Analysis of Verbal and Nonverbal Behaviors in Online Bibliobattle Focusing on Listener Responses

Akihiro Nagai
Future University Hakodate
Hokkaido, Japan
a-nagai@sumilab.org

Yasuyuki Sumi
Future University Hakodate
Hokkaido, Japan
sumi@fun.ac.jp

***Abstract*—This paper presents an attempt to read the success of a presentation in conveying information and gaining empathy from listener responses. The success of a presentation is strongly influenced not only by the speaker but also by the listener's reactions, such as nodding and other nonverbal behaviors and speech during the Q&A session. Therefore, we analyze the relationship between the participants' verbal and nonverbal behavior data obtained from online Bibliobattle and their voting results. Specifically, we focus on the amount of speech and nodding by the listeners, and in addition, we examine the relationship between their voting results and the results of questionnaires on empathy and comprehension. The results suggest that while some listeners expressed nodding as a signal of empathy, there were not a few who did not, and that these differences may be related to personality.**

***Index Terms*—online presentation, empathy, verbal and nonverbal behaviors, listener's response**

I. Introduction

The quality of a presentation is considered to depend mainly on the skill of the presenter, including speaking skills and the use of presentation materials. On the other hand, the reaction of the audience is also considered to affect the success or failure of the presentation. For example, nods of agreement and facial expressions that confirm understanding of the content while listening to a presentation, as well as speech during question and answer (Q&A) sessions, contribute greatly to the success of a presentation. Such opportunities are naturally fostered in face-to-face presentations, but in online presentations, which have become widespread in recent years, many audience members have their cameras and microphones turned off, so listener reactions are not conveyed to the speaker. This causes frustration for the presenter, and as a result, the quality of the presentation seems to be lowered. Therefore, this study examines the relationship between the quantity of participants' verbal and nonverbal behaviors and presentation quality in online presentations. In doing so, we examine the impact of the transmission of nonverbal information by controlling the number of participants who have their cameras turned on.

In this study, Bibliobattle was employed as a presentation game. Bibliobattle is a game based on book reviews proposed by Taniguchi et al. [1]. Several people bring in books they have read and found interesting, each person in turn gives a 5-minute presentation and a 3-minute Q&A session, and at the end everyone, including the audience, votes on which book they would like to read the most to determine the winner. The game was chosen as the subject of this study because the time for each presentation is fixed, the conditions are controlled, and the quality of the presentations is evaluated by the votes.

This paper is based on results obtained from a project, JPNP23025, commissioned by the New Energy and Industrial Technology Development Organization (NEDO).

II. Related Work

There have been many attempts to understand conversation as a collaborative process that includes not only utterances and gestures from the speaker but also listener responses. For example, Bavelas et al. [2] analyzed listener responses by focusing on the role of gaze.

Otsuka's research group has conducted various studies on understanding the structure of verbal and nonverbal behavior in conversation, focusing on face-to-face conversations between about four people. Otsuka et al. [3] used eye gaze and head movements as cues to estimate participants' turn taking. Kumano et al. [4] built a prediction model of empathy among conversation participants by focusing on their gaze and facial expressions. Ishii et al. [5] focused on eye gaze behaviors, the timing of speech utterances and the end of speech utterances with the aim of estimating empathy skills in multiparty meetings. Their results showed that eye gaze behavior with speech differs according to a person's empathy skill level, and that it is effective to use both eye gaze behavior and dialogue acts to estimate empathy skills. We, too, try to understand the listeners' empathy from their verbal and nonverbal responses, especially nodding.

Matsumura et al. [6] analyzed empathy through listener responses in Bibliobattle. They examined the relationship between participants' nonverbal behaviors and the results of Bibliobattle voting, paying particular attention to the co-occurrence of multiple listener behaviors (e.g., backchannel and nodding). Our study aims at a more detailed understanding by focusing on verbal as well as nonverbal behaviors. In addition, we also discuss the value of listener responses to the presenter by observing online Bibliobattle rather than face-to-face environments.

While most of the previous studies have focused on face-to-face conversations, this study focuses on online conversations. One contribution of this study is to confirm that listener's

2836-3795/24/$31.00 ©2024 IEEE
DOI 10.1109/COMPSAC61105.2024.00331

nodding plays an important role in online conversations as well as in face-to-face conversations. Another value of this study is to focus on the effects of what is possible only in an online environment, namely, the ability to turn on and off individual visual and auditory communication channels.

