WIP: Relationship Between Eye Gaze Points and Comprehension Level in On-Demand Learning Contents in Elementary School Students

Satori Hachisuka
Interfaculty Initiative in Information Studies
The University of Tokyo
Chiba, Japan
s_hachisuka@iii.u-tokyo.ac.jp

Kayoko Kurita
Center for Research and Development of Higher Education
The University of Tokyo
Tokyo, Japan
kurita@he.u-tokyo.ac.jp

Shin'ichi Warisawa
Graduate School of Frontier Sciences
The University of Tokyo
Chiba, Japan
warisawa@edu.k.u-tokyo.ac.jp

***Abstract*— This work-in-progress research paper aims to elucidate the appropriate form of teacher images in online learning content and establish the knowledge necessary to construct content with learning effects. Our previous research found that it is adequate to maintain and improve the level of understanding when the presence or absence of the teacher's image matches the learners' preferences. Therefore, our research question was whether a similar trend could be observed among the younger generation as elementary school students. Furthermore, we hypothesized that students would be more likely to prefer content that presented an image of a teacher and that they would spend a lot of time fixating on the teacher's image. We experimented to verify this hypothesis. In this paper, we conducted quizzes for 21 elementary school students (4th to 6th grade) that simulated on-demand learning. Then, we calculated the correct answer rates for the quizzes and measured the eye gaze points during the quizzes. The quizzes consisted of five multiple-choice questions that included common sense and riddles. In addition, two types of content were created and used: one that included the teacher's face image and one that did not. We got the result when the presence or absence of a teacher image in the video content matched the participant's individual preferences, and learning effects improved until four weeks after the experiment. Additionally, the participants whose correct answer rates improved from the first quizzes gazed longer time rate at the options area during the explanation sections. According to the results, we consider it essential for students to solve the questions themselves again while watching the explanation's contents with the teacher's image matched their preferences to enhance comprehension of the teacher's explanations.**

Keywords— answer correctness, comprehension level, eye gaze tracking, learning effect, on-demand learning

I. INTRODUCTION

From 2020 to early 2023, a period marked by the global imperative to curb the spread of coronavirus infections, online education emerged as a widespread alternative. As the world gradually eases pandemic-related restrictions recently, educational settings are navigating a return to pre-pandemic norms, but some institutions continue to embrace online learning. These differences in trends are often based not only on countries and regions but also on school classifications such as universities, high schools, junior high schools, and elementary schools, or on school-specific policies. Though we understand the importance of a face-to-face educational environment, we propose not a simple return to a face-to-face format across the board but a careful identification of essential elements of how online classes should be conducted and their practical use, and we want to come up with guidelines for this. We believe this is an important research topic that can only be accomplished with a generation that has received face-to-face and online education.

Mayer (2001) [1] proposed an effective method for teaching using multimedia, which is currently used in many educational environments and describes the role of teachers in both face-to-face and online formats. Furthermore, even before the spread of the coronavirus infection, Bhat et al. (2015) [2] and Pi et al. (2017) [3] were conducting research on how teachers can project their image in online education materials, such as Massive Open Online Course (MOOC). There has been debate about the teacher's projection method and size. Furthermore, Goh et al. (2017) [4] and Colliot and Jamet (2018) [5] also show that the presence or absence of teacher images does not affect the learning effect. Still, the experimental participants were university students or older in both cases. The impact on learners below high school age has yet to be clarified.

The authors have previously investigated the influence of the presence or absence of teacher images in on-demand learning materials on learning motivation and learning effectiveness for junior high and high school students. In particular, an experiment with 53 junior high school students revealed that when the presence or absence of teacher images matches learners' preferences, it effectively maintains and improves understanding [6]. Therefore, in this paper, we experimented with elementary school students younger than junior high school students to examine whether a similar trend could be observed. We also conducted eye gaze measurements to explore how elementary school students looked at the teacher's facial image when viewing on-demand teaching materials and understand the contents and improved comprehension. We hypothesized that elementary school students, younger than junior high school students, would prefer content that displays teacher images and would spend a longer time fixating on images of teachers.

II. Experiment Method

In this study, we experimented using videos containing quiz questions and explanations that we created independently as on-demand learning materials. The quiz content consisted of the same questions as those we had previously given to junior high and high school students to compare the results' differences between age groups. We prepared two types of videos, one with the teacher's face image displayed and the other without the image of the teacher's face, and the percentage of correct answers to the quiz and the level of understanding of the explanations of the group that watched each video were measured. Participants took the first quizzes (primary) while watching the video contents, and they took similar quizzes (secondary) on the same day. After four weeks, they took the same quizzes (primary) as the first time online. We calculated the learning retention rate based on viewing the explanatory videos. Details of the experimental settings are shown below.

