Exploiting Object-of-Interest Information to Understand Attention in VR Classrooms

Efe Bozkir*†
Human-Computer Interaction
University of Tübingen

Philipp Stark*‡
Human-Computer Interaction
Hector Research Institute of Education Sciences and Psychology
University of Tübingen

Hong Gao§
Human-Computer Interaction
University of Tübingen

Lisa Hasenbein¶
Hector Research Institute of Education Sciences and Psychology
University of Tübingen

Jens-Uwe Hahn‖
Hochschule der Medien Stuttgart

Enkelejda Kasneci**
Human-Computer Interaction
University of Tübingen

Richard Göllner††
Hector Research Institute of Education Sciences and Psychology
University of Tübingen

Abstract

Recent developments in computer graphics and hardware technology enable easy access to virtual reality headsets along with integrated eye trackers, leading to mass usage of such devices. The immersive experience provided by virtual reality and the possibility to control environmental factors in virtual setups may soon help to create realistic digital alternatives to conventional classrooms. The importance of such settings has become especially evident during the COVID-19 pandemic, forcing many schools and universities to provide the digital teaching. Researchers foresee that such transformations will continue in the future with virtual worlds becoming an integral part of education. Until now, however, students' behaviors in immersive virtual environments have not been investigated in depth. In this work, we study students' attention by exploiting object-of-interests using eye tracking in different classroom manipulations. More specifically, we varied sitting positions of students, visualization styles of virtual avatars, and hand-raising percentages of peer-learners. Our empirical evidence shows that such manipulations play an important role in students' attention towards virtual peer-learners, instructors, and lecture material. This research may contribute to understanding of how visual attention relates to social dynamics in the virtual classroom, including significant considerations for the design of virtual learning spaces.

Index Terms: Human-centered computing—Human computer interaction (HCI)—Empirical studies in HCI—; Computing methodologies—Computer graphics—Graphics systems and interfaces—Virtual reality; Applied computing—Education—Interactive learning environments—; Applied computing—Education—Computer-assisted instruction—

1 Introduction

Everyday use of head-mounted displays (HMDs) is increasing as virtual reality (VR) technology and virtual environments are already being used in various domains such as gaming and entertainment. In addition, some of the consumer-grade HMDs are coming to market with integrated eye trackers that may help to assess human attention during immersion and allow for more interactive virtual environments. It is likely that, in the near future, such tools will become widely used mobile devices similar to today's mobile phones or smart watches. To this end, not only should researchers strive to improve the capabilities of these devices, but scrutiny should also be given to understanding human behavior and attention while using such technology.

Measures of eye movements obtained through eye-tracking are effective indicators of human states and visual behavior to some extent; however, they are dependent on application or task [17]. Analyzing and modeling human attention using this data in a specific domain may not be transferable to other domains. Thus, when assessing human attention in digital environments, or more particularly in VR for the application in educational technology, specific domain knowledge and configurations should be considered. There is already some history of training and teaching in digital or virtual setups [14, 19]. Today, due to the COVID-19 pandemic, virtual or digital education has become more popular and even a necessity in many cases. Currently, many schools and universities are carrying out their teaching responsibilities remotely via platforms such as Zoom[1] or Webex[2]. Such platforms lack the possibility of instructor-student interaction beyond audio and video features and encounter privacy concerns if videos are recorded and stored during classes. VR setups offer the immersion, interaction, and privacy preservation that current remote learning platforms lack. In addition, as VR allows users to easily control the environmental settings, it is possible to evaluate different classroom manipulations and subsequent effects on human behavior, a step that is exponentially more difficult in real world classrooms.

In this work, we exploit object-of-interest information by using eye-gaze and three main sets of objects in immersive VR. We focus on virtual peer-learners, virtual instructor, and screen to understand visual attention through the design of a virtual classroom and a lecture about computational thinking. We choose these objects-of-interests since they are of particular interest with regard to attention towards social dynamics and learning. Our study has three different design factors: Different sitting positions of participating students, different visualization styles of virtual avatars including an instructor and peer-learners, and different hand-raising behaviors of virtual peer-learners. Different sitting positions include seating participating students in the front or back of the virtual classroom. In addition, different visualization styles of avatars consists of two conditions

*Both authors contributed equally to this research.
†e-mail: efe.bozkir@uni-tuebingen.de
‡e-mail: philipp.stark@uni-tuebingen.de
§e-mail: hong.gao@informatik.uni-tuebingen.de
¶e-mail: lisa.hasenbein@uni-tuebingen.de
‖e-mail: hahn@hdm-stuttgart.de
**e-mail: enkelejda.kasneci@uni-tuebingen.de
††e-mail: richard.goellner@uni-tuebingen.de

