A Comparison of Praising Skills in Face-to-Face and Remote Dialogues

Toshiki Onishi[1,*]**, Asahi Ogushi**[1,*]**, Yohei Tahara**[2,*]**, Ryo Ishii**[3,†]**,**
Atsushi Fukayama[3,†]**, Takao Nakamura**[3,†]**, Akihiro Miyata**[2,‡]

[1]Graduate School of Integrated Basic Sciences, Nihon University, Tokyo, Japan
[2]College of Humanities and Sciences, Nihon University, Tokyo, Japan
[3]NTT Human Informatics Laboratories, NTT Corporation, Kanagawa, Japan
*{chto22002, chas22009, chyo19046}@g.nihon-u.ac.jp
†{ryo.ishii.ct, atsushi.fukayama.ap, takao.nakamura.vp}@hco.ntt.co.jp
‡miyata.akihiro@acm.org

Abstract

Praising behavior is considered to an important method of communication in daily life and social activities. An engineering analysis of praising behavior is therefore valuable. However, a dialogue corpus for this analysis has not yet been developed. Therefore, we develop corpuses for face-to-face and remote two-party dialogues with ratings of praising skills. The corpuses enable us to clarify how to use verbal and nonverbal behaviors for successfully praise. In this paper, we analyze the differences between the face-to-face and remote corpuses, in particular the expressions in adjudged praising scenes in both corpuses, and also evaluated praising skills. We also compare differences in head motion, gaze behavior, facial expression in high-rated praising scenes in both corpuses. The results showed that the distribution of praising scores was similar in face-to-face and remote dialogues, although the ratio of the number of praising scenes to the number of utterances was different. In addition, we confirmed differences in praising behavior in face-to-face and remote dialogues.

Keywords: multimodal interaction, dialogue corpus, praise

1. Introduction

Praising behavior is considered an important method of communication in daily life and social activities. Praise is a verbal and nonverbal behaviors expression of approval that is directed toward the behavior and character of the target (Brophy, 1981; Jenkins et al., 2015; Kalis et al., 2007). In addition, praising behavior is considered to be a complex social communication rather than simple one-way communication of intention from a person who gives praise to a person who receives praise (Henderlong and Lepper, 2002). Many studies on praising behavior have analyzed the results of questionnaires and evaluations conducted after specific tasks to understand the effects and influences of praising behavior (Henderlong and Lepper, 2002; Ennis et al., 2018). An engineering analysis of praising behavior is valuable and would contribute to communication in the scenes of business and education, as well as to the study of pedagogy and psychology. However, a dialogue corpus for this analysis has not yet been developed. Therefore, we have not been able to clarify how to use verbal and nonverbal behaviors for successfully praise. In addition, the influence of the COVID-19 pandemic has increased the opportunities for remote communication. We consider that praising skills are also important in such situations. However, a corpus for the study of praising skills is not available.

To the best of our knowledge, this is the first study to develop corpuses of face-to-face and remote two-party dialogues with ratings of praising skills. Our corpus of face-to-face and remote dialogues enables to clarify how to use verbal and nonverbal behaviors for success-

fully praise. In this paper, we analyze the differences between our face-to-face and remote corpuses. Human verbal and nonverbal behaviors are known to differ in face-to-face and remote dialogues in term of the information transmitted between the parties (Doherty-Sneddon et al., 1997). In particular, nonverbal behavior is considered to be conveyed to a lesser degree in remote dialogue. In the case of praising behavior, the method of conveying verbal and nonverbal behaviors as well as the important behaviors of praising a partner are considered to differ in face-to-face and remote dialogues. An understanding of praising behaviors in the two dialogue environments enables successful praising in both face-to-face and remote dialogues. The contribution of this paper is an initial examination of the differences between face-to-face and remote corpuses, focusing on the evaluation of praising skills and behaviors such as head motion, gaze behavior, and facial expression during praising.