A. Participants

We adopted 21 healthy elementary school students as experimental participants. We determined that participants met the requirement of being healthy based on guardians' and participants' self-reporting that they were participating in regular classes without any special support. The participants were 48% male (10 students) and 52% female (11 students). The composition of the grades was 38% (eight students) in 4th grade, 29% (six students) in 5th grade, and 33% (seven students) in 6th grade. The participants were 10 to 12 years old, and they participated in the experiment based on their own free will, with the consent of their guardians.

B. Procedure

The experiments consisted of two days, and the interval between the first and second day was four weeks (Fig. 1). On the first day, we conducted the following steps in the face-to-face format: (1) Informed consent, (2) An explanation of the experiment, (3) Quizzes (primary), (4) Semi-structured interview, and (5) Similar quizzes (secondary). On the second day, four weeks after the first day of the experiment, we conducted the same quiz as (3) quiz (primary) on the first day in an online environment using Google Forms. On the first day, (2) explanation of the experiment and (3) implementation of the quiz (primary) were reproduced by playing videos that imitated online teaching materials. Two types of videos were prepared, one with the teacher image displayed and one without, and the experiment participants were set to watch one of them. The only difference between the two videos was the presence or absence of the teacher image as visual information, and the same material was used for the audio information. Also, the implementation of (3) quizzes (primary) on the first day was divided into three parts: (3-1) quiz questions, (3-2) answering time, and (3-3) correct answers and explanations (Fig. 2). By listening to the correct answers and explanations (3-3), the experiment participants could understand whether their answers were correct or incorrect and how to solve the problem. Furthermore, during the quizzes (primary) section, the gaze points of the participants were measured using a non-contact, calibration-free eye mark recorder (EMR ACTUS, nac Image Technology Inc.) We conducted this experiment with the approval of the University of Tokyo Ethics Review Committee (Approval No. 21-123).

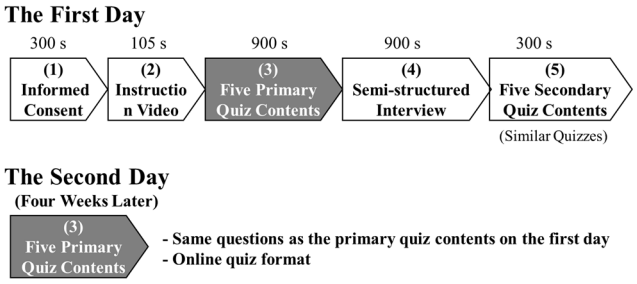

Fig. 1. The experiment procedure of the first and second days. The contents of (5) the Secondary Quiz were similar to (3) the Primary quiz; therefore, participants could solve the questions if they understood the explanation in the video.

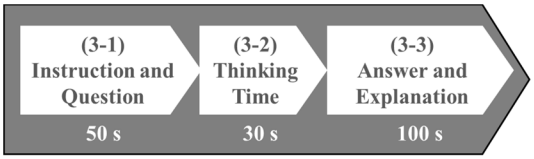

Fig. 2. The flow of one quiz content. The quiz consisted of three parts: (3-1) quiz questions, (3-2) answering time, and (3-3) correct answers and explanations. The experiment participants could understand how to solve the quiz from the explanation in (3-3).

C. Quiz Contents

We asked five questions in each of the primary and secondary quiz sections. To avoid the direct influence of the experiment participants' strengths and weaknesses in the class subjects, the contents of the quizzes were designed based on common sense or on-the-spot inspiration. For example, the order of the three colors of general vehicle traffic lights and their reason (Fig. 3). Participants in the experiment chose one option out of two to four that they thought was correct and answered verbally, along with the reason for their choice. The order of questions was counterbalanced within the experiment participants. The reason for giving verbal answers was to avoid interfering with eye gaze measurement and to allow experiment participants to think and answer while looking at the content presentation display.

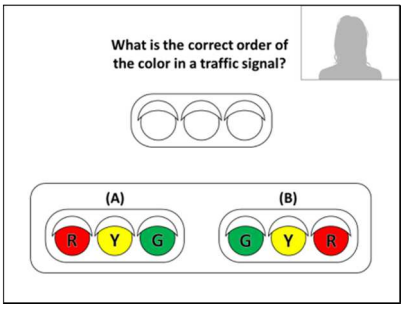

Fig. 3. The example of displayed quiz content with a teacher image. Japanese language was used in the actual experiment. The participants in the experiment chose one option that they thought was the correct answer from among the options and orally answered the question along with the reason for the choice.

