Metacognition-EnGauge: Real-time Augmentation of Self-and-Group Engagement Levels Understanding by Gauge Interface in Online Meetings

KO WATANABE, RPTU Kaiserslautern-Landau & DFKI GmbH, Germany
ANDREAS DENGEL, RPTU Kaiserslautern-Landau & DFKI GmbH, Germany
SHOYA ISHIMARU, Osaka Metropolitan University & DFKI Lab Japan, Japan

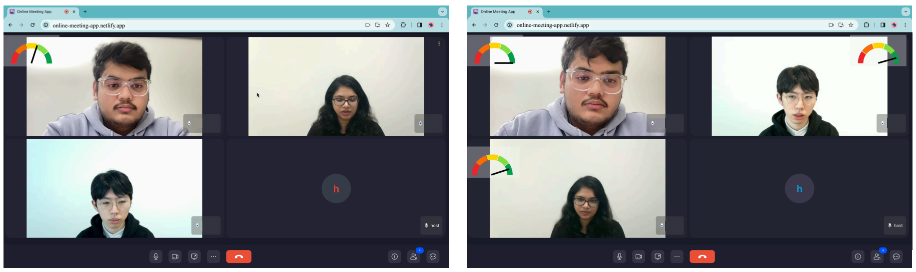


Fig. 1. Engagement levels are presented as gauge in the online meeting. Self-and-group metacognition augmentation are shown.

Engagement is one of the core cognitive states in communication. Increased engagement improves the quality and immersion of the conversation. In this demonstration, we aim to present a metacognition augmentation application called *Metacognition-EnGauge*. This work aims to discover three research questions. Does self-metacognition augmentation of engagement levels support participant to be engaged in the meetings? Does group-metacognition augmentation of engagement levels support participant to be engaged in the meetings? Do participants prefer none, self, or group-metacognition augmentation intervention in the online meetings? To answer these questions, we conducted online meeting experiments with 18 participants. As a result, we found that self-metacognition augmentation outperformed well on increasing average engagement levels. According to the survey, participant preference shows group-metacognition as the best intervention.

CCS Concepts: • **Applied computing** → **Computer-assisted instruction**; *Interactive learning environments*; *Distance learning*.

Additional Key Words and Phrases: Metacognition Augmentation, Affective Computing, Social Interaction, Social Cognitive State

ACM Reference Format:
Ko Watanabe, Andreas Dengel, and Shoya Ishimaru. 2024. Metacognition-EnGauge: Real-time Augmentation of Self-and-Group Engagement Levels Understanding by Gauge Interface in Online Meetings. In *The Augmented Humans International Conference (AHs 2024), April 4–6, 2024, Melbourne, VIC, Australia.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3652920.3653054

1 INTRODUCTION

Online meetings have expanded around the world with the COVID-19 pandemic and opened up the world to collaboratively work together remotely. The opportunity encouraged a massive number of people to work anytime and anywhere around the world. However, online meetings often have difficulty communicating due to the lack of non-verbal behavior information [7]. Online meeting participants need supports on understanding others' engagement.

In this study, we demonstrate an application *Metacognition-EnGauge*. Figure 1 shows an application augmenting participants to understand self-and-group engagement levels during meetings. The engagement recognition uses *EnGauge* [6], the deep-learning-based engagement levels prediction model. Our study has three research questions,

RQ1 Does self-metacognition augmentation of engagement levels support participant to be engaged in the meetings?
RQ2 Does group-metacognition augmentation of engagement levels support participant to be engaged in the meetings?
RQ3 Do participants prefer none, self, or group-metacognition augmentation intervention in the online meetings?

By answering these questions, we will understand the potential approach for supporting online meeting participants to become more engaged in online communication.

2 RELATED WORK

Holstein et al. has proposed a method for the teachers to check the students' cognitive states as a dashboard [1]. In the dashboard, each student has an indicator display floating above the head. The work proposed visualizing multiple people's cognitive states combined as a dashboard. *Jumple* project aims to augment the virtual physical education experience [5]. The work supports visualized interactivity with remote participants. Niwa et al. state the importance of AI agents to augment human cognitive abilities [3]. Their approach is to implement a similar-looking AI avatar as a participant to investigate the trustworthiness level of agent statements. Lastly, Kytö et al. has proposed a method for visualizing public and personal information using digital profiles [2]. The work states that giving meta-information about a person communicating will support face-to-face communication smoothly. Concerning previous research, augmentation of the cognitive state positively supports change in behaviors and understanding humans.

