

A Trial Attempt by a Museum Guide Robot to Engage and Disengage the Audience on Time

Madhumita Ghosh
Department of Intelligent Interaction Technologies
University of Tsukuba
Tsukuba, Japan
E-mail: madhumita2ghosh@gmail.com

Hideaki Kuzuoka
Department of Intelligent Interaction Technologies
University of Tsukuba
Tsukuba, Japan
E-mail: kuzuoka@iit.tsukuba.ac.jp

Abstract— Social robots may perform a variety of services like conveying information, guiding an audience, providing assistance to troubled customers and so on. In these situations, there might be a number of people who might need such assistance at the same time. Therefore, it is important to engage the audience, convey the information, and disengage the audience in time to allow the next group of people to have an opportunity to ask for help. A typical example of such a situation is a guided tour by a robot in a museum. We conducted a field experiment where we tested whether a robot can engage and disengage the audience at predetermined points in time so as to complete the tour in a predictable amount of time. With the help of social scientists, we created some verbal and non-verbal behaviors for the robot that are usually employed by professional museum guides to manage an audience. We analyzed the audience's disengagement behavior (i.e., point of disengagement) by analyzing the video data captured during the experiment. We also confirmed the audience's engagement with the exhibits by analyzing the questionnaire data from the experiment. This was our pilot trial and we experienced a few challenges. In this paper, we describe our observations as well as the challenges. Our pilot trial indicates that a combination of the behaviors that we programmed into the robot guide can disengage the audience more efficiently than when those behaviors are used separately or not used at all.

Keywords- *Human-Robot Interaction; Museum guide robot; Engagement; Social cues; Social robots*

I. INTRODUCTION

We are gradually moving towards an era when robots will be used by the common person just like computers are being used today. For the commercial success of robots, just like computers today, robots need to be useful for a variety of important purposes, in addition to being easy-to-use. When robots become popular, more people will be interested in building robotic applications, just like the interest in developing iPhone applications today. This will meet the first of the two future needs mentioned above. We can make robots easy-to-use by developing speech or gesture-based user interfaces. This is very similar to human-to-human communication. We can use elements of human-to-human communication in human-robot interaction (HRI) to make robots more user-friendly.

Another problem that arises during human-robot interaction is the problem of losing interest. Since common

people perceive robots as just a machine, therefore, they might feel that the rules of etiquette usually followed in human society need not be applied during human-robot interaction. As a result, it has been observed in our previous experiments that the audience doesn't always feel compelled to stay with the robot till the end of the interaction. There are a number of problems that arise from this situation. Firstly, when people lose interest in the interaction halfway through, the purpose of interaction is not served. For example, if the purpose of the interaction was to provide some information to the audience, then the audience only receives a part of the information. Therefore, it becomes difficult to recommend the use of robots. Secondly, if the interaction ends mid-way, then it is difficult to study the interaction as a whole and understand the key points during the interaction that had a positive impact on the audience. This also means that it will be difficult to design better robots and supporting systems for a more interesting experience. Thirdly, if common people do not find it interesting to interact with robots, then it will be difficult to make robots commercially successful in the future.

To solve these problems, we are investigating how to make robots more engaging. We chose the museum as an example of a social environment where people of all ages are likely to interact with robots. In this paper, we describe our ongoing research and experiments to understand how robots acting as museum guides can engage the audience during an interaction. Also, after studying the requirements for a guide robot in the context of a museum environment, we understand that timely disengagement is also important. Visitors wait for the curator to finish the current guided tour and come back to guide the next set of visitors through the exhibits. When a guided tour does not end within the designated amount of time, the visitors who are waiting to join in the next round of the tour have to wait for a long time. Depending on the day of the week, if there is a big crowd, then some people might give up and leave. Therefore, it is important to disengage the audience from the exhibits at the right time so that the guided tour can be completed within the predetermined amount of time.

II. BACKGROUND

In our lab, we have developed robots that have been deployed in social spaces like museums. Some aimed to study the effect of human ploys to attract attention [1], others to study the way the audience arranges itself around a robot

[2]. Social robots have also been introduced into social spaces like shopping malls [3] and train stations [4]. Even social scientists have participated in the efforts to analyze the effect of technology on people in social environments [5]. In these previous researches, some verbal and non-verbal behaviors observed in human-to-human communication have been used to make the robot seem more interesting.

In our research, we are studying how robots can engage people specifically in a museum environment. By conducting a context-specific study, we want to reveal not only those elements that all robots can use to engage the user, but we also want to reveal context-specific elements that museum guide robots can employ to engage as well as disengage the audience.