[1]https://www.zoom.us/
[2]https://www.webex.com/

2642-5254/21/$31.00 ©2021 IEEE
DOI 10.1109/VR50410.2021.00085

that are cartoon- and realistic-styled avatars. Lastly, different hand-raising behaviors include 20%, 35%, 65%, and 80% of the peer-learners raising their hands to answer questions during the lecture. To the best of our knowledge, this is the first work that assesses students' attention by using object-of-interest information in an immersive VR classroom through the manipulation of sitting positions of students, visualization styles of peer-learners and instructor, and hand-raising behaviors of peer-learners collectively. Such manipulations may be important indicators of students' visual attention towards lecture contents and social dynamics in the classroom and should be taken into consideration when designing VR classrooms.

2 Related Work

Since our work benefits from VR in education and in eye tracking research, we discuss the state-of-the-art along these two lines. Various studies using VR in education settings assess the mechanisms of attention or social dynamics by using pre- or post-tests or by relying on head movement behavior as a proxy for gaze. Using eye tracking in addition to such information presents the possibility of a deeper understanding of visual and situational attention during immersive experiences.

2.1 Virtual Reality in Education and Classrooms

VR offers great promise for supporting teaching and learning procedures, especially when digital learning, physical inabilities, ethical concerns, and situational limitations are considered. An extensive review of immersive VR in education and its pedagogical foundations are discussed in [14] and [18], respectively. We focus on research on VR in education and immersive VR classrooms in this section.

The effectiveness of learning in virtual and augmented reality (VR/AR) compared to tablet-based applications and the impact of VR-based systems on students' achievements are studied in [30] and [2], respectively, and these works indicate several advantages of VR-based conditions. In addition, it has been found that students' motivation increases when VR is used as a teaching tool in art history [9] and social studies [11]. VR not only supports the effectiveness of learning, but also can improve instructor teaching skills [21].

Apart from VR applications in teaching and learning, the design and degree of realism in VR classrooms have also been studied. Presence of a virtual instructor was found to increase the engagement and progress of users [42]. Furthermore, the processes of synthesizing virtual peer-learners by using previous learner comments [25] and designing VR classrooms by replicating real conditions [40] which may affect learning are considered.

Several works focused on understanding visual attention and behavior in immersive VR classrooms. Bailenson et al. [4] and Blume et al. [5] studied learning outcomes according to sitting positions and offer compelling evidence that students seated in the front have better learning outcomes. Few studies, however, took head movements into consideration [13, 31, 34, 39] in such setups. In [13], the immersive VR classroom was used as a tool to study attention measures for attention deficit/hyperactivity disorder (ADHD), whereas in [31] reliability of virtual reality and attention was studied with continuous performance task (CPT) for clinical research. Social interaction using head movements was studied in [39] with users' head movements found to shift between the interaction partner and target. Some studies argued for eye tracking measurements, especially in clinical research for diagnosis or attention related tasks [27, 33]. However, none of the previous works have focused on social interactions and dynamics in the immersive VR classroom in an everyday setting by using object-of-interest information and eye movements.

2.2 Eye Tracking in Virtual Reality

Eye tracking and gaze estimation are considered challenging tasks in a real world setting because it is difficult to control factors such as occlusions or illumination changes [16, 47]. However, in most of the VR setups, eye trackers are located inside of HMDs. This creates not only a more controlled and reliable environment for eye tracking, but also provides a unique opportunity to analyze and process human visual behavior during the VR experience.

Eye tracking has been used in many applications and shown to be helpful for various tasks in VR such as guiding attention in panoramic videos using central and peripheral cues [37], predicting motion sickness by using 3D Convolutional Neural Networks [24], synthesizing personalized training programs to improve skills [22], foveated rendering using saccadic eye movements and eye-dominance [3, 29], evaluation and diagnoses of diseases such as Parkinson's disease [32], re-directed walking using blinking behavior [23], or continuous authentication using eye movements [48]. While these works have used either the eye tracking or gaze data to derive more meaningful information for related tasks, assessing visual attention via eyes and gaze-based interaction is more relevant for classroom setups in particular. Bozkir et al. [6] assessed visual attention using gaze guidance and pupil dilations in a time-critical situation, whereas Khamis et al. [20] discussed gaze-based interaction using smooth pursuit eye movements in VR. In addition, Sidenmark and Lundström [41] analyzed eye fixations on interacted objects during hand interaction in VR and found that interaction with stationary objects may be favorable. Aforementioned works indicate that eye movements can be used reliably in VR setups. Moreover, considering that the majority of objects in a classroom are stationary or have limited spatial movement, visual attention extracted from such data may provide valuable insight into human behavior. While exploiting objects-of-interests could be considered as a primitive task, it forms the foundation of more complex tasks necessary to understand visual attention.