3 METHODOLOGY

3.1 Engagement levels estimation model

This study uses the engagement level detection model EnGauge [6]. The model uses 24 participant data of high, middle, and low engagement levels. The data is collected by a role-playing approach. High-level engagement participants focus only on speaking, middle-level engagement participants focus on speaking and taking notes, and low-level engagement participants ignore the conversation and focus on reading a document. The prediction pipeline uses a facial cropped image into MobileNetV2 [4] based deep-learning model to classify the engagement.

3.2 Architecture of the application

The engagement level estimation model proposed in Section 3.1 can predict three-level engagement classified results. In order to make the result reliable in real-time and sequential, we applied time-concerned engagement level percentage prediction. Our application collects facial image from a webcam. We utilize HTML and JavaScript to capture a webcam frame image at five-second intervals during online meetings by recording the video screen-view. The captured webcam frame image is then post request to the engagement detection model mounted server. The server then predicts the requested participant webcam frame image into a high, middle, or low state of engagement level. This result is then stored into an array format as a result of engagement prediction in time-series (X_i). The engagement level estimation in percentage is done using following formula:

$$Y(\%) = \frac{\sum_{i=1}^{N} X_i(\%)}{N} \quad (1)$$

Y represents the result of the engagement level in percentage. It will be a response from the server towards the client application. X_i is an engagement classification result, it is either 100% (high engagement), 50% (middle engagement), or 0% (low engagement). N is the number of elements. In this pilot application we choose $N = 5$, which means the last five results of engagement classification inside X_i used to calculate Y. This approach supports gaining engagement-level prediction results into a time-sequential concern output. Once the client application receives a response from the server, the client application moves the gauge interface needle regarding the value. The higher the engagement, the more the gauge needle moves towards 180 degrees of the gauge, and when it gets lower, it shifts to zero degrees. Therefore, the participants can check the prediction result every five seconds as feedback. We extracted the engagement level result into CSV to compare participant engagement levels in each experiment condition.

4 DATA COLLECTION

4.1 Participants

We recruited 18 participants (Four females, 13 males, and one prefer not to say). Nationalities are ten Japanese, six Indian, one German, and one Chilean. Participants were between 23 and 31 years old, and the mean age was 26 years old. They are either workers or university students in Japan and Germany. We obtained consent from all participants before the experiment. For participants in Germany, including general data protection regulation (GDPR). All participants were allowed to opt out of the experiment at anytime.

4.2 Experiment Condition

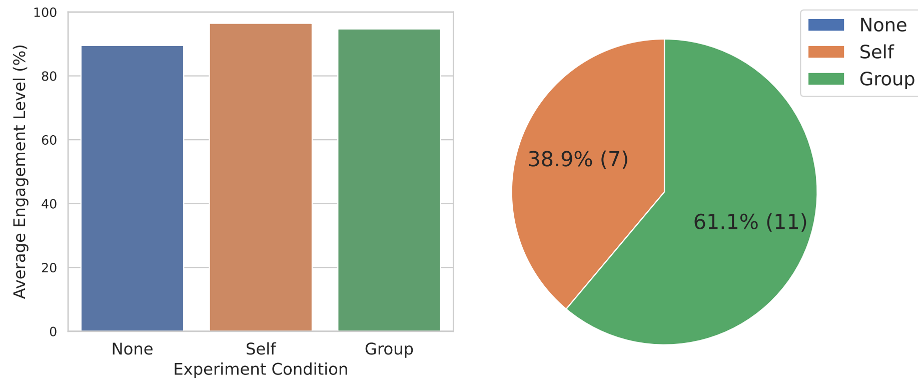
In this study, we request participants to join from anywhere in remote conditions. In each experiment, three participants are asked to be in a group per session. In each session, we asked the group to discuss for ten minutes. We asked participants to join under three different online meeting conditions. Condition 1 (C-1) is an online meeting without *EnGauge* intervention. Condition 2 (C-2) is an online meeting with self-metacognition augmentation by *EnGauge* intervention. Condition 3 (C-3) is an online meeting with group-metacognition augmentation by *EnGauge* intervention.

Figure 1 explains the conditions of C-2 and C-3. The self-metacognition augmentation gives one gauge interface feedback to the participant. The group-metacognition augmentation gives all (three) participants' engagement gauge interface feedback to the participant. We asked participants to work on *would you rather questions* for the group discussion. We prepared three different topics. Topic A (T-A) asked to discuss "Would you rather live in a hot or cold country?". Topic B (T-B) asked to discuss "Would you rather choose to gain physical strength or brain intelligence?". Topic C (T-C) asked to discuss "Would you rather go to the future or the past?".

These topics were chosen because they are open questions for anyone to answer from their perspective. We were also required to ask participants to give supportive reasons for their decision. Therefore, participants will keep on discussing until the ten minutes session time is over. The topic order is counterbalanced between groups to avoid differences in engagement based on topic interest.