D. Data Analysis

a) Learning Retention Rate: Regarding the correctness of the answer to the quiz, we defined it as not only the correct choice but also the reason for selecting the correct option. This definition excludes cases where a person answers correctly by chance without knowing the correct cause. This paper used the percentage of correct answers to quizzes based on this definition as an index for data analysis. For the quizzes four weeks later, the answer was considered correct only if the answer was correct and the participant remembered the explanations or answers they had watched on the first day of the experiment, and cases, where the answer was correct by chance, were excluded. Data regarding teachers' preferences for the presence or absence of images were classified based on the responses obtained during the semi-structured interview. The quiz answer rate after four weeks was 100%, so in this paper, we analyzed the learning retention rate based on the correct answer rate using data from all 21 students.

b) Eye Gaze Measurement: Regarding the fixation points, we focused on the change in the correct answer rate before and after watching the explanatory video on the first day of this experiment. We compared the fixation points of the group whose correct answer rate improved before and after viewing and the group whose correct answer rate did not improve. We conducted this analysis for 11 people with images of the teacher in the experimental content. As an analysis method, we divided the areas in the explanatory video into five regions (Text, Explanatory illustrations, Teacher's face image, Multiple choices, and Others) and calculated the proportion of gaze time for each area.

III. Results and Considerations

As a result of the impression evaluation obtained from the interviews, 48% of the participants in this experiment preferred the presentation of teacher's facial images in the learning content, almost the same as 45% in a previous study targeting junior high school students. Furthermore, the most common reason for not liking the presentation of facial images was that it interfered with concentration, accounting for 64% of the experiment participants who said they did not like the presentation of facial images in the semi-structured interview.

a) Learning Retention Rate: First, to confirm whether the presence or absence of a teacher image was related to the correct answer rate, the average correct answer rate between groups with and without a teacher image in each quiz content was compared by *t*-test (Fig. 4). As a result, there were no significant differences in accuracy regarding whether the teacher image was presented or not in either quiz content. This result shows the same tendency as when we conducted previous experiments targeting junior high school students.

Next, in this paper, to confirm the result obtained in our previous research that the learning retention rate improves when learners' preferences match the presence or absence of teacher image display, we divided 21 students according to the interview data. The first group was a group whose preference for the presence or absence of a teacher image obtained in the interview matched the presence or absence of a teacher image in the content actually presented during the experiment (consistent with preference group), and the other group did not (inconsistent with preference group). The Bonferroni multiple comparison test results for these two groups are shown in Fig. 5. In both groups, the correct answer rates of the secondary quiz and the quiz conducted four weeks later were significantly higher than the initial rate of the primary quiz conducted before presenting the commentary content. However, when comparing the results of the secondary quiz shown in the center of the graph in Fig. 5 with the results of the primary quiz administered four weeks later shown on the right side of the graph, a significant performance improvement was observed only in the group the teacher image existence matched the participants' preferences. The results showed that the correct answer rate improved over time when a display method that matched the preferences of the experiment participants was used. This result suggests that it is preferable to match the presence or absence of a teacher image in online learning content to the learners' preferences. However, in this paper, 16 people (76%) in the group matched the preference, and five (24%) did not, resulting in a biased number.

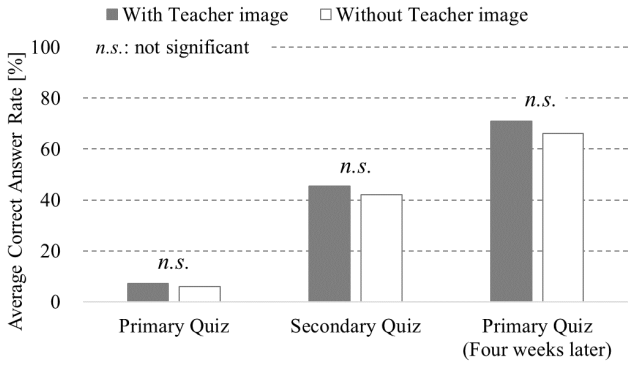

Fig. 4. The average correct answer rate between groups with and without a teacher image in each quiz. There were no significant differences in answer correctness regarding whether the teacher image was presented or not in each quiz.

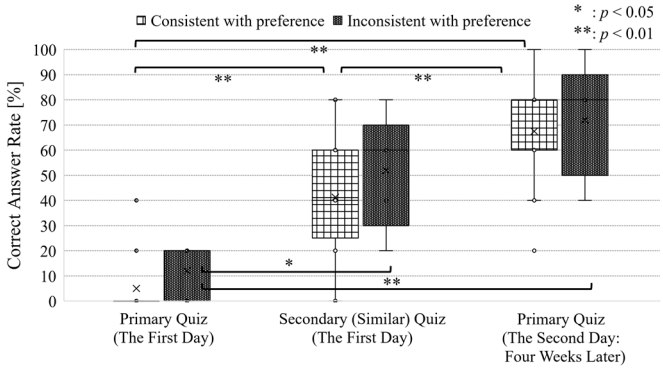

Fig. 5. Correct answer rates of the group which teacher image existence match the preference of participants (consistent with preference group) and didn't match the participants' preference (inconsistent with preference group). The correct answer rate improved over time when a display method that matched the preferences of the experiment participants was used.