III. Conducting Online Bibliobattles

A. Purpose and Summary of Data Recording

Data from online Bibliobattles were recorded for the purpose of this study, which was to estimate presentation content transfer and empathy acquisition from listener responses. Data recording was conducted over three days, with totally four battles; in each battle, the four participants brought a book to present. Those four participants were asked to give a presentation on a different book on all four battles. In addition to these four participants, we had four people participate as audience. In other words, eight people participated in four Bibliobattles. Each presentation consisted of a 5-minute presentation and a 3-minute Q&A session, and the winner was decided by voting after the four presentations.

B. Recording Environment

BiblioBattle was conducted via online meeting. Zoom meetings were used as the online meeting tool. Because we wanted to investigate the effect of the visibility of audience responses on the presentation, we asked participants to turn their cameras on and off as directed by the experimenter. The amount of cameras that were turned on was varied from battle to battle to ensure that conditions were the same for each presenter. On the other hand, in order to record the amount of nodding, utterances, and their contents, a smartphone was placed in front of each participant in addition to the computer used for Zoom meeting participation to record video and audio of the upper body.

C. Data Recording Methods

In addition to the eight participants (four presenters and four audience members), one experimenter participated in the Zoom meeting as a moderator. The Zoom meeting was recorded as well as the Bibliobattle was conducted. Figure 1 shows an example of the Zoom screen during the Bibliobattle. Each participant was asked to participate in the Zoom meeting from a quiet space such as his/her home. We also ask them to set up a smartphone next to each participant's computer to record their own upper body video and audio.

D. Data Recording Flow

One experimenter participated as a moderator and facilitated the Bibliobattle. He instructed who was the presenter for each presentation and who was to turn on the camera. Each presentation was followed by five minutes of presentation and three minutes of Q&A. After four presentations, the winner was decided by a vote of all eight participants. After the game, the participants were asked to answer a questionnaire after each presentation, in which they were asked to rate their interest in the book presented and their level of understanding of the book on a 5-point scale from 1 to 5, respectively.

Fig. 1. Example of screen during online Bibliobattle.

IV. Data Analysis

This study analyzes verbal and nonverbal behaviors during online biblio-battles, as well as questionnaire evaluations. For the verbal analysis, the amount of speech was collected from audio transcriptions. For the nonverbal analysis, we focused on head movements during biblio battles recorded with a smartphone. The questionnaire included items such as interest in the books introduced in the Bibliobattle, evaluation of comprehension, and which presentations were interesting. This chapter describes the methodology and data used for the analysis.

A. Data Coding of Verbal Behavior

The audio data of each participant was transcribed using Adobe Premiere Pro, which was chosen because it allows for noise reduction and transcription at the same time. In this study, linguistic behavior was analyzed from the volume of each participant's speech. In Bibliobattle, there is a time limit, which sets an upper limit to the amount of time that each participant can speak, and we believe that the more each participant speaks, the greater the impact within the Bibliobattle session. For this reason, the number of characters of the speech output from the transcription was tabulated. The amount of speech was counted separately during the presentation and during the Q&A session in order to read the influence of each type of speech during the presentation and during the Q&A session.

B. Data Coding of Nonverbal Behavior

Only the up-and-down movements of the head movements were acquired, and nonverbal behaviors such as nodding were analyzed. Nonverbal behaviors were recorded by the smartphone provided to each participant during the Bibliobattle, and were output as quantitative data using OpenFace. OpenFace outputs the head movement data in terms of head position and rotation. The pitch angle motion (i.e., the rotational motion corresponding to the nodding of the head) was used to measure the head motion by obtaining the sum of the differences from one frame ago, i.e., the cumulative amount of head motion. The data were formatted for analysis so that they could be divided by each participant, by each schedule, by each presentation, and by the presentation and Q&A session.