2. Related work

2.1. Studies on praising behavior

Many studies have been conducted on praising behavior to clarify their functions and effects. However, the majority of the studies have focused on praising behaviors in educational scenes, while few have focused on praising behaviors in daily life and business scenes. Brophy (1981) suggested that praising behavior by teachers is widely recommended as a means of reinforcing students and has various functions other than reinforcing students' behavior and performance. Teacher praise has been shown to have a variety of

functions beyond reinforcing students' behavior and academic performance, including (1) spontaneous expressions of surprise and admiration, (2) maintaining balance in response to previous remarks, (3) proxy reinforcement (praising one student to tell the rest of the class that they should do so), (4) avoidance of negative expressions (teaching through praise without negative expressions), (5) a means of relieving tension, (6) student elicitation, (7) transition rituals (signaling the end of a task and the transition to the next task), and (8) comforting awards and encouragement. Henderlong and Lepper (2002) described the effects of praising behavior on children's intrinsic motivation and perseverance as complex and diverse, ranging from valuable to marginally valuable to detrimental. In addition, they described that praising behavior may have different effects on the motivation of the receiver depending on the receiver's traits such as age, gender, and culture. Moreover, praising behavior should be based on the results of behaviors, the behaviors should be clearly described, and praising behavior should be sincere in order to increase the effectiveness of praising behavior (Ennis et al., 2018).

2.2. Estimation of personality traits and performance

Many studies have analyzed behaviors and abilities in specific tasks and scenes using the verbal and nonverbal behaviors of humans. Batrinca et al. (2016) investigated the automatic recognition of the Big Five personality traits (extraversion, agreeableness, conscientiousness, emotional stability, and openness to experience) from audio and video data collected in two scenarios: human-machine interaction and human-human interaction. Aran and Gatica-Perez (2013) investigated the prediction of personality traits of individuals participating in small-group discussions. Lin and Lee (2018) proposed a framework that models the vocal behaviors of both the target speaker and his/her contextual interlocutors to improve the prediction performance for scores of ten different personality traits in the ELEA corpus. Jayagopi et al. (2012) presented a framework for defining and extracting group behavioral cues characterizing speaking and looking patterns in face-to-face interactions. Ramanarayanan et al. (2015) presented a comparative analysis of three different feature sets to predict different human-rated scores of presentation proficiency. Park et al. (2014) presented their computational approach in using verbal and nonverbal behavior from multiple modalities of communication to predict a speaker's persuasiveness in online social multimedia content and showed that having prior knowledge of the speaker's sentiment partiality contributes to better prediction of the level of persuasiveness. Nguyen et al. (2014) proposed a computational framework to automatically predict hirability in real job interviews using the nonverbal cues of applicants' and interviewers' extracted from audio and visual modalities. Sanchez-

Cortes et al. (2012) proposed a computational framework to infer emergent leadership in newly formed groups from nonverbal behavior by combining speaking turns, prosodic features, visual activity, and motion. Ishii et al. (2018) investigated gaze behavior and dialogue act categories during turn-keeping/changing changes depending on the empathy skill level, which is measured using Davis' IRI. Soleymani et al. (2019) investigated verbal and nonverbal behaviors during intimate self-disclosure.

3. Research goal

As discussed in 2.1, many studies related to praising behavior have been reported. These studies clarified the functions and effects of praising behavior. Therefore, they have not clarified how to behave in actual dialogue scenes. In addition, as discussed in 2.2, many studies have analyzed behaviors and abilities in specific tasks and scenes using verbal and nonverbal behaviors. These studies have analyzed speakers' personality traits and abilities, such as presentation, empathy, and self-disclosure during communication. However, few studies have analyzed praising behavior using human verbal and nonverbal behaviors. Therefore, the behaviors that are expressed during praising behavior in actual dialogue have not been clarified. In this study, we attempt to analyze the praising behavior from an engineering point of view and to clarify which behavior is important in praising behavior. However, there has not been a dialogue corpus available to analyze these behaviors.

We develop corpuses of face-to-face and remote two-party dialogues with ratings of praising skills, for the first time to the best of our knowledge. These corpuses of face-to-face and remote dialogues enables us to clarify how to use verbal and nonverbal behaviors to praise successfully. Our research goal is an initial examination of the differences between face-to-face and remote corpuses, focusing on the evaluation of praising skills and behaviors such as head motion, gaze behavior, and facial expression during praising.

