in proper manner, in order to obtain a high-quality recording.

- The participant (protagonist) is briefed about the experiment.

- The (protagonist) participant and the (support) participant have six conversations, following six different scenarios. Each conversation is about one minute in duration.

- Nao robot gives feedback after each conversation.

- The (protagonist) participant fills a questionnaire after each conversation, in order to rate his/her opinion about the feedback delivered by the robot.

- At the end of the experiment, the (protagonist) participant complete the Godspeed questionnaire in order to rate the Nao robot in the role of a social mediator.

The experiment included 20 participants in total (17 males and 3 females) with an average age of 23 and standard deviation of 2.42. Each experiment lasted about 20 minutes.

4.2 Assessment of Social Mediator

At the end of each conversation, the (protagonist) participant filled an assessment form about the received feedback. The questions concern different aspects of the feedback, including the content of feedback, likability, and timing (see

Gestures	Description
	Normal: "Good, carry on". If a good and normal conversation is going on Nao provides this feedback.
	Uninterested: "You both seem uninterested". Nao will invite the speakers to contribute more to the discussion, when both of the speakers have not been speaking for a period of time.
	Slow down: "You are talking a lot". Nao will ask the speaker to slow down when he/she is speaking too much.
	Calm down: "Please calm down". Nao will ask the speaker to calm down if the speaker is being too aggressive.
	Speak louder: "I am sorry, I cannot hear you". When one or both of the speakers are speaking too softly, Nao will ask them to increase their volume.
	Too noisy: "Please lower your volume". When the speakers are being too loud, Nao will give feedback about the noise.

Table 1: This table shows the gesture in one column and their corresponding description in the second.

	Question
Q1	Did you notice when the socio-feedback system was addressing you?
Q2	Did you notice when the socio-feedback system was addressing others?
Q3	Was the timing of socio-feedback appropriate?
Q4	Did the socio-feedback system interrupt the conversation?
Q5	Was the interaction natural?
Q6	Did you understand the message given by the socio-feedback?
Q7	Did you agree with the given feedback?
Q8	Did you enjoy using the socio-feedback system?

Table 2: Questions of the assessment form.

Table 2). At the end of all six conversations, the (protagonist) participant rated his/her experience of Nao as social mediator via the Godspeed questionnaire. In order to keep the assessments consistent, we adopted a 5-likert scale for both questionnaires.
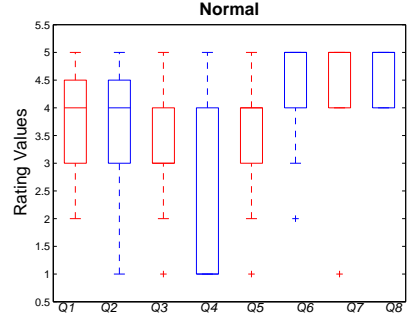
5. RESULTS AND DISCUSSIONS

In this section, we present results of the assessments made by participants. Fig.4 shows the box plots of user rating values for each question; each subfigure shows the rating values for each feedback. Rating value of 1 means the minimum score and 5 means maximum.
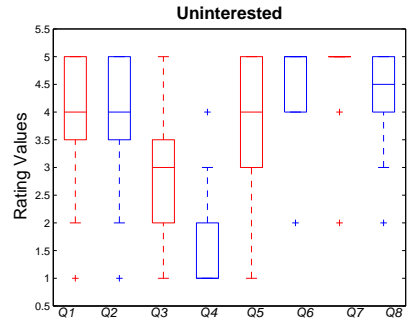
It can be seen from Fig.4 that sociofeedback via Nao received high ratings in most cases. $Q1$ and $Q2$ asked the participants if they could tell when Nao was addressing them or the other speaker. The high values for all the cases show that the participants could distinguish among feedbacks meant for them and feedbacks meant for the other speaker. In $Q3$, we asked participants about the timing of the feedback. Although, most participants stated that Nao gave feedback timely, there is still room for improvement. Ratings of $Q4$ suggests that participants at times felt they were interrupted by Nao. The timing can be improved by waiting for the speaker to stop his/her sentence or by getting attention of the speaker using some gesture and then delivering the feedback message. Furthermore, in $Q5$ and $Q6$ high ratings reveal that interaction between Nao and the participants was fairly natural and Nao spoke with clarity. In $Q7$, we asked if the participants agreed with the feedback they received. The rating for this question is really high showing participants' agreement with the feedback. Similarly, high ratings in $Q8$ confirm that participants like the feedback from Nao. The average ratings for each question ($Q1-Q8$) can be seen in Table 3. Each column shows average rating values of assessment questions for different scenarios.