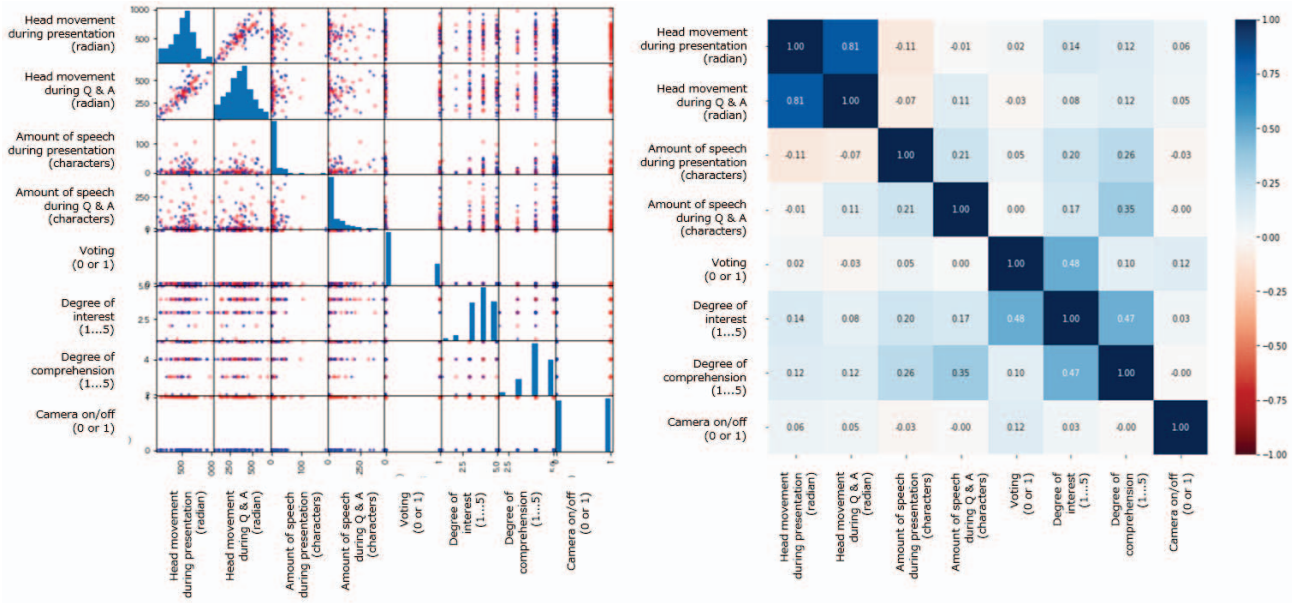

Fig. 2. Analysis of the relationship between the amount of verbal and nonverbal behavior and questionnaire results. Left: Scatterplot (bars indicate frequency), color coding in scatterplot indicates camera ON (red)/OFF (blue). Right: Correlation coefficients between elements.

C. Questionnaire Data

The questionnaire was administered at the end of the biblio-battle and included items that evaluated the relationship between the presenter and the participants, their knowledge of the books introduced during the presentation, and their understanding of the content of the presentation. We attempted to clarify the relationship between these questionnaire items and each participant's verbal and nonverbal behavior. The items of the questionnaire that are of interest are as follows.

- How well did you understand the presenter's explanation of the book? (5-point scale, 1: did not understand, 5: understood)
- How much did the presenter's description of the book make you want to read it? (5-point scale, 1: don't want to read, 5: want to read)
- Were there any books you knew about? (multiple answers allowed)
- Have you read any of these books? (Multiple answers allowed)

V. Results and Discussion

A. Trends observed in the cumulative data of verbal and nonverbal behaviors of all participants

The data on verbal and nonverbal behaviors were organized, and the following elements were subjected to analysis.

- Head movement during presentation (radian)
- Head movement during Q&A (radian)
- Amount of speech during presentation (characters)
- Amount of speech during Q&A (characters)
- Whether you voted for the presenter (0 or 1)
- Camera's ON/OFF (0 or 1)
- How well did you understand the presenter's explanation of the book? (5-point scale, 1: did not understand, 5: understood)
- How much did the presenter's description of the book make you want to read it? (5-point scale, 1: don't want to read, 5: want to read)

Scatter plots and correlations of the aggregated data are shown in Figure 2. Here, the scatter plots and correlation coefficients for the following elements are visualized as a heat map. The color coding in the scatter plots is based on whether the camera is ON (red) or OFF (blue).