4. Dialogue corpus

We developed corpus of face-to-face and remote two-party dialogues with ratings of praising skills for the first time to the best of our knowledge. In our previous work (Onishi et al., 2020), we recorded face-to-face two-party dialogues, annotated the dialogue data, and evaluated the praising skills. In this paper, we developed the face-to-face corpus with increased number of dialogue participants and annotators from previous work. In addition, we newly recorded remote two-party dialogues, annotated the dialogue data, and evaluated praising skills.

4.1. Face-to-face dialogue
4.1.1. Recording of two-party dialogue
We recorded face-to-face two-party dialogues to record verbal and nonverbal behaviors. The participants in the

Figure 1: The photograph of two-party face-to-face dialogue.

two-party dialogues were 34 university students in their twenties (28 males and six females) who were divided into 17 pairs. Among the 17 pairs, 14 pairs included participants meeting for the first time, two pairs included acquaintances, and one pair included friends. To begin recording dialogues, we requested participants to prepare two or more episodes about what they had been working hard to prepare materials for the dialogue. The participants were seated facing each other and separated by 180 cm apart, as shown in Figure 1. The dialogues were recorded using a video camera to record each participant's head and face behaviors and a microphone to record each participant's voice. Each pair of participants (participants A and B) performed dialogues (1) to (3) in accordance with the experimenter's instructions.

(1) A self-introduction (5 min).

(2) Dialogue with participant A as a praiser and participant B as a receiver (5 min).

(3) Dialogue with participant B as a praiser and participant A as a receiver (5 min).

We recorded 17 pairs of dialogues (1) to (3) for a total of 255 minutes of two-party dialogues. Dialogue (1) (self-introduction) was not used in our analysis because many of the pairs were meeting for the first time, and its purpose was simply to relieve the tension between participants. In dialogues (2) and (3), the receiver was instructed to discuss the things that they had been working hard to accomplish. To ensure that the participants conversed naturally regarding a variety of topics, we also allowed them to discuss topics that they had not prepared beforehand. The praiser was instructed to praise the receiver. However, we allowed the participants to raise questions and react freely to avoid any unnatural dialogues that would have involved unilateral praising. This procedure was approved by the ethics committee.

	Scenes	Mean (sec)	Max (sec)	Min (sec)
Utterance (praiser)	2701	1.324	23.117	0.062
Utterance (receiver)	3413	2.040	26.234	0.018
Praising scenes	228	2.018	9.127	0.368

Table 1: Information about utterance scenes and praising scenes in face-to-face corpus.

4.1.2. Annotation of dialogue data

We annotated the dialogue data recorded in 4.1.1. We used ELAN (Brugman and Russel, 2004), a tool for annotating video and audio data, to manually annotate utterance scenes in the video and audio data of each participant. The results of the utterance scenes are presented in Table 1.

Utterance scenes: a continuous voice intervals with a silent interval of less than 400 ms.

In our previous work (Onishi et al., 2020), we used continuous voice intervals of less than 200 ms of silence. However, we confirmed that the utterances were separated into small intervals. In this paper, we set continuous voice intervals of less than 400 ms so that the intervals of utterances were natural.

4.1.3. Evaluation of praising skills

The evaluation of praising skills was conducted by five third-party annotators who did not participate in the two-party dialogue. The annotators did not have any training or qualifications to avoid the influence of prior knowledge or preconceptions. They made the following judgments and evaluations for each utterance scene of the praiser extracted in 4.1.2, referring to the video data recorded from the video camera set up in front of the praiser and the audio data recorded from the microphone attached to the praiser.

- Judgment of whether the praiser praises the dialogue partner in the scene.

- If the praiser was praising the dialogue partner, the evaluation of praising skills on a 7-point Likert scale from 1 (I do not think the praiser is successfully praising) to 7 (I think the praiser is successfully praising).

From the above judgments and evaluations, we defined praising scenes and praising scores. The information of praising scenes is shown in Table 1.

Praising scenes: the scene in which three or more annotators judged that the praiser was praising in each utterance scene.

Praising scores: the mean of the evaluations by annotators who judged that the praiser was praising in each praising scene.