3 Methodology

The main focus of this work is to investigate object-of-interest information in different manipulations of an immersive VR classroom. We focus on three objects that may be considered as the most important objects in the current setup, namely peer-learners, instructor, and screen.

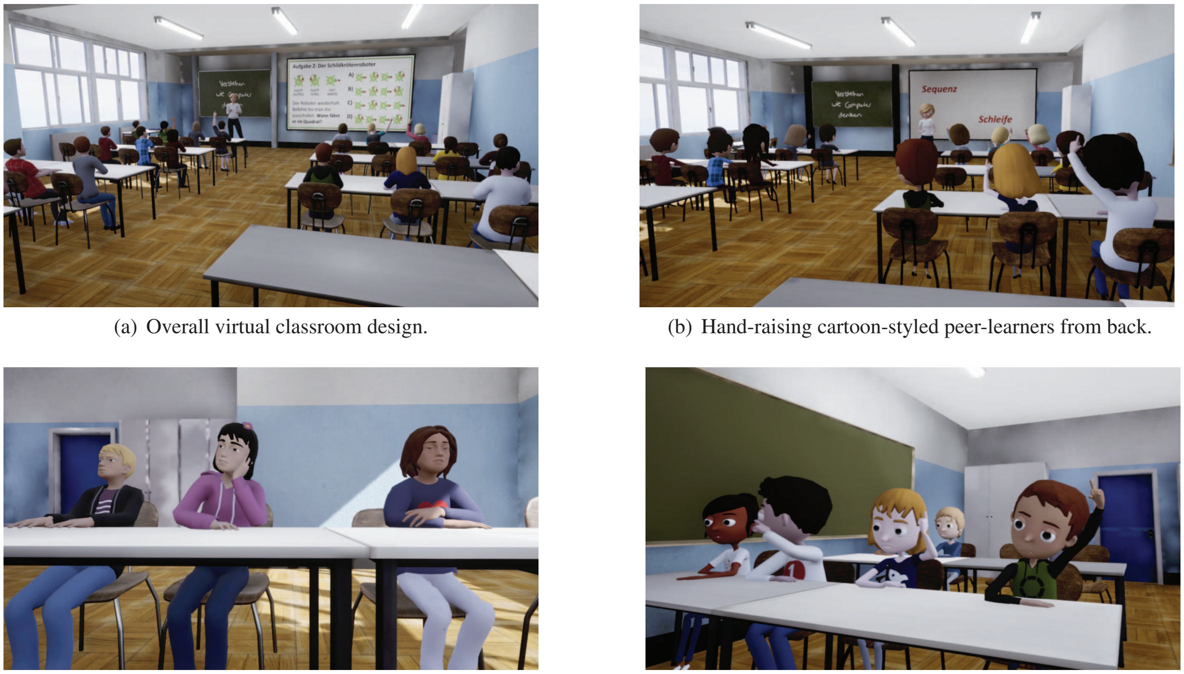
3.1 Participants

381 volunteer sixth-grade students (179 female and 202 male) between 10 to 13 years old ($M = 11.5$, $SD = 0.6$) were recruited for the experiment. In this age group, students are able to use an HMD, but do not have much experience with VR. They also had no background knowledge about the lecture content. Data from 101 participants were removed due to hardware related problems, incorrect calibration, low eye tracking ratio (lower than 90%), and synchronization issues. The average number of participants per condition was 17.5 ($SD = 5.2$). Finally, we used the data of 280 participants (140 female and 140 male) with the aforementioned average age and standard deviation. For each condition group separately, participants' gender was also equally distributed ($M = 0.58, SD = 0.08$). The study was approved by the ethics committee of the University of Tübingen prior to the experiments. Participants and their parents or legal guardians provided written informed consent in advance.

3.2 Apparatus

For the experiments, HTC Vive Pro Eye devices with integrated Tobii eye trackers were used. The HTC Vive Pro Eye has a refresh rate of 90 Hz and field of view of 110°. The integrated eye tracker has 120 Hz sampling rate. The screen resolution per eye was set to 1440×1600. Unreal Game Engine v4.23.1[3] was used to render the virtual classroom.