At the end of the experiment the participants were asked to fill a god speed questionnaire. The purpose was to get user opinion about the robot in the role of a social mediator. Table 4 shows the user rating value averages for each of the god speed criteria.
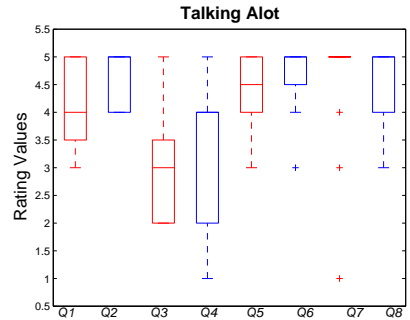
The Godspeed questionnaire contains a collection of measures for evaluating a social robot, including anthropomorphism (similarity to human form), animacy (life likeness),
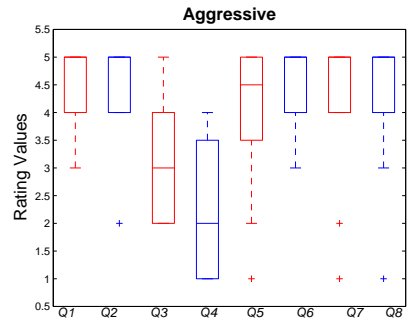

(a) Ratings for "Normal" scenario.

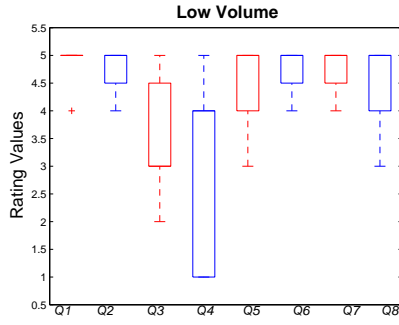(b) Ratings for "Uninterested" scenario.

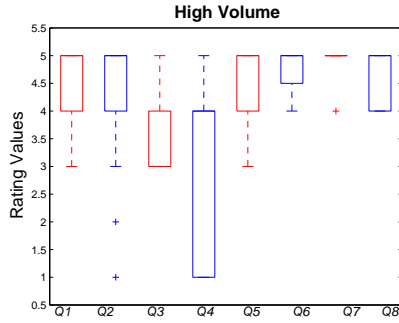(c) Ratings for "Talking a lot" scenario.

(d) Ratings for "Aggressive" scenario.

(e) Ratings for "Low Volume" scenario.

(f) Ratings for "High Volume" scenario.

Figure 4: Box plots of participant's ratings for each of the six scenarios.

Characteristics	Average Values
Anthropomorphism	
Machine/human like	4
Moving rigidly/elegantly	3
Animacy	
Mechanical/organic	4
Inert/interactive	4
Likability	
Dislike/like	5
Unfriendly/friendly	4
Perceived Intelligence	
Ignorant/knowledgeable	4
Unintelligent/intelligent	4
Perceived Safety	
Calm/agitated	3
Quiescent/surprised	3

Table 4: Average values of Godspeed questionnaire.

the size of the robot. Nao is a small robot (2 feet); when people interact with Nao while they are standing, the safety value is usually very good [37]. In our case, Nao is seated very close to the participants (see Fig.1), which may explain why the safety value is high in our experiments.

We also asked whether they would like to receive sociofeedback or not. Out of 20 participants, 19 responded in favor of receiving sociofeedback. We also inquired about their preferred platform for sociofeedback. We asked them to choose between android application, Skype VoIP application, smart glasses, Nao robot, and virtual human.
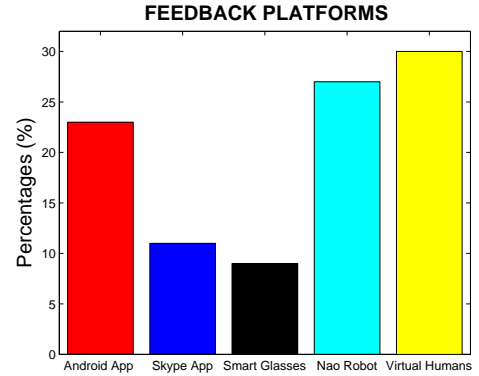
Figure 5: Feedback platforms shown with respective percentage of participants' choice.