We evaluated the rate of concordance of praising scores among annotators using intraclass correlation coefficients (ICC) (Shrout and Fleiss, 1979). We calculated

Figure 2: The photograph of two-party remote dialogue.

	Scenes	Mean (sec)	Max (sec)	Min (sec)
Utterance (praiser)	16351	1.233	11.180	0.450
Utterance (receiver)	27834	1.959	13.430	0.450
Praising scenes	236	1.998	10.610	0.470

Table 2: Information about utterance scenes and praising scenes in remote corpus.

the ICC for each combination of three to five annotators, and calculated the weighted average by considering the number of samples. The results were $ICC(2, k) = 0.571$. This result suggests that praising scores are reliable data with a medium level of concordance among annotators. In our previous study (Onishi et al., 2020), an annotator judged whether utterances were praising, and participants who were the receiver in each dialogue rated the praising skills of the partner. However, we were concerned that the judgments and evaluations would be subjective, so the judgments and evaluations were performed by five annotators.

4.2. Remote dialogue

4.2.1. Recording of two-party dialogue

We recorded two-party remote dialogues using an online communication tool to record verbal and nonverbal behaviors. The participants of the dialogues were 40 people (20 males and 20 females) in their twenties to fifties, and each participant talked three times with a different partner. We recorded a total of 60 pairs of dialogues. The age and gender of the participants in each group were the same, and they had not met each other. The participants were seated in front of a PC, as shown in Figure 2. The recording of dialogues used Zoom[1], an online communication tool. We recorded the video of each participant using the camera on a PC set up in front of the participants, and the audio of each participant using the microphone on a PC. The participants (participant A and B) performed the following three dialogues according to the experimenter's instructions, as in 4.1.1.

(1) A self-introduction (5 min).

(2) Dialogue with participant A as a praiser and participant B as a receiver (10 min).

(3) Dialogue with participant B as a praiser and participant A as a receiver (10 min).

We recorded 60 pairs of dialogues (1) to (3) for a total of 1500 minutes of two-party dialogues. The topic of dialogues was specific experiences and stories about oneself having worked hard in the past. We asked the receiver to continue the dialogue on the above themes. To avoid unnatural dialogues, we told the receiver to talk about their experiences in natural dialogues rather than speeches or presentations. We asked the praiser to praise their partners' experiences when they felt necessary to praise. In addition, we allowed the praiser to raise questions, as appropriate.

4.2.2. Annotation of dialogue data

We automatically annotated utterance scenes for each participant's voice data using ASR[2], an automatic speech recognition system. The results of the utterance scenes are presented in Table 2.

Utterance scenes: a continuous voice intervals with a silent interval of less than 400 ms, as in 4.1.2.

4.2.3. Evaluation of praising skills

The evaluation of praising skills was conducted by five third-party annotators who were not participating in the two-party dialogue, using the same method as in 4.1.3. Specifically, annotators made the following judgments and evaluations for each utterance scene of the praiser extracted in 4.2.2, referring to the video data recorded from the PC set up in front of the praiser and the audio data recorded from the microphone.

- Judgment of whether the praiser praises the dialogue partner in the scene.

- If the praiser was praising the dialogue partner, the evaluation of praising skills on a 7-point Likert scale from 1 (I do not think the praiser is successfully praising) to 7 (I think the praiser is successfully praising).

From the above judgments and evaluations, we defined praising scenes and praising scores.

Praising scenes: the scene in which three or more annotators judged that the praiser was praising in each utterance scene.

---

[1]Zoom Cloud Meetings, Zoom Video Communications Inc., https://zoom.us/.
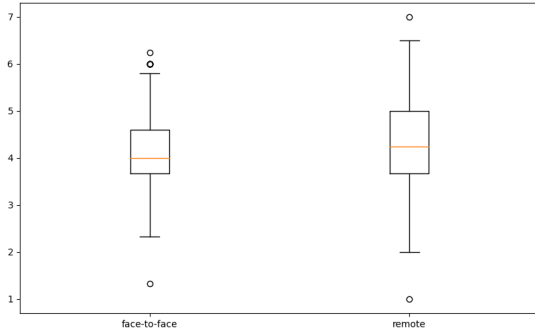
[2]Automatic speech recognition (ASR) system developed by NTT Corp.