[3]https://www.unrealengine.com/



(a) Overall virtual classroom design.

(b) Hand-raising cartoon-styled peer-learners from back.

(c) Realistic-styled peer-learners.

(d) Hand-raising cartoon-styled peer-learners.

Figure 1: Views from the virtual classroom.

3.3 Experimental Design

The virtual classroom consists of 4 rows of desks organized in 2 columns. Next to each desk, chairs are located to let virtual peer-learners sit. There are 24 virtual peer-learners in the environment and all of them sit on chairs during the entirety of the lecture. Some of the chairs are kept empty so as not to overcrowd the virtual classroom. In addition, the virtual classroom includes other objects, which exist in real classrooms such as board, screen, cupboard, clock, and windows. The lecture content is visualized on the white screen. Additionally, the virtual instructor walks around the podium, replicating behavior similar to that of a real instructor. Fig. 1 (a), (b), (c), and (d) show the overall design, hand-raising peer-learners, realistic-styled peer-learners, and cartoon-styled peer-learners, respectively.

The content of the virtual lecture is about computational thinking [44] and the lecture takes $\approx$ 15 minutes in total, including 4 phases. These four phases are grouped as "Introduction to the topic", "Knowledge input", "Exercises", and "Summary" and take $\approx$ 3, $\approx$ 4.5, $\approx$ 5.5, and $\approx$ 1.5 minutes, respectively. The topic of the virtual lecture is visible on the board as "Understanding how computers think". The first phase starts with the virtual instructor entering the classroom. After staying for a while, the instructor leaves the classroom for about 20 seconds. During this time, participants have the opportunity to explore the classroom, look around, and acclimate themselves with the virtual environment. During the initial phase of the lecture, the instructor asks five questions, and some of the virtual peer-learners raise their hands to interact. In the second phase, the instructor describes two terms, "sequence" and "loop", and shows these terms on the white screen. After the descriptions, the instructor asks four questions about each term and some of the peer-learners raise their hands to answer them. In the third phase, the instructor assigns two exercises and allows students some time to think about them. Later, choices for each exercise are provided by the instructor and, this time, peer-learners raise their hands to vote on the correct answer out of the presented options. In the fourth phase, the instructor summarizes the lecture without asking any questions, which means that peer-learners do not raise their hands. In addition, no hand-raise is expected from the participants as hand poses are not measured during the experiments.

Our study is conceptualized in a between-subjects design. We evaluated three design factors, namely sitting positions of the participants, visualization styles of virtual avatars, and hand-raising percentages of virtual peer-learners. Participants were seated either in the front or back rows, which means that the participants seated in the front had one row in front of them, whereas participants seated in the back had three rows between them and the screen. Both conditions were aligned in the aisle side of the desks that were on the right side of the classroom. This manipulation can give insights about students' attention during a lecture, when they have either the overview over whole class and see most of their virtual peer-learners or when they are positioned closer to instructor and screen the lecture is presented on. Participants encountered either cartoon- or realistic-styled virtual avatars in the environment, including the virtual instructor and peer-learners. The cartoon-styled avatars have larger heads and tinier arms and legs as compared to the realistic-styled avatars. Since the animation and design of more realistic looking avatars is time and cost expensive, it should be interesting to investigate the impact of such manipulation. In addition, various hand-raising percentages of virtual peer-learners consist of four levels, namely 20%, 35%, 65%, and 80%. This means that when a question is asked during the lecture by the virtual instructor, a corresponding percentage of virtual peer-learners raise their hands to answer the question. The last two manipulations are of particular interest, regarding the question how social avatars should be designed in a virtual classroom and how they are perceived by students. Under which condition do students use social information and how does visualization and certain behaviour influence students attention. This helps to simulate and evaluate social dynamics and engagement during the virtual lecture using visual attention. In total, our 2 (factor 1) $\times$2 (factor 2) $\times$4 (factor 3) between-subjects design leads to 16 treatment groups.

3.4 Procedure

In the beginning of the experiment, the assistants introduced the experiment and its process to the participants. Participants had the opportunity to familiarize themselves with the hardware and the VR environment. Afterwards, the actual experiment and data collection began. Firstly, the eye tracker was calibrated. Then, the experiment was started with assistants pressing a start button. At the end of the virtual lecture, the participants were told to take the HMD off by a message which was displayed in the virtual environment. Virtual lectures were carried out without any breaks. After watching the virtual lecture, participants filled out questionnaires about their perceived realism and experienced presence which were conceptualized for the VR classroom according to [26, 38].