As a result, it was confirmed that the amount of speech during the Q&A session correlated well with the level of comprehension of the presentation content, that the level of comprehension of the presentation content correlated with the level of interest in the book, and that the level of interest in the book strongly contributed to voting. However, the cumulative data for all participants confirmed that the reactions during the presentation, i.e., the amount of nodding and the amount of speech during the Q&A session, did not correlate with the voting results.

B. Analysis focusing on the individuality of each participant

Aggregated data from all participants would only provide results that buried the individuality of each participant. Therefore, a closer look at the data for individual participants identified three distinct groups, each with different tendencies.

For the two participants in the first group, nodding was strongly correlated with comprehension and interest. In other words, interest in the presentation was easily observed as a

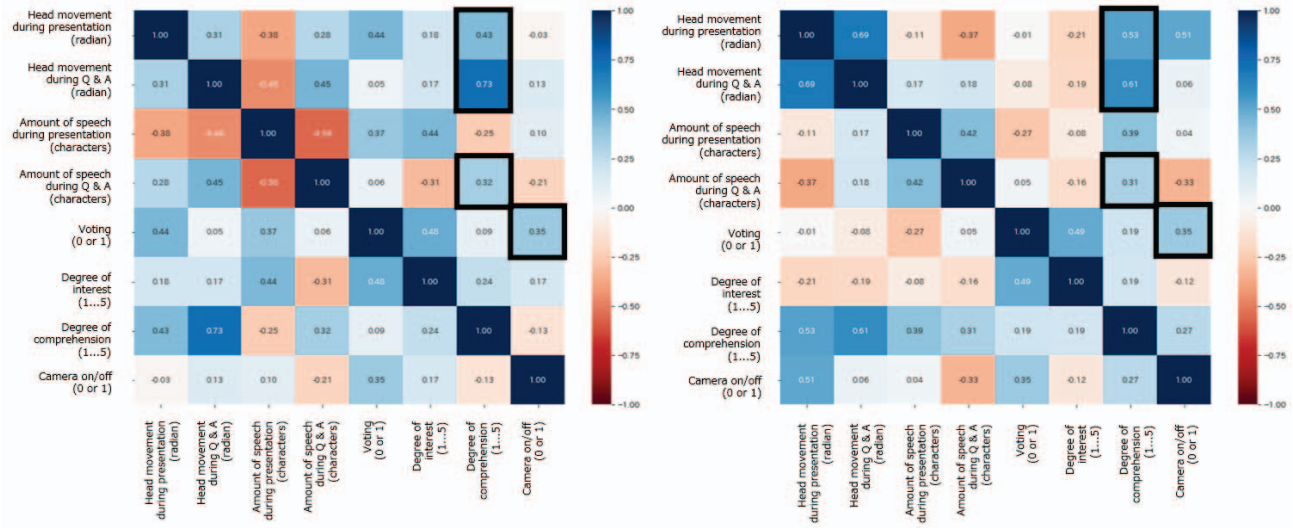

Fig. 3. Participants with high correlations between verbal and nonverbal behaviors and comprehension.