Figure 3: The distribution of praising scores in face-to-face and remote dialogues.

	Mean	SD	Max	Min
face-to-face	4.124	0.808	6.250	1.333
remote	4.280	0.931	7.000	1.000

Table 3: The statistics of praising scores in face-to-face and remote dialogues.

Praising scores: the mean of the evaluations by annotators who judged that the praiser was praising in each praising scene.

The information of the praising scenes is shown in Table 2. We calculated the ICC for each combination of three to five annotators, and calculated the weighted average by considering the number of samples. The results were $ICC(2,k) = 0.701$. This result suggests that praising scores are reliable data with a high level of concordance among annotators.

5. Comparison of face-to-face and remote dialogues

5.1. Praising scores

We compared the praising scenes and praising scores that were extracted and calculated in 4.1.3 and 4.2.3, respectively, between face-to-face and remote dialogues. First, we compared praising scores between face-to-face and remote dialogues. The distribution of praising scores in face-to-face and remote dialogues is shown in Figure 3. The statistics of praising scores in face-to-face and remote dialogues are shown in Table 3. We performed an unpaired t-test between face-to-face and remote dialogue using praising scores. The result was $t(457) = 1.931(p = 0.054)$, indicating no significant difference. Thus, we consider that the dialogues recorded in this paper are face-to-face and remote dialogues with a uniform distribution of praising scores.

Second, we compared the praising scenes in face-to-face and remote dialogues. The percentage of praising scenes was 8.4% of all utterances in face-to-face dialogue, as shown in Table 1. On the other hand, the percentage of praising scenes was 1.4% among all utter-

	Mean	SD	Max	Min
face-to-face	6.909	4.297	18	1
remote	1.967	2.446	15	0

Table 4: The statistics of the number of Praise scenes in face-to-face and remote dialogues.

ance scenes in the remote dialogue, as shown in Table 2. Thus, the number of scenes in the remote dialogue was the same as in the face-to-face dialogue. However, the overall dialogue had fewer praising scenes in remote dialogue. In addition, we checked how many praising scenes existed in one round of dialogue. The statistics of the number of praising scenes in face-to-face and remote dialogues are shown in Table 4. Thus, we confirmed that face-to-face dialogues have one or more praising scenes. On the other hand, the remote dialogues have one or more times of the praising scene, although some dialogues have no praising scene. Thus, we considered that praise behavior is more likely to be generated in face-to-face dialogue, while praise behavior is less likely to be generated in remote dialogues.

5.2. Facial behaviors

5.2.1. Feature extraction

We extracted features related to head, gaze, and action units from video data captured by a video camera or PC set up in front of the participants using OpenFace (Baltrušaitis et al., 2016), a face image processing tool.

Head motion: We used the variance (_var), median (_med), and 10th (_p10) and 90th percentile values (_p90) of the rotation angles around the x-axis (pose_Rx), y-axis (pose_Ry), and z-axis (pose_Rz) of the head when the face was viewed from the video camera side and the x-axis was from left to right, the y-axis was from bottom to top, and the z-axis was from the front to the back.

Gaze behavior: We used the variance, median, and 10th and 90th percentile values of the angles in the x-axis (gaze_Ax) and y-axis (gaze_Ay) of gaze when the face was viewed from the video camera side, and the x-axis was from the left to the right, and the y-axis was from the bottom to the top.

Action units: Action units (Ekman and Friesen, 1977) represent the fundamental actions of individual muscles or muscle groups. We used the variance, median, and 10th and 90th percentile values of the intensity of the action units used in OpenFace. The action units used in this paper are listed in Table 5.

5.2.2. Comparison of features

We confirmed the difference in behavior between face-to-face and remote dialogues in successfully praising

Item	Content	Item	Content
AU01	Inner brow raiser	AU14	Dimpler
AU02	Outer brow raiser	AU15	Lip corner depressor
AU04	Brow lowerer	AU17	Chin raiser
AU05	Upper lid raiser	AU20	Lip stretcher
AU06	Cheek raiser	AU23	Lip tightener
AU07	Lid tightener	AU25	Lips part
AU09	Nose wrinkler	AU26	Jaw drop
AU10	Upper lip raiser	AU45	Blink
AU12	Lip corner puller		

Table 5: The list of Action Units.