Each session took ≈ 45 minutes in total. The experiments were carried out in groups of ten participants who were randomly allocated to one of the 16 treatment groups by using a random number generator to ensure the random distribution of conditions within groups. To maintain natural behavior, participants selected the physical seat in the experiment room freely without being informed about experimental conditions. Although research assistants helped with technical issues regarding the use of the HMD, participants were blinded to the true purpose and design of the study, as it was solely introduced as a learning experience.

3.5 Data Processing and Measurements

During the experiments, head location and pose, gaze, and eye related data along with experimental condition were collected. Head movements are particularly helpful for mapping eye-gaze in the virtual environment. These were saved in data sheets for each participant using anonymous identifiers which ensured the privacy of the participants.

As gaze data reported by the eye tracker can be affected negatively by blinks or noisy sensor measurements, we applied a linear interpolation on the gaze vectors to clean the data. Afterwards, using head pose and interpolated gaze data, we applied ray-casting [35] to map the gaze into the 3D virtual environment. The objects in the 3D environment are surrounded by dedicated colliders; therefore, we were able to calculate 3D gaze points and gazed objects using the procedure visualized in Fig. 2.

However, gazed objects may not directly represent visual attention as participants can gaze on some objects unconsciously for a very short time. To overcome this issue, we set an attention threshold of 200 ms, meaning that we count the objects as object-of-interest if participants stay with their gaze on the objects for at least the amount of the attention threshold. As we assume that both fixations and saccades can occur during attending one object, the selected threshold is larger than classical fixation thresholds applied in eye tracking literature for both conventional [36] or VR eye tracking [1] setups. While we also experimented with various threshold values, our results show similar trends across different thresholds.

In addition to the data related to visual attention, self-reported perceived realism and experienced presence were obtained at the end of the experiments with 4-point Likert scales ranging from 1 ("completely disagree") to 4 ("completely agree") with 6 (e.g., "I felt like the teacher and the classmates could be real people") and 9 (e.g., "During the virtual lecture, I almost forgot that I was wearing the VR glasses") items, respectively.

In this study, we focused on three main objects in the virtual classroom, namely peer-learners, virtual instructor, and screen, when we extracted object-of-interest information. We decided that these objects may have a significant impact on social dynamics in the classrooms and for overall course of lecture. In our analyses, the attention time on each peer-learner is aggregated and the object of "peer-learners" represents the aggregated object and related attention. In addition, in our classroom setup there is one board and one white screen behind the instructor as depicted in Fig. 1 (a). The lecture content is provided on the white screen only; therefore, in our analysis we refer to the white screen when mentioning screen object.

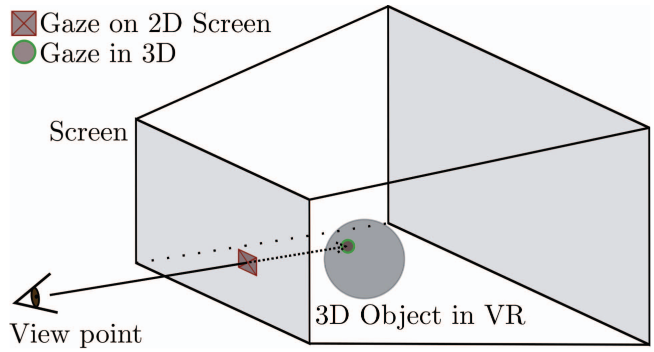


Figure 2: Ray-casting procedure to obtain 3D gazed object.

3.6 Research Hypotheses

Our hypotheses correspond to the experimental factors of sitting positions, avatar visualization styles, and various hand-raise percentages of virtual peer-learners, respectively. Furthermore, since we analyze behaviors towards three different objects in the virtual classroom, namely peer-learners, instructor, and screen, for simplicity we call attention to attending these objects-of-interests for the rest of the paper.