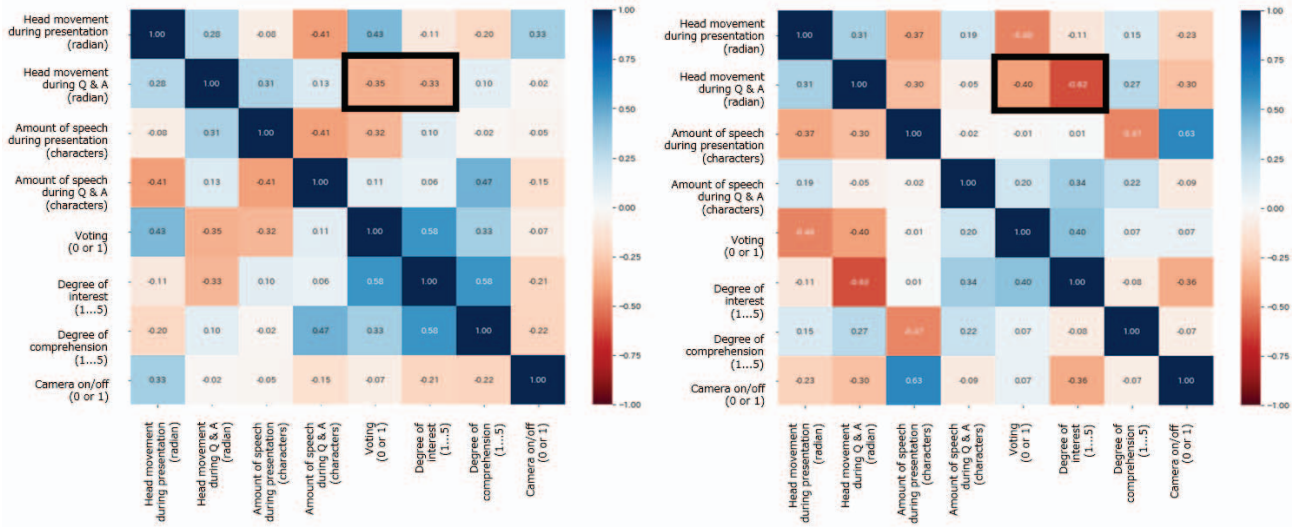

Fig. 4. Participants with an inverse correlation between the amount of nodding and the level of interest/comprehension.

nonverbal behavior. The correlation between verbal and nonverbal behavior and subjective evaluation of the presentation in the data of these two individuals is shown in Figure 3.

The highlighted celles in Figure 3 show significant correlation between verbal/nonverbal behaviors and the level of understanding of the book. These are considered to be the participants whose nonverbal behaviors are likely to be manifested in behaviors that increase their understanding of the book, such as speech during the Q&A session, and in their understanding of the content of the presentation.

On the other hand, some participants had an inverse correlation between their nonverbal behavior and their level of interest and comprehension. A correlation chart of their data is shown in Figure 4.

This group showed inverse correlations between nonverbal behavior and level of interest, and between nonverbal behavior and level of comprehension. In particular, the inverse correlation between head movements during the Q&A session showed that interest and comprehension did not increase in response to nodding, which was the opposite of the response of the previous group.

Interestingly, personality assessment showed that the former group was more extroverted and the latter group more introverted. The Big Five (OCEAN) Personal Traits were used as a personality assessment. It might be interpreted that extroverts are more likely to manifest their interest in externally observable behaviors, while introverts are more likely to focus quietly on objects of interest and react rather

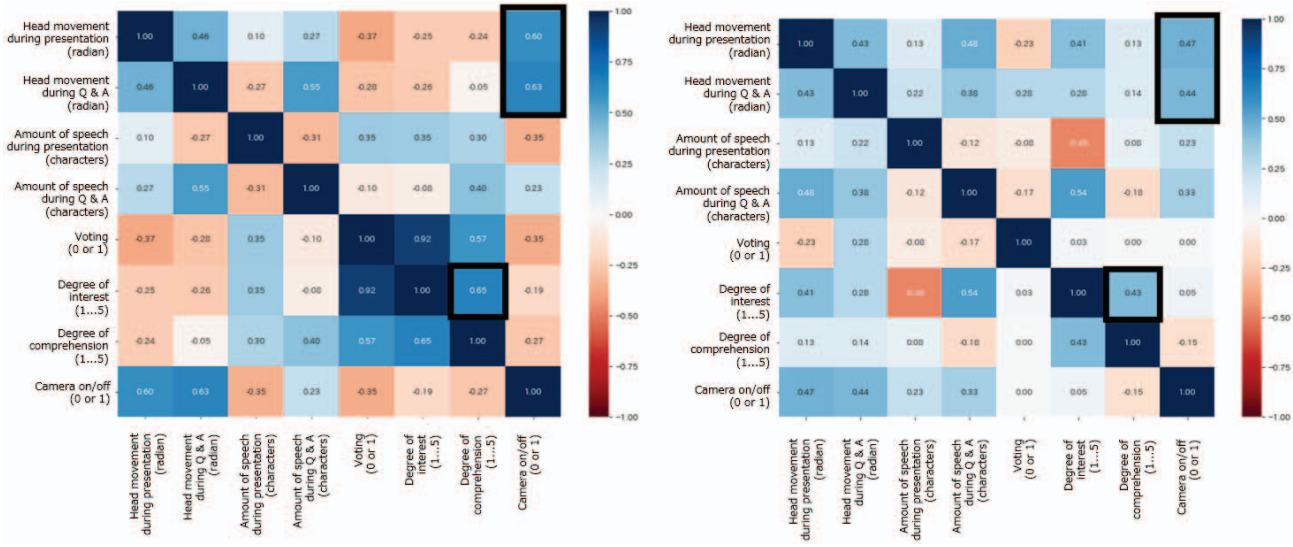

Fig. 5. Participants nodding significantly depending on whether the camera is on or off

socially to objects of less interest.