b) Eye Gaze Measurement: Fig. 6 shows the average gaze time ratio of the gaze areas of the group whose correct answer rate improved and the group whose correct answer rate did not improve before and after viewing the explanation video. We compared the time ratio results between the primary and secondary quizzes on the first day. Based on this result, when the proportion of gaze time was compared between the two groups using a *t*-test, no significant difference was found in the proportion of gaze time to the teacher image. This result shows a similar trend to the learning retention rate results. On the other hand, a significant difference was observed in the multiple choices area at a significance level of 1%. These results suggest that looking at the options while watching the explanation improves the learning effect. This may be because the students were re-solving the problem themselves while watching the explanation. We suggested this as a new hypothesis.

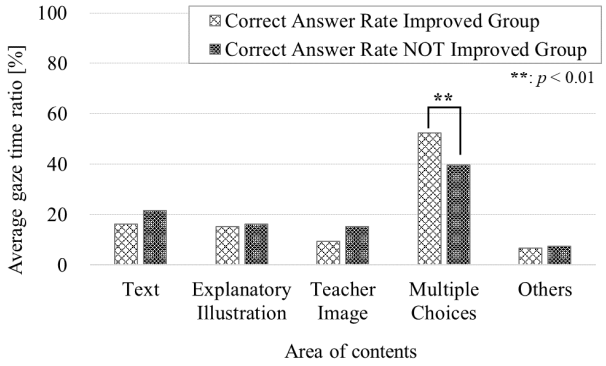

Fig. 6. The average gaze time ratio of the gaze areas of two groups. One group consisted participants whose correct answer rate improved between the primary quiz and secondary quiz on the first day. The other group didn't improve the answer correctness.

IV. Conclusion and Future Work

The results of this paper show that when considering learning content for elementary school students, adjusting the presence or absence of images of teachers in the learning content to suit the learners' preferences may contribute to improving learning effectiveness. On the other hand, the interviews revealed that when presenting images of teachers, it is necessary to devise measures that do not interfere with understanding the learning content. Furthermore, even when a teacher image is presented, it is considered essential to encourage the user to look at the options and solve the problem again when viewing the explanation. The limitations of this paper are the small number of participants and the bias in the number of participants between the analyzed groups. To confirm the credibility of the results obtained in this paper, increasing the number of participants is necessary and crucial for the potential impact of this research.

In our future work, we aim to increase the number and diversity of experimental participants further to examine the applicability to actual learning environments and learning subjects. Online education generally involves learners simultaneously watching the same standardized content, including teacher images. This method lacks the individual support that teachers provide to students in a traditional classroom environment. In other words, it is hard for teachers to offer support based on each learner's personality traits and conditions in general online group lessons by timing their responses accordingly. Therefore, by advancing this research, it is expected that even in online education, rather than simply providing uniform content, it will be possible to adjust the presence or absence of teacher images, as well as the timing and position of their display, thereby improving students' learning effectiveness and engagement. Additionally, by further examining diversity, it may be possible to establish a method of displaying teacher images appropriate for different age groups and grade levels. Furthermore, this will lead to individually optimized learning for learners who require exceptional support, considering their preferences and physical and mental diversity.

References

[1] R. E. Mayer, Multimedia learning. Cambridge University Press, United Kingdom, 2001.

[2] S. Bhat, P. Chinprutthiwong, M. Perry, "Seeing the instructor in two video styles: preferences and patterns," Proceedings of the 8th International Conference on Educational Data Mining, 2015, pp. 305-312.

[3] Z. Pi, J. Hong, and J. Yang, "Does instructor's image size in video lectures affect learning outcome?" in Journal of Computer Assisted Learning, 33, 2017, pp. 347-354.

[4] W. W. Goh, E. Nyub, S. Y. Won, and C. L. Lim, "The importance of teacher's presence and engagement in MOOC learning environment: A case study," 2017 IEEE Conference on e-Learning, 2017, pp. 127-132.

[5] T. Colliot, E. Jamet, Understanding the effects of a teacher video on learning from a multimedia document: an eye-tracking study, Educational Technology Research and Development, 66, 2018, pp. 1415-1433.

[6] S. Hachisuka, K. Kurita, and S. Warisawa, "Relationship between preference for the existence of teacher images and learning effects in online learning for junior high school students," Conference Proceedings of IEEE International Conference on Teaching, Assessment and Learning for Engineering, 2023, pp. 247-250.