3.6.1 Visual Attention in Different Sitting Positions (H1)

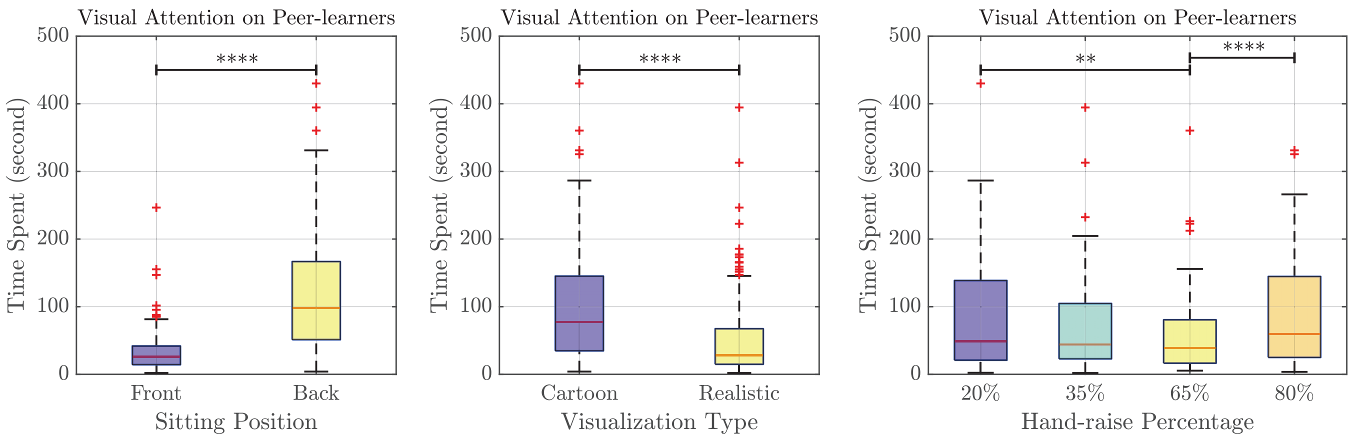
We expect that participants seated in the front condition have less attention on peer-learners, naturally because they do not have as many peer-learners sitting in front of them as opposed to the participants sitting in the back. In addition, the participants that are located in the front are closer to the virtual instructor and the screen that visualizes lecture content. Due to the proximity and having fewer moving and occluding objects in their field of view (FOV), we hypothesize that these participants have more attention time on both virtual instructor and screen than the participants sit in the back.

3.6.2 Visual Attention in Different Visualization Styles of Virtual Avatars (H2)

We hypothesize that attention time on peer-learners in the cartoon-styled visualization is longer than in the realistic-styled visualization as cartoon-styled peer-learners are more exciting for participants when ages of our interest group are taken into consideration. In addition, we assume that participants look at the realistic-styled instructor for longer than at cartoon-styled instructor as participants may consider the realistically rendered instructor more credible in a learning environment. Lastly, we do not expect any differences in terms of attention towards virtual screen that lecture content is visualized, as the visualization style of the screen does not change.

3.6.3 Visual Attention in Different Hand-raising Behaviors of Peer-learners (H3)

We hypothesize that attention time on peer-learners increases with a higher number of virtual peer-learners raising their hands when questions are asked, as this would create a visually more dynamic classroom. Additionally, we expect that if fewer virtual peer-learners raise their hands, this will lead participants to keep their attention either on the instructor or the lecture screen due to having less amount of visual distractors when questions are provided by the virtual instructor.



(a) Comparison between sitting positions. (b) Comparison between visualization types. (c) Comparison between hand-raising behaviors.

Figure 3: Attention towards virtual peer-learners for different classroom manipulation configurations. $**$ and $****$ correspond to the significance levels of $p < .01$ and $p < .0001$, respectively.

4 Results

In this section, we analyze the total amount of time spent on each object-of-interest (OOI), which we call visual attention, between different conditions. For each OOI, we applied a 3-way full factorial ANOVA for statistical comparison using alpha level of 0.05. For non-parametric analysis, we transformed the data using the aligned rank transform (ART) [46] before applying ANOVAs. For the pairwise comparisons, we used Tukey-Kramer post-hoc test as the sample sizes were not equal. While the main focus of this work is to assess visual attention using OOI information, here we report experienced presence and perceived realism questionnaires to support our main results. We obtained mean values of 2.91 for experienced presence and perceived realism with $SD = 0.55$ and $SD = 0.57$, respectively, without any significant differences between conditions.

4.1 Visual Attention on Peer-learners

Total time spent on peer-learners for different sitting positions, avatar visualization styles, and various hand-raising behaviors are depicted in Fig. 3 (a), (b), and (c), respectively. Total time spent on peer-learners is significantly longer in the back seated condition ($M =$ 115.07 sec, $SD = 85.28$ sec) than it is in the front seated condition ($M = 33.59$ sec, $SD = 32.45$ sec) with ($F(1,264) = 156.23$, $p <$ $.0001$, $\eta^2 = .36$).