The results can be seen in Fig.5. 30% of the participants chose virtual humans as their preferred feedback delivery platform. Nao robot was chosen 27%, android application 22%, skype application 12% and smart glasses were chosen by 9% of the participants.

We also asked the participants to leave any suggestion that they might have about the experiment. Mostly we got good comments by the participants. Some participants suggested improvements about the feedback.These suggestions were about feedback timing and about making interaction more natural. We intend to work on further improvements of our setup in light of these suggestions.

likeability (personal likeness of the participant), perceived intelligence, and perceived safety of a robot. The scores for likeability are the highest. In other words, the participants seemed to like Nao, and perceived it as friendly. Anthropomorphism also has good ratings. The robot is rated highly human-like but the motions of the robot can be improved to make it more elegant. The animacy of Nao is also rated high by the participants. Thus Nao was considered as highly interactive. Likewise, the participants perceived the robot as knowledgeable and intelligent. However, Nao received moderate ratings for its perceived safety suggesting there is a room for improvement to make participants more comfortable in the presence of Nao. Perceived safety is related to

| Question	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
Normal	4	4	3	2	4	5	4	5
Uninterested	4	4	3	1	4	4	5	4
Talking Alot	4	5	3	3	4	5	5	5
Aggressive	5	5	3	2	4	5	4	5
Low Volume	5	5	4	3	4	5	5	5
High Volume	5	4	4	3	5	5	5	5
Total Average	4.5	4.5	3.3	2.3	4.2	4.8	4.7	4.8

Table 3: Average rating values of each assessment question. Each column shows the rating values for each question where each row represents social scenario.

6. CONCLUSION AND FUTURE WORK

In this paper, we presented a user study to assess how Nao is perceived by people in the role of social mediator in two-person dialogs. In this setting, Nao robot monitors an ongoing conversation, and provides feedback to the participants regarding speaking mannerism and other behaviors. We aimed to investigate how effectively the feedback from the humanoid robot is perceived by humans. To this end, we conducted a survey with 20 participants, where six different feedback messages delivered by the robot were evaluated. The participants were part of a scenario-based dialog. Participants were engaged in a discussion, and the feedback messages were delivered by Nao to the participants in real-time. They followed a scripted conversation and enacted a certain scenario. Sociofeedback system analysed the conversation and based on the analysis Nao provided the feedback. Participants provided us with an assessment after receiving each feedback. They assessed the content, timing, relevance and their liking of the feedback. At the end, each participant completed a Godspeed questionnaire to assess the robot in the role of social mediator. We observed that the participants clearly liked receiving feedback from Nao robot. The agreement scores are very high showing that participnats agreed with the provided feedback. There is room for improvement in the timing of the feedback. We will try to improve the timing in future experiments. The results of the Godspeed questionnaire suggest that the participants really liked the humanoid robot Nao in this experiment. The ratings for all Godspeed criteria are high. Only with regard to perceived safety, the evaluation was only mildly positive; this may be explained by the fact that the robot was sitting near the participants. However, the average rating is still acceptable, and this issue may not be very critical.

Overall, this study suggests that sociofeedback by the Nao robot can be accurately identified and is appreciated by participants. In future, we aim to further improve the social state estimation. To this end, we will collect multi-modal (audio and video) datasets for training the system. Secondly, we will attempt to scale the proposed system to multi-party dialogs. We also intend to further improve the feedback delivery of Nao and make it look more interactive, so that in future it can be a part of real world group discussions.

7. ACKNOWLEDGEMENTS

8. REFERENCES

[1] T. Fong, I. Nourbakhsh, and K. Dautenhahn, "A survey of socially interactive robots," *Robotics and autonomous systems*, vol. 42, no. 3, pp. 143–166, 2003.

[2] H. Li, J.-J. Cabibihan, and Y. K. Tan, "Towards an effective design of social robots," *International Journal of Social Robotics*, vol. 3, no. 4, pp. 333–335, 2011.