III. METHOD

A. Field Study and the Development of our System

To begin our investigation, we visited the Science Museum in Tokyo and interviewed a curator who regularly conducts guided tours. Thereafter, we conducted a field study to understand the flow of events that take place during a guided tour. We studied the gestures and the script that the curator uses for the guided tour of a particular exhibition hall in the museum. We also collaborated with social scientists who have studied the interaction between humans and machines of different kinds [6]. One of our collaborators has conducted her doctoral research on the non-verbal cues that museum curators use to manage a group during a guided tour and to engage and disengage the audience at the appropriate point in time [7]. We have used the findings from this research as the foundation for developing our museum guide robot.

Thereafter, we conducted several pilot experiments in our lab. After making several rounds of improvements, we conducted an experiment at the Tokyo Science Museum (Fig. 1). We deployed our museum guide robot to conduct a guided tour of three exhibits in an exhibition hall which displayed the historic advancements in bicycle design. Some of the bicycles displayed there were decades or even centuries old. We attempted to engage the audience in an interaction where the robot revealed an interesting history of the exhibits. We also solicited feedback from the participants through a questionnaire.

B. Experimental Setup

To engage the audience throughout the interaction, we programmed the robot to make eye-contact at TRPs (Transition Relevance Places) [8]. We also worked on the script that the robot would use to relate the interesting history of each exhibit. We worked with the curator to make the script interesting and infuse humor wherever possible.

We selected two key elements from the plethora of verbal and non-verbal gestures that museum guides use to manage the audience, as revealed by the work of our collaborators. We chose these two elements because we realized – after a comparative analysis – that these elements are relatively easier to incorporate in our robot platform Talk Torque 2. Since the robot’s movements are limited by the

number of DOFs (Degrees of Freedom) of its joints, we chose the ‘summative assessment’ (verbal gesture) and the ‘lean back gesture’ (non-verbal gesture) as tools of disengagement. Summative assessment is performed ritually at the end of the explanation of an exhibit by museum guides to disengage the audience from the current exhibit and prepare them to move on to the next exhibit. The summative assessment is usually the guide’s personal opinion about the exhibit or a reflection based on the content of the explanation just concluded. This usually signals the end of the explanation to the audience. The lean back gesture is performed towards the end of the summative assessment and involves taking a step back, away from the current exhibit, a few seconds before the summative assessment ends. This non-verbal gesture emphatically displays to the audience that it is time to move away from that exhibit. Our hypothesis is that a combination of these two elements will be able to effectively disengage the audience from the current exhibit and will help to prevent the ‘lingering effect’. The lingering effect happens when the audience members linger around the current exhibit even though the guide or the rest of the group has moved on.

We created four conditions for the experiment. In the first condition, the robot guide uses both the summative assessment (SA) and lean back (LB) gestures. In the second condition, the robot uses only the SA and in the third condition, the robot uses only the LB gesture. In the fourth condition, the robot does not use either the SA or the LB gesture.

Our hypothesis:

1. During the condition in which the robot uses both SA and LB gestures, the audience will disengage more effectively than in the conditions where only one of these elements is used.

2. There will be some lingering effect in the condition in which the robot does not use either the SA or the LB gestures.

We conducted our experiment on 16 sets of subjects on a single day at the Science Museum. Each subject set usually consisted of one to four people. The subjects signed a consent form to allow us to videotape the experiments and at the end of each guided tour, the subjects filled out a questionnaire.



Figure 1. A robot-guided tour in progress.

C. Our Approach to Analyzing the Data

In the following paragraphs we refer to the exhibit that the robot just finished explaining as the ‘current exhibit’ and the following exhibit in the tour as the ‘next exhibit’.

We analyzed the video data to ascertain the point of disengagement for each subject. We decided the point of disengagement to be the point in time when the subject looked away from the current exhibit and continued to look towards the next exhibit or towards the robot or elsewhere, without returning their gaze towards the current exhibit. We analyzed whether the subject disengaged from the current exhibit during the (a)summative assessment (SA) or the last statement (in the conditions in which SA was not used) in the explanation, (b)during the lean back gesture, (c)when the robot was turning around to prepare to move to the next exhibit, (d)during the transition from one exhibit to the next, or (e)never disengaged at all (Fig.2 to 5). To confirm whether our analysis was accurate, we recruited two independent raters. We sliced each video

from the beginning of the summative assessment/last statement to the point where the robot is in transit towards the next exhibit, into ten time slices and numbered each slice. The number was inserted as a label into the video clips. Then we asked the independent raters to judge each time slice and determine whether the subject in that time slice was engaged with the current exhibit or had finally disengaged from the current exhibit and continued to remain disengaged.