Attention towards peer-learners as different visualization styled avatars differs significantly. Cartoon-styled peer-learners ($M =$ 98.67 sec, $SD = 82.79$ sec) drew significantly more attention than the realistic-styled peer-learners ($M = 55.28$ sec, $SD = 65.65$ sec) with ($F(1,264) = 54.13$, $p < .0001$, $\eta^2 = .17$).

Furthermore, for different hand-raising manipulations, attention time on the peer-learners differs significantly with ($F(3,264) = 6.93$, $p < .001$, $\eta^2 = .07$). Particularly, the total time spent on peer-learners in the 80% condition ($M = 88.95$ sec, $SD = 78.15$ sec) is significantly longer than in the 65% condition ($M = 59.23$ sec, $SD = 65.19$ sec) with ($F(3,264) = 6.93$, $p < .0001$, $\eta^2 = .07$). In addition, the total time spent in the 20% condition ($M = 88.62$ sec, $SD = 87.53$ sec) is significantly longer than in the 65% condition ($M = 59.23$ sec, $SD = 65.19$ sec) with ($F(3,264) = 6.93$, $p = .005$). In summary, attention time towards extreme levels of hand-raising percentages are longer than for intermediate levels.

Additionally, we found some significant interaction effects regarding the attention time on the peer-learners. The time on peer-learners in the hand-raising condition depends on the sitting position of the students with ($F(3,264) = 3.88$, $p = .0097$, $\eta^2 = .041$), as well as the attention time on peer-learners in the avatar visualization styles condition depends on the sitting position with ($F(1,264) = 11.37$, $p < .001$, $\eta^2 = .039$) and vice versa. A small interaction effect was found between the hand-raising condition and the avatar visualization styles with ($F(3,264) = 3.36$, $p = .02$, $\eta^2 = .036$).

4.2 Visual Attention on Instructor

Total time spent on instructor for different sitting positions, avatar visualization styles, and various hand-raising behaviors are depicted in Fig. 4 (a), (b), and (c), respectively. The participants that are seated in the front ($M = 190.07$ sec, $SD = 93.13$ sec) attended to the virtual instructor significantly more than the participants seated in the back ($M = 80.37$ sec, $SD = 60.78$ sec) with ($F(1,264) = 144.16$ $p < .0001$, $\eta^2 = .34$).

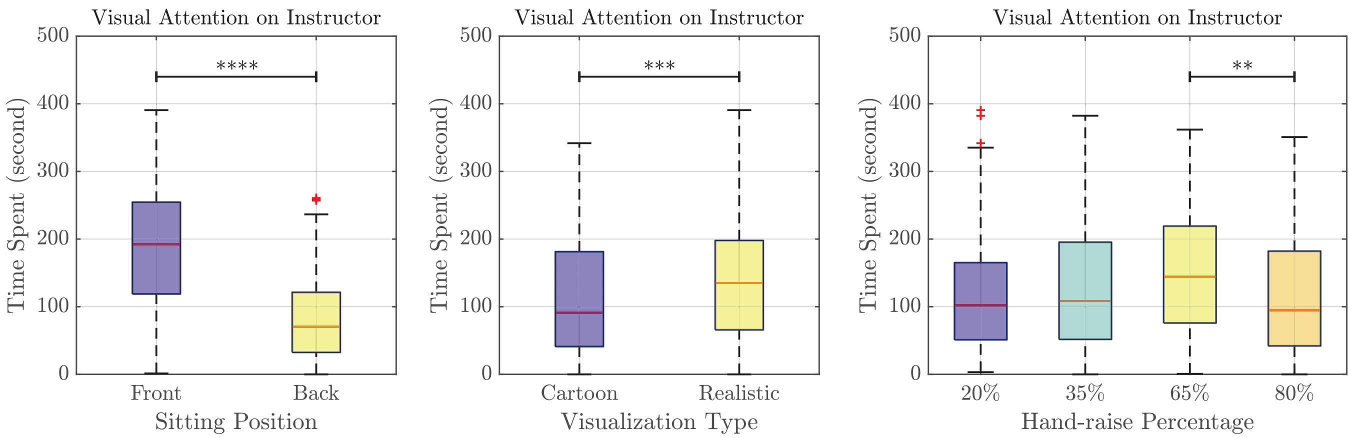
The virtual instructor drew significantly more attention in the realistic-styled avatar condition ($M = 145.98$ sec, $SD = 96.63$ sec) than in the cartoon-styled avatar condition ($M = 114.82$ sec, $SD =$ 89.83 sec) with ($F(1,264) = 11.81$, $p < .001$, $\eta^2 = .04$).