[3] T. Kanda, M. Shiomi, Z. Miyashita, H. Ishiguro, and N. Hagita, "A communication robot in a shopping mall," *IEEE Transactions on Robotics*, vol. 26, no. 5, pp. 897–913, 2010.

[4] S. Thrun, M. Bennewitz, W. Burgard, A. B. Cremers, F. Dellaert, D. Fox, D. Hahnel, C. Rosenberg, N. Roy, J. Schulte, and D. Schulz, "Minerva: A second-generation museum tour-guide robot," in *In Proceedings of IEEE International Conference on Robotics and Automation (ICRA 99)*, 1999.

[5] F. Tanaka, A. Cicourel, and J. R. Movellan, "Socialization between toddlers and robots at an early childhood education center," *Proceedings of the National Academy of Science*, vol. 104, pp. 17954–17958, Nov. 2007.

[6] B. Graf, C. Parlitz, and M. Hägele, "Robotic home assistant care-o-bot® 3 product vision and innovation platform," in *HCI (2)*, pp. 312–320, 2009.

[7] K. Williams and C. Breazeal, "A reasoning architecture for human-robot joint tasks using physics-, social-, and capability-based logic," in *IROS*, pp. 664–671, 2012.

[8] D. François, D. Polani, and K. Dautenhahn, "Towards socially adaptive robots: A novel method for real time recognition of human-robot interaction styles," in *Humanoids*, pp. 353–359, 2008.

[9] F. Papadopoulos, K. Dautenhahn, and W. C. Ho, "Exploring the use of robots as social mediators in a remote human-human collaborative communication experiment," *Paladyn*, vol. 3, no. 1, pp. 1–10, 2012.

[10] K. Dautenhahn, "Socially intelligent robots: dimensions of human–robot interaction," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 362, no. 1480, pp. 679–704, 2007.

[11] D. Bohus and E. Horvitz, "Dialog in the open world: platform and applications," in *Proceedings of the 2009 international conference on Multimodal interfaces*, pp. 31–38, ACM, 2009.

[12] A. A. Salah, T. Gevers, N. Sebe, and A. Vinciarelli, *Human Behavior Understanding: First International Workshop, HBU 2010, Istanbul, Turkey, August 22, 2010, Proceedings*, vol. 6219. Springer, 2010.

[13] A. A. Salah, J. Ruiz-del Solar, Ç. Meriçli, and P.-Y. Oudeyer, "Human behavior understanding for robotics," in *Human Behavior Understanding*, pp. 1–16, Springer, 2012.

[14] S. D. J. D. D. T. Umer Rasheed, Yasir Tahir and N. Magnenat-Thalmann, "Real-time comprehensive sociometrics for two-person dialogs," in *Human Behavior Understanding*, Springer, 2013.

[15] A. S. Pentland, *Honest signals*. MIT press, 2010.

[16] K. H. S. D. J. D. D. T. Yasir Tahir, Umer Rasheed and N. Magnenat-Thalmann, "Nao as social mediator: A user study," in *Robots in public spaces:towards multi-party, short-term, dynamic human-robot interaction*, Springer, 2013.

[17] C. Bartneck, E. Croft, and D. Kulic, "Measuring the anthropomorphism, animacy, likeability, perceived intelligence and perceived safety of robots," in *Metrics for HRI Workshop, Technical Report*, vol. 471, pp. 37–44, Citeseer, 2008.

[18] C. Breazeal and B. Scassellati, "A context-dependent attention system for a social robot," *rn*, vol. 255, p. 3, 1999.

[19] C. L. Sidner and M. Dzikovska, "A first experiment in engagement for human-robot interaction in hosting activities," in *Advances in Natural Multimodal Dialogue Systems*, pp. 55–76, Springer, 2005.

[20] M. E. Pollack, L. Brown, D. Colbry, C. Orosz, B. Peintner, S. Ramakrishnan, S. Engberg, J. T. Matthews, J. Dunbar-Jacob, C. E. McCarthy, *et al.*, "Pearl: A mobile robotic assistant for the elderly," in *AAAI workshop on automation as eldercare*, vol. 2002, pp. 85–91, 2002.