Feature	Mean (face-to-face)		Mean (remote)
AU07_var	**0.213**	>	0.132
AU45_var	0.121	<	**0.301**
gaze_Ay_p10	0.011	<	**0.141**
AU04_p10	0.188	<	**0.334**
AU06_p10	0.797	<	**1.556**
AU07_p10	1.071	<	**1.768**
AU09_p10	**0.059**	>	0.009
AU26_p10	**0.206**	>	0.064
gaze_Ay_med	0.074	<	**0.184**
AU06_med	1.158	<	**1.924**
AU07_med	1.619	<	**2.213**
AU09_med	**0.186**	>	0.060
AU23_med	0.053	<	**0.156**
AU26_med	**0.519**	>	0.330
gaze_Ay_p90	0.140	<	**0.251**
AU06_p90	1.523	<	**2.295**
AU07_p90	2.145	<	**2.601**
AU09_p90	**0.405**	>	0.230
AU23_p90	0.251	<	**0.531**
AU45_p90	0.672	<	**1.018**

Table 6: The features with significant differences between face-to-face and remote dialogues.

scenes using the features in 5.2.1. Specifically, we confirmed the differences between the same features of the high group of praising scores in face-to-face and remote dialogues. In this paper, we defined low, medium, and high groups of praising scores for both face-to-face and remote dialogues.

Low group: praising scenes with a praising score of 3.8 points or less in both face-to-face and remote dialogues. (The total of face-to-face dialogue was 82 scenes, and remote dialogue was 87 scenes.)

Middle group: praising scenes with a praising score greater than 3.8 points and less than 4.4 points in face-to-face dialogue (65 scenes in total), and praising scenes with a praising score greater than 3.8 points and less than 4.6 points in remote dialogue (63 scenes in total).

High group: praising scenes with a praising score of 4.4 points or higher in face-to-face dialogue (81 scenes in total), and praising scenes with a praising score of 4.6 points or higher in remote dialogue (87 scenes in total).

In this paper, we use the features in the high group of praising scores as an initial step of the feature analysis. We performed an unpaired t-test for each feature in face-to-face and remote dialogues, and we confirmed a significant difference at the 1% level for the features shown in Table 6. Therefore, we consider that the features shown in Table 6 exhibit different behaviors when praising in face-to-face and remote dialogues. From Table 6, we confirmed that the dominant features in face-to-face dialogues are related to AU09 and AU26. AU09 represents the behavior of wrinkling the nose, and AU26 represents the behavior of lowering the chin and opening the lips. We consider these behaviors to be actively used in face-to-face dialogue. We confirmed that the dominant features in the remote dialogue are related to gaze_Ay, AU04, AU06, AU23, and AU45. gaze_Ay represents the behavior of moving the head to the left or right, AU04 represents the behavior of lowering the eyebrows, AU06 represents the behavior of lifting the cheeks, AU23 represents the behavior of closing the lips tightly, and AU45 represents the behavior of blinking. We consider these behaviors to be

actively used in remote dialogue. Although the variance of AU07 was dominant in face-to-face dialogue, the median, 10th percentile, and 90th percentile values were dominant in remote dialogue. AU07 represents the behavior of eyebrow tensing. This behavior is considered to be more intense in remote dialogue, while the eyebrows are tensed in face-to-face dialogue.

6. Discussion

From the analysis in 5.1, we confirmed that the distribution of praising scores is similar in face-to-face and remote dialogues, although the ratio of the number of praising scenes to the number of utterances is different. In particular, the number of praising scenes for the number of utterance scenes was lower in remote dialogues than in face-to-face dialogues. Therefore, we administered a questionnaire to the participants who were the receivers after the dialogue. The participants answered the following questions on a 7-point Likert scale ranging from 1 (disagree) to 7 (agree).

Q1: The partner praised me even if my actions were trivial.

Q2: The partner accepted what I had to say without criticism.