4.3 Post Survey

Participants were asked to answer post-survey after each of the three experimental conditions. For post-survey, we asked "What was the conclusion against the topic?", "Please give the supportive reasons for the conclusion.", "In your



(a) Average engagement level of participants in each interventions.

(b) Participant preference for the intervention.

Fig. 2. Comparison of engagement in each different intervention conditions. Overall average engagement for all participants is highest for self-metacognition augmentation intervention. The participant preference is on group-metacognition augmentation intervention.

perspective, how engaging were you in the meeting? (Scale of 0 to 10)", "How was the online meeting application? (Scale of 0 to 10)", and "Write any comments in general about the experiment." for all condition questionnaires. After each C-2 and C-3, we asked "How was the *Metacognition-EnGauge* system? Any difficulties? (Scale of 0 to 10)", "Did you have any change in behaviors when using the application?", "Can you give us opinions on improving the application?". After all experiment conditions, we asked "Which condition do you prefer? (Choose one from C-1, C-2, or C-3)".

5 RESULT AND DISCUSSION

In this section, we state the results of the experiment and discuss the answers to three research questions. Figure 2 shows the average engagement levels of participants in each conditions, and the participant preference of the intervention. For research question 1 and 2, self-and-group metacognition augmentation intervention both made average engagement level during online meeting higher than having no intervention. Comparing self and group, the engagement became highest for the self-metacognition augmentation. Towards the answer for the research question 3, we found that majority of participant prefer group-metacognition augmentation.

6 CONCLUSION

In this work, we verified the effectiveness of real-time augmentation of self-and-group engagement levels understanding using *Metacogniton-EnGauge*. Our study shows that self-and-group engagement levels feedback by gauge interface improve engagements throughout the meetings. The average engagement through the meeting performed best for self-metacognition augmentation. According to the survey, we found that most participants prefer group-metacognitive feedback.

ACKNOWLEDGMENTS

This work was supported by DFG International Call on Artificial Intelligence "Learning Cyclotron" (Project Number: 442581111).

REFERENCES

[1] Kenneth Holstein, Gena Hong, Mera Tegene, Bruce M. McLaren, and Vincent Aleven. 2018. The classroom as a dashboard: co-designing wearable cognitive augmentation for K-12 teachers. In *Proceedings of the 8th International Conference on Learning Analytics and Knowledge* (Sydney, New South Wales, Australia) *(LAK '18)*. Association for Computing Machinery, New York, NY, USA, 79–88. https://doi.org/10.1145/3170358.3170377

[2] Mikko Kytö, Ilyena Hirskyj-Douglas, and David McGookin. 2021. From Strangers to Friends: Augmenting Face-to-face Interactions with Faceted Digital Self-Presentations. In *Proceedings of the Augmented Humans International Conference 2021* (Rovaniemi, Finland) *(AHs '21)*. Association for Computing Machinery, New York, NY, USA, 192–203. https://doi.org/10.1145/3458709.3458954

[3] Masayasu Niwa, Katsutoshi Masai, Shigeo Yoshida, and Maki Sugimoto. 2023. Investigating Effects of Facial Self-Similarity Levels on the Impression of Virtual Agents in Serious/Non-Serious Contexts. In *Proceedings of the Augmented Humans International Conference 2023* (Glasgow, United Kingdom) *(AHs '23)*. Association for Computing Machinery, New York, NY, USA, 221–230. https://doi.org/10.1145/3582700.3582721

[4] Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, and Liang-Chieh Chen. 2018. Mobilenetv2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 4510–4520.

[5] Soohyun Shin, Jaekyung Cho, and Seong-Woo Kim. 2021. Jumple: Interactive Contents for the Virtual Physical Education Classroom in the Pandemic Era. In *Proceedings of the Augmented Humans International Conference 2021* (Rovaniemi, Finland) *(AHs '21)*. Association for Computing Machinery, New York, NY, USA, 268–270. https://doi.org/10.1145/3458709.3458964

[6] Ko Watanabe, Tanuja Sathyanarayana, Andreas Dengel, and Shoya Ishimaru. 2023. EnGauge: Engagement Gauge of Meeting Participants Estimated by Facial Expression and Deep Neural Network. *IEEE Access* (2023).

[7] Ko Watanabe, Yusuke Soneda, Yuki Matsuda, Yugo Nakamura, Yutaka Arakawa, Andreas Dengel, and Shoya Ishimaru. 2021. Discaas: Micro behavior analysis on discussion by camera as a sensor. *Sensors* 21, 17 (2021), 5719.

Received 5 March 2024; accepted 5 March 2024