The video data for 44 subjects were analyzed. Since there were 3 exhibits in our robot-guided tour, 132 units of video data were available. We discarded 19 units of video data because it was impossible to determine the engagement/disengagement status of the subject in those videos due to poor video quality, orientation and distance from the subject, or because someone blocked the field of view of the camera. The cases where the raters could not determine the engagement/disengagement status of the subject were also discarded from our analysis. There were 39 cases where the two raters were in agreement. We applied

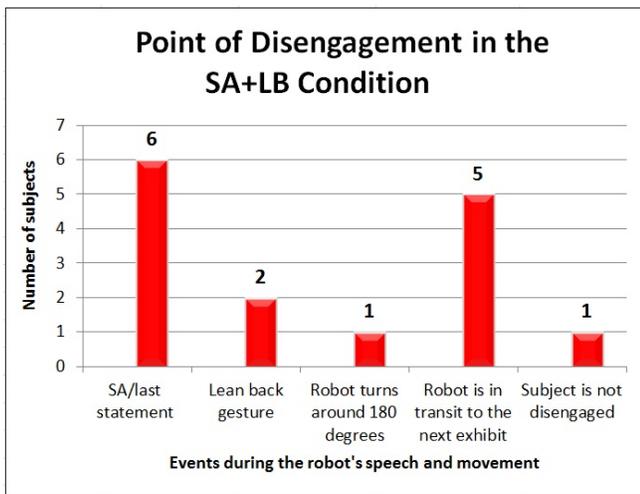


Figure 2. Point of disengagement in the summative assessment + lean back gesture condition (Graph 1).

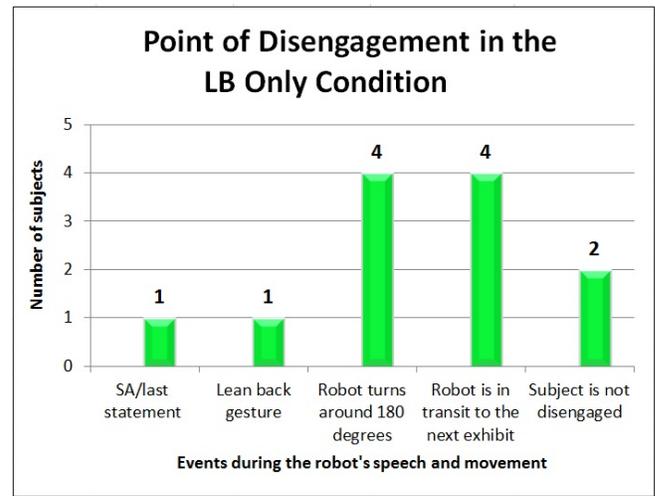


Figure 4. Point of disengagement in the lean back only condition (Graph 3).

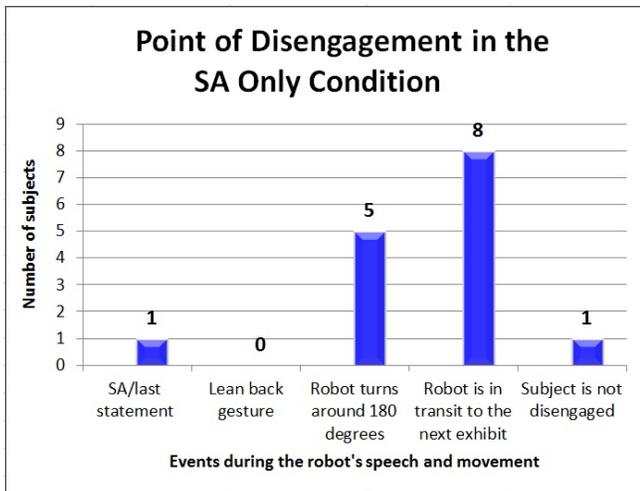


Figure 3. Point of disengagement in the summative assessment only condition (Graph 2).

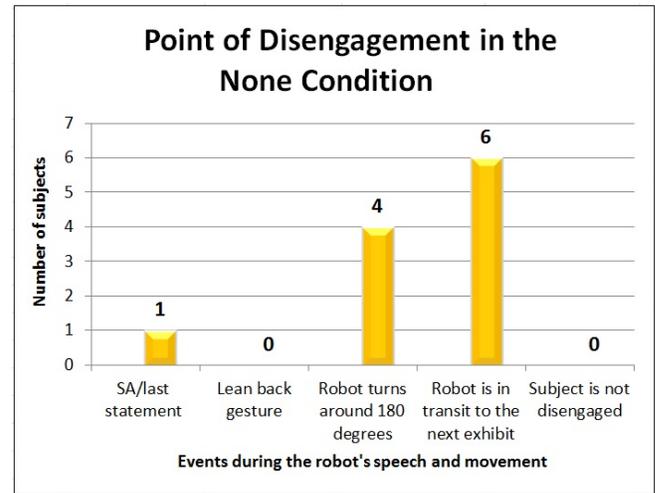


Figure 5. Point of disengagement in the None condition (Graph 4).