The results are shown in Figure 4. In Q1, we performed an unpaired t-test on the results of face-to-face and remote dialogues, and the results showed no significant difference. Thus, the participants who were receivers in both face-to-face and remote dialogues felt that their partners praised them even for trivial points. On the other hand, in Q2, we performed an unpaired t-test on the results of face-to-face dialogue and remote dialogues, and the results showed a significant difference
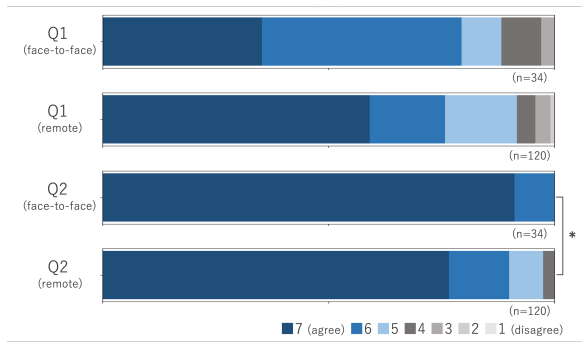
Figure 4: The results of Q1 and Q2 in face-to-face and remote dialogues. * indicates significance level of 1% or less ($p < .01$)

at the 1% level. Thus, face-to-face dialogue is more likely to make participants feel that their partner is listening to them. From the above, we consider that face-to-face dialogue is more likely than remote dialogue to receive and praise their partner's stories. In this paper, as an initial study, we analyzed the distribution of praising scores and the statistics of praising scenes. In the future, we would like to analyze praising skills in face-to-face and remote dialogues by analyzing the content of utterances, interviewing dialogue participants, and investigating what annotators liked when they judged praise.

In addition, from the analysis in 5.2.2, we confirmed the difference in the behavior of praise between face-to-face and remote dialogues. In this paper, as an initial study, we compared the behaviors of head motion, gaze behavior, and facial expression in the high group of praising scores and compared the behaviors of face-to-face dialogues and remote dialogues. In the future, we would like to conduct a more detailed analysis to clarify how to behave to successfully praise.

7. Conclusion

In this paper, we discussed the differences between face-to-face and remote corpuses, focusing on the evaluation of praising skills and behaviors such as head motion, gaze behavior, and facial expression during praising. We developed corpuses of face-to-face and remote two-party dialogues with ratings of praising skills, for the first time to the best of our knowledge. In particular, we analyzed the expressions in adjudged praising scenes in the face-to-face and remote corpuses, as well as evaluated the praising skills. In addition, we compared the differences in head motion, gaze behavior, and facial expression between face-to-face and remote dialogues in high-rated praising scenes. The results showed that the distribution of praising scores was similar in face-to-face and remote dialogues, although the ratio of the number of praising scenes to the number of utterance scenes was different. In addition, the results of the questionnaire suggested that face-to-face

dialogue were more likely than remote dialogue to receive and praise their partner's stories. Moreover, the difference in behavior in face-to-face and remote dialogues was confirmed. The corpus developed in this paper, as well as the results of the analysis in this paper, are expected to contribute to the analysis of praising. As this is a preliminary study, we have only compared the evaluation values and behaviors in scenes where the evaluation of praising skills was high between face-to-face and remote dialogues. Therefore, we could not determine which behaviors enable successful praising. In the future, we would like to analyze the content of utterances, conduct interviews with dialogue participants and annotators, and conduct a detailed analysis of behaviors. In addition, we plan to compare which behaviors contribute to successful praise in face-to-face and remote dialogues by constructing machine learning models for predicting praising skills.

8. Bibliographical References

Aran, O. and Gatica-Perez, D. (2013). One of a kind: inferring personality impressions in meetings. In *Proceedings of the 15th ACM International Conference on Multimodal Interaction (ICMI'13)*, pages 11–18.

Baltrušaitis, T., Robinson, P., and Morency, L. (2016). Openface: An open source facial behavior analysis toolkit. In *IEEE Winter Conference on Applications of Computer Vision (WACV'16)*, pages 1–10.