The last group to be introduced was characterized by the influence of camera on/off on nonverbal behavior. A correlation chart of the participants categorized in Figure 5 is shown.

This group's head movements became more active when the camera was turned on, which may have been influenced by their psychological need to have some kind of reaction to the camera when they were in the picture. Although the change in behavior when the camera was switched on was not observed in the analysis of the tendency of all the participants, it was possible to find some participants who were affected by the camera as an individual characteristic.

C. Impact on presentation with and without visual feedback from the audience

One of the main motivations for this study was to find out whether the audience's camera on contributed to the quality of the presentation.

For each presentation, we quantified the number of camera-ons by audiences, the number of votes received, the total number of interest levels by audiences, and the total number of comprehension levels by audiences, and examined the correlations between them.

Figure 6 summarizes the results. Each value in the figure is as follows.

- Vote: Number of votes obtained
- Interest: Total audience interest score
- Comprehension: Total audience comprehension score
- Camera: Number of audience members who had their cameras on

The figure on the left plots the data for each presentation. Four Bibliobattles were held, with four presentations each battle, so there are 16 plots. The color of the plots corresponds to the presenter's camera on/off, with red indicating on and blue off. The figure on the right shows the correlation between each value.

Strong correlations were observed between the number of votes and the level of interest, and between the level of interest and the level of understanding. However, camera on/off was largely independent of the participants' level of interest and comprehension. A weak inverse correlation was observed between the number of cameras on and the number of obtained votes, which was the opposite of what we expected.

Possible causes include the fact that the participants were known to each other, that seeing the other's face may increase tension, and that the voting results may depend purely on the attractiveness of the book and be independent of the quality of the presentation. Therefore, for more detailed discussions, it may be necessary to control the books used for the Bibliobattle.

VI. Conclusions

In this study, we analyzed verbal and nonverbal behaviors in online Bibliobattle, focusing on listener responses. Data from Bibliobattle in an online environment were collected, and questionnaire data were collected along with each participant's quantified verbal and nonverbal behaviors.

As a result, it was confirmed that the amount of speech during the Q&A session correlated well with the level of comprehension of the presentation content, that the level of comprehension of the presentation content correlated with the level of interest in the book, and that the level of interest in the book strongly contributed to voting. However, the cumulative data for all participants confirmed that the reactions during the presentation, i.e., the amount of nodding and the amount of speech during the Q&A session, did not correlate with the voting results.

Therefore, a closer look at the data for individual participants identified three distinct groups, each with different