Cohen’s Kappa to the remaining cases and discarded 31 cases where Kappa coefficient was less than 0.7. We finally had 53 cases available for our analysis and we present the results in the next section.

IV. RESULTS

In this section, we present our findings in a set of figures. We discuss the implications of these results in the following section. Please note that in the results, a unit of interaction indicates the interaction between one subject and the robot with respect to a single exhibit.

We analyzed the video material recorded during this experiment and found that the audience was easily engaged in the interaction throughout the experiment. This was evident from the fact that the subjects made eye contact with the robot whenever the robot would look in their direction. During the explanation of the exhibits, the subjects laughed whenever there was some humorous content. Finally, all the subjects answered correctly when asked in the questionnaire about which of the exhibits had the oldest history. It proved that the subjects were listening attentively to all the explanations.

Graphs 1, 2, 3, and 4 (Fig.2 to 5) show the point of disengagement of the subjects for all the 53 units of interaction that we analyzed. We can see in Graph 1 that in majority of the cases, the subject disengaged from the current exhibit during the delivery of the summative assessment. The number of subjects that disengaged during the summative assessment and the lean back gesture taken together also form the majority of the cases analyzed for this condition over the number of cases in which the subject did not respond to the robot’s attempt to induce disengagement behavior, even though it is by a very slim margin. In Graph 2, we see that the summative assessment when delivered without the lean back gesture, fails to disengage the subject. The subject does not disengage until the robot is on its way to the next exhibit. This might immediately give us the impression that the lean back gesture is the one that is really accomplishing the target of disengaging the subject in the SA+LB condition. However, as we can see from Graph 3, that is not the case. When the lean back gesture is used by itself, it shows similar poor results as in the SA only condition. Most subjects disengage only when the robot starts turning around or moving towards the next exhibit. This indicates that a combination of the summative assessment and the lean back gesture has a strong effect on the subjects’ disengagement behavior. This is further reinforced by the results in the None condition as displayed in Graph 4. In the absence of both the disengagement-inducing behaviors, the subject disengages only when the robot moves in a very explicit manner, as opposed to the subtle movement of the lean back gesture.

The table in Fig. 6 summarizes the trends observed in the subjects’ disengagement behavior in the different conditions. As we can see from the table, overall, people disengaged mostly during the robot’s transit to the next exhibit. But, it is possible to argue that the act of transitioning to the next exhibit is the equivalent of forcing a disengagement rather than making it happen in a more subtle, covert fashion, by

using the SA and LB gestures. More importantly, if we consider that the leftmost column in the table represents events in a chronological order, then we have to note that when both the summative assessment and the lean back gesture is used, 40% of the subjects disengaged during the delivery of the summative assessment and waited for the robot to finish explaining about the current exhibit and start moving towards the next exhibit. Compared to this, in all the other conditions, the disengagement rate was below 10% in this phase (delivery of the SA/last statement). Based on these results, we can suggest that using a combination of the two disengagement-inducing behaviors can help to disengage the subject more readily. Whether such requirements exist, depend upon the context in which human-robot interaction takes place. In the context described in this paper, it can be beneficial to disengage the audience from an exhibit while summarizing the talk about the exhibit.

We did not observe any lingering behavior in the subjects during the None condition. From the results of our experiment, we find that hypothesis 1 is supported but hypothesis 2 is not supported.

V. DISCUSSION

The results presented in the previous section indicate that in the context described in this paper, a combination of summative assessment and lean back gesture can be useful as a disengagement tool. Although these results need to be verified further, they encourage us to continue our investigation into the idea of using the findings from ethnomethodological studies [9] of human-to-human interaction to develop robot behaviors for more effective human-robot interaction. Since social science provides us with a treasure trove of subtle human behaviors that are used regularly in human-to-human communication to manage and smoothen the communication process, therefore it presents a sea of possibilities for human-robot interaction researchers to test and incorporate these behaviors into robots.