[21] H. Ishiguro, T. Ono, M. Imai, T. Maeda, T. Kanda, and R. Nakatsu, "Robovie: an interactive humanoid robot," *Industrial robot: An international journal*, vol. 28, no. 6, pp. 498–504, 2001.

[22] A. Billard, "Robota: Clever toy and educational tool," *Robotics and Autonomous Systems*, vol. 42, no. 3, pp. 259–269, 2003.

[23] C. D. Kidd, W. Taggart, and S. Turkle, "A sociable robot to encourage social interaction among the elderly," in *Robotics and Automation, 2006. ICRA 2006. Proceedings 2006 IEEE International Conference on*, pp. 3972–3976, IEEE, 2006.

[24] K. C. Welch, U. Lahiri, Z. Warren, and N. Sarkar, "An approach to the design of socially acceptable robots for children with autism spectrum disorders," *International Journal of Social Robotics*, vol. 2, no. 4, pp. 391–403, 2010.

[25] I. Fujimoto, T. Matsumoto, P. R. S. De Silva, M. Kobayashi, and M. Higashi, "Mimicking and evaluating human motion to improve the imitation skill of children with autism through a robot," *International Journal of Social Robotics*, vol. 3, no. 4, pp. 349–357, 2011.

[26] G. Schiavone, D. Formica, F. Taffoni, D. Campolo, E. Guglielmelli, and F. Keller, "Multimodal ecological technology: From childâĂŹs social behavior assessment to child-robot interaction improvement," *International Journal of Social Robotics*, vol. 3, no. 1, pp. 69–81, 2011.

[27] J.-J. Cabibihan, H. Javed, M. Ang Jr, and S. M. Aljunied, "Why robots? a survey on the roles and benefits of social robots in the therapy of children with autism," *International Journal of Social Robotics*, pp. 1–26, 2013.

[28] A. Powers and S. Kiesler, "The advisor robot: tracing people's mental model from a robot's physical attributes," in *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*, pp. 218–225, ACM, 2006.

[29] M. E. Foster, A. Gaschler, M. Giuliani, A. Isard, M. Pateraki, and R. Petrick, "Two people walk into a bar: Dynamic multi-party social interaction with a robot agent," in *Proceedings of the 14th ACM international conference on Multimodal interaction*, pp. 3–10, ACM, 2012.

[30] G. Tur, A. Stolcke, L. Voss, S. Peters, D. Hakkani-Tur, J. Dowding, B. Favre, R. Fernández, M. Frampton, M. Frandsen, *et al.*, "The calo meeting assistant system," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 18, no. 6, pp. 1601–1611, 2010.

[31] S. Al Moubayed, J. Beskow, G. Skantze, and B. Granström, "Furhat: a back-projected human-like robot head for multiparty human-machine interaction," in *Cognitive Behavioural Systems*, pp. 114–130, Springer, 2012.

[32] J. Harris and E. Sharlin, "Exploring the affect of abstract motion in social human-robot interaction," in *RO-MAN, 2011 IEEE*, pp. 441–448, IEEE, 2011.

[33] M. Saerbeck and C. Bartneck, "Perception of affect elicited by robot motion," in *Proceedings of the 5th ACM/IEEE international conference on Human-robot interaction*, pp. 53–60, IEEE Press, 2010.

[34] A. Beck, A. Hiolle, A. Mazel, and L. Cañamero, "Interpretation of emotional body language displayed by robots," in *Proceedings of the 3rd international workshop on Affective interaction in natural environments*, pp. 37–42, ACM, 2010.

[35] J. Ham, R. Bokhorst, and J. Cabibihan, "The influence of gazing and gestures of a storytelling robot on its persuasive power," in *International conference on social robotics*, 2011.

[36] A. Delaborde and L. Devillers, "Use of nonverbal speech cues in social interaction between human and robot: emotional and interactional markers," in *Proceedings of the 3rd international workshop on Affective interaction in natural environments*, pp. 75–80, ACM, 2010.

[37] K. Werner, J. Oberzaucher, and F. Werner, "Evaluation of human robot interaction factors of a socially assistive robot together with older people," in *Complex, Intelligent and Software Intensive Systems (CISIS), 2012 Sixth International Conference on*, pp. 455–460, IEEE, 2012.