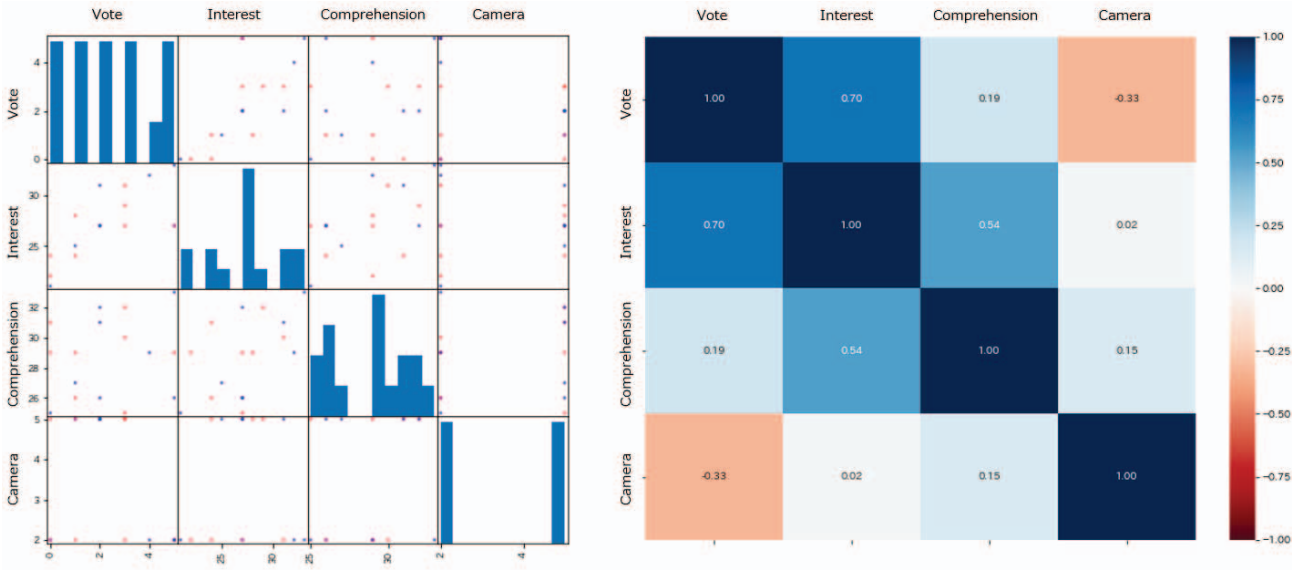

Fig. 6. Effect of the number of cameras turned on.

tendencies. There were participants for whom nodding was strongly correlated with comprehension and interest, i.e., their interest in the presentation was easily observed as a nonverbal behavior. On the other hand, there were participants for whom the amount of nodding correlated inversely with interest in the book and voting. Personality tests showed that the former group was more extroverted and the latter group more introverted. The last group was significantly affected by the camera being on or off, with significantly more nodding when the camera was on and a weak correlation with the amount of speech during the Q&A session. However, this did not lead to voting.

One of the main motivations for this study was to find out whether the audience's camera on contributed to the quality of the presentation. As the result, camera on/off was largely independent of the participants' level of interest and comprehension. A weak inverse correlation was observed for voting, which was the opposite of what was initially expected.

Because of the small number of participants in the experiment analyzed, excessive generalization of the results presented here should be avoided. However, it is worthwhile to confirm that our intuition, e.g., responses such as nodding, correlate with comprehension of the content of the presentation, and the finding that the positive or negative correlation between empathy for the presentation and response (amount of nodding and utterance in Q&A session) seems to be related to personality is also interesting. A better understanding of presentation quality from the interaction between presenter and listeners requires a discussion based on data with more and diverse participants.

REFERENCES

[1] T. Taniguchi, H. Kawakami, and O. Katai, "Bibliobattle: Informal community scheme based on book review sessions," in *Proc. 8th Int. Workshop on Social Intelligence Design (SID 2009)*, 2009, pp. 92–98.

[2] J. B. Bavelas, L. Coates, and T. Johnson, "Listener responses as a collaborative process: The role of gaze," *Journal of Communication*, vol. 52, no. 3, pp. 566–580, January 2006.

[3] K. Otsuka, H. Sawada, and J. Yamato, "Automatic inference of cross-modal nonverbal interactions in multiparty conversations: "who responds to whom, when, and how?" from gaze, head gestures, and utterances," in *Proceedings of the 9th International Conference on Multimodal Interfaces (ICMI '07)*, 2007, pp. 255–262.

[4] S. Kumano, K. Otsuka, D. Mikami, M. Matsuda, and J. Yamato, "Analyzing interpersonal empathy via collective impressions," *IEEE Transactions on Affective Computing*, vol. 6, no. 4, pp. 324–336, 2015.

[5] R. Ishii, K. Otsuka, S. Kumano, R. Higashinaka, and J. Tomita, "Analyzing gaze behavior and dialogue act during turn-taking for estimating empathy skill level," in *Proceedings of the 20th ACM International Conference on Multimodal Interaction (ICMI '18)*, 2018, pp. 31–39.

[6] K. Matsumura, Y. Sumi, and M. Sugiya, "Analyzing listeners' empathy by their nonverbal behaviors in bibliobattle," *Journal of Information Processing*, vol. 25, pp. 361–365, 2017.