Batrinca, L., Mana, N., Lepri, B., Sebe, N., and Pianesi, F. (2016). Multimodal personality recognition in collaborative goal-oriented tasks. *IEEE Transactions on Multimedia*, 14(4):659–673.

Brophy, J. (1981). Teacher praise: A functional analysis. *Review of educational research*, 51(1):5–32.

Brugman, H. and Russel, A. (2004). Annotating multimedia / multi-modal resources with elan. In *Proceedings of the 4th International Conference on Language Resources and Language Evaluation (LREC'04)*, pages 2065–2068.

Doherty-Sneddon, G., Anderson, A., O'malley, C., Langton, S., Garrod, S., and Bruce, V. (1997). Face-to-face and video-mediated communication: A comparison of dialogue structure and task performance. *Journal of experimental psychology: applied*, 3(2):105–125.

Ekman, P. and Friesen, W. (1977). Manual for the facial action coding system. *Palo Alto: Consulting Psychologists Press*.

Ennis, R., Royer, D., Lane, K., Menzies, H., Oakes, W., and Schellman, L. (2018). Behavior-specific praise: An effective, efficient, low-intensity strategy to support student success. *Beyond Behavior*, 27(3):134–139.

Henderlong, J. and Lepper, M. (2002). The effects of praise on children's intrinsic motivation: A review and synthesis. *Psychological Bulletin*, 128(5):774–795.

Ishii, R., Otsuka, K., Kumano, S., Higashinaka, R., and Tomita, J. (2018). 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)*, pages 31–39.

Jayagopi, D., Sanchez-Cortes, D., Otsuka, K., Yamato, J., and Gatica-Perez, D. (2012). Linking speaking and looking behavior patterns with group composition, perception, and performance. In *Proceedings of the 14th ACM international conference on Multimodal interaction (ICMI'12)*, pages 433–440.

Jenkins, L., Floress, M., and Reinke, W. (2015). Rates and types of teacher praise: A review and future directions. *Psychology in the Schools*, 52(5):463–476.

Kalis, T., Vannest, K., and Parker, R. (2007). Praise counts: Using self-monitoring to increase effective teaching practices. *Preventing School Failure: Alternative Education for Children and Youth*, 51(3):20–27.

Lin, Y. and Lee, C. (2018). Using interlocutor-modulated attention blstm to predict personality traits in small group interaction. In *Proceedings of the 20th ACM International Conference on Multimodal Interaction (ICMI'18)*, pages 163–169.

Nguyen, L., Frauendorfer, D., Mast, M., and Gatica-Perez, D. (2014). Hire me: Computational inference of hirability in employment interviews based on nonverbal behavior. *IEEE Transactions Multimedia*, 16(4):1018–1031.

Onishi, T., Yamauchi, A., Ishii, R., Aono, Y., and Miyata, A. (2020). Analyzing nonverbal behaviors along with praising. In *Proceedings of the 22nd ACM International Conference on Multimodal Interaction (ICMI'20)*, pages 609–613.

Park, S., Shim, H., Chatterjee, M., Sagae, K., and Morency, L. (2014). Computational analysis of persuasiveness in social multimedia: A novel dataset and multimodal prediction approach. In *Proceedings of the 16th International Conference on Multimodal Interaction (ICMI'14)*, pages 50–57.

Ramanarayanan, V., Leong, C., Chen, L., Feng, G., and Suendermann-Oeft, D. (2015). Evaluating speech, face, emotion and body movement time-series features for automated multimodal presentation scoring. In *Proceedings of the 17th ACM on International Conference on Multimodal Interaction (ICMI'15)*, pages 23–30.

Sanchez-Cortes, D., Aran, O., Mast, M., and Gatica-Perez, D. (2012). A nonverbal behavior approach to identify emergent leaders in small groups. *IEEE Transactions Multimedia*, 14(3):816–832.

Shrout, P. and Fleiss, J. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological bulletin*, 86(2):420–428.

Soleymani, M., Stefanov, K., Kang, H., Ondras, J., and Gratch, J. (2019). Multimodal analysis and estimation of intimate self-disclosure. In *Proceedings of the 21st ACM International Conference on Multimodal Interaction (ICMI'19)*, pages 59–68.