Our experimental results suggest that the above disengagement tool is effective in ‘disengaging the audience early on’, rather than, ‘not using this tool means that the subject will not disengage or there will be a lingering effect’. The data presented in the table clearly indicates that it is possible to disengage the audience by simply moving on. However, the audience might find this rude and abrupt. Using a more sophisticated disengagement mechanism like the one suggested in this paper gives an impression of friendliness or politeness on the part of the robot guide.

Condition-wise Trends in Disengagement				
Events during the robot's speech and movement	SA+LB condition	SA only condition	LB only condition	None condition
SA/last statement	40.00%	6.67%	8.33%	9.09%
Lean back gesture	13.33%	0%	8.33%	0%
Robot turns around 180 degrees	6.67%	33.33%	33.33%	36.36%
Robot is in transit to the next exhibit	33.33%	53.33%	33.33%	54.54%
Subject is not disengaged	6.67%	6.67%	16.67%	0%
Total units of interaction analyzed	15	15	12	11

Figure 6. Percentage of subjects that disengaged during the different stages of the tour, in each condition.

VI. LIMITATION

Since the subjects usually participated in a group, either because they were with their families or because several people lined up to participate in the experiment at the same time and hence we chose to put people in groups of 4~6 people to prevent a big crowd in the waiting area. Most families participated in groups of 3~4 people. We will make it a part of our future work to study the difference in disengagement behavior when a person interacts with a robot one-on-one and when they are a part of a group.

VII. CONCLUSION

In this paper, we reported our very first pilot trial in a real world setting with common people as our subjects. We tested whether a robot could incorporate subtle verbal and non-verbal gestures to engage and disengage an audience in the specific context of a guided tour of exhibits in a museum. We found that the behaviors that we selected for our very first step into this investigation show promise as a disengagement tool. We feel encouraged to continue our collaboration with social scientists to test and incorporate other elements from human-to-human communication into social robots.

ACKNOWLEDGMENT

We are grateful to our collaborators, Prof. Paul Luff and Dr. Katie Ann Best, for sharing their insights about the social behavior of museum guides. We also want to convey our sincere thanks to Yūki Akamine and Deokyong Shin for helping us with the logistics to conduct this field trial and for assisting us in the fieldwork.

REFERENCES

- [1] Hideaki Kuzuoka, Karola Pitsch, Yuya Suzuki, Ikkaku Kawaguchi, Keiichi Yamazaki, Akiko Yamazaki, Yoshinori Kuno, Paul Luff, Christian Heath, "Effect of restarts and pauses on achieving a state of mutual orientation between a human and a robot," *Computer Supported Collaborative Work (CSCW'08)*, pp. 201–204, San Diego, California, USA, 2008.
- [2] Hideaki Kuzuoka, Yuya Suzuki, Jun Yamashita, Keiichi Yamazaki, "Reconfiguring spatial formation arrangement by robot body orientation," *5th ACM/IEEE international conference on Human-robot interaction*, pp. 285–292, Osaka, Japan, 2010.
- [3] Takayuki Kanda, Masahiro Shiomi, Zenta Miyashita, Hiroshi Ishiguro, Norihiro Hagita, "An affective guide robot in a shopping mall," *4th ACM/IEEE international conference on Human robot interaction*, pp. 173–180, San Diego, California, USA, 2009.
- [4] Masahiro Shiomi, Daisuke Sakamoto, Takayuki Kanda, Carlos Toshinori Ishi, Hiroshi Ishiguro, Norihiro Hagita, "A semi-autonomous communication robot - a field trial at a train station," *3rd ACM/IEEE international conference on Human robot interaction*, pp. 303–310, Amsterdam, Netherlands, 2008.
- [5] Dirk vom Lehn, Christian Heath, Jon Hindmarsh, "Exhibiting Interaction: Conduct and Collaboration in Museums and Galleries," *Symbolic Interaction*, vol. 24, May 2001, pp. 189-216, doi: 10.1525/si.2001.24.2.189.
- [6] Paul Luff, Jon Hindmarsh, Christian Heath, *Workplace Studies, Recovering Work Practice and Information System Design*, Cambridge University Press, 2000.
- [7] Katie-Anne Best, "The interactional organization of the guided tour: the work of tour guides in museums, galleries and historic houses," doctoral dissertation, King's College London, 2008.
- [8] Akiko Yamazaki, Keiichi Yamazaki, Yoshinori Kuno, Matthew Burdelski, Michie Kawashima, Hideaki Kuzuoka, "Precision timing in human-robot interaction: coordination of head movement and utterance," *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 131–139, Florence, Italy, 2008.
- [9] Harold Garfinkel, *Studies in Ethnomethodology*, Prentice-Hall, 1967.