Furthermore, attention time on the instructor is found to differ significantly between different hand-raising behaviors of the peer-learners with ($F(3,264) = 3.54$, $p = .015$, $\eta^2 = .04$). In particular, the total time spent on virtual instructor in the 65% condition ($M =$ 152.46 sec, $SD = 91.48$ sec) is significantly longer than the 80% condition ($M = 117.39$ sec, $SD = 91.12$ sec) with ($F(3,264) = 3.54$, $p = .009$, $\eta^2 = .04$). Overall, more attention is drawn by the virtual instructor in the intermediate levels of hand-raising than the extreme levels. There were no interaction effects found for attention time on instructor.

4.3 Visual Attention on Screen

Total time spent on the screen, where the lecture content visualized for different sitting positions, avatar visualization styles, and various hand-raising behaviors are depicted in Fig. 5 (a), (b), and (c), respectively. The participants that are seated in the front ($M = 218.65$ sec, $SD = 78.70$ sec) attended to the lecture screen for a significantly longer period of time than the back seated participants ($M = 154.21$ sec, $SD = 96.88$ sec) with ($F(1,264) = 42.5$, $p < .0001$, $\eta^2 = .14$).

We did not find significant effects on screen attention between cartoon- and realistic-styled avatar conditions ($F(1,264) = 1.9$, $p =$ $.17$, $\eta^2 < .01$); however, attention time in realistic style ($M = 193.35$ sec, $SD = 92.30$ sec) was slightly longer than cartoon style ($M =$ 173.95 sec, $SD = 96.11$ sec).



(a) Comparison between sitting positions. (b) Comparison between visualization types. (c) Comparison between hand-raising behaviors.

Figure 4: Attention towards virtual instructor for different classroom manipulation configurations. $**$, $***$, and $****$ correspond to the significance levels of $p < .01$, $p < .001$, and $p < .0001$, respectively.

In addition, the total attention time on the screen is found to differ significantly between different hand-raising conditions with ($F(3,264) = 5.74$, $p < .001$, $\eta^2 = .06$). In particular, attention time on screen is longer in the 65% hand-raising condition ($M = 222.03$ sec, $SD = 94.90$ sec) than in the 80% condition ($M = 156.06$ sec, $SD = 88.25$ sec) with ($F(3,264) = 5.74$, $p < .001$, $\eta^2 = .06$). In addition, attention time in the 65% condition is also significantly longer than in the 35% hand-raising condition ($M = 174.87$ sec, $SD = 81.28$ sec) with ($F(3,264) = 5.74$, $p = .025$). The overall trend of attention on the lecture screen is similar to virtual instructor with the intermediate conditions being higher than the extreme conditions. There were no interaction effects found for attention time on screen.

5 Discussion

We discuss experimental results particularly for social interaction and dynamics in VR classrooms, usability of eye tracking data, and the advantages of such classrooms along with their limitations.

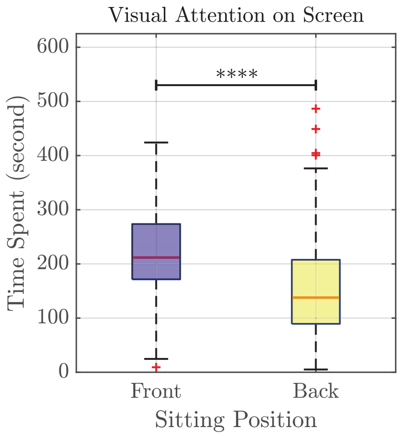
5.1 Social Dynamics in VR Classroom

We discuss our findings about social dynamics in the VR classroom in three parts, particularly based on **H1**, **H2**, and **H3** which are related to different sitting positions, different avatar visualization styles, and different hand-raise behaviors of peer-learners, respectively.

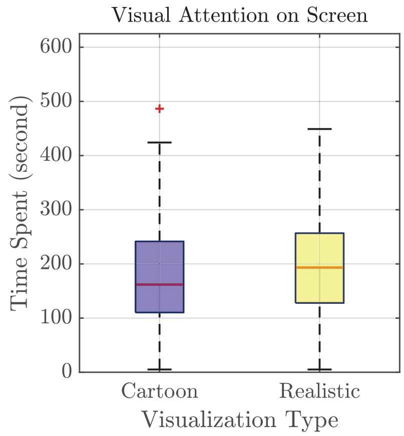
In our analyses, we found that the participants seated in the front of the classroom attended less on the peer-learners than the participants in the back, which was expected because they had fewer peers in their FOV, unless they turn back of the classroom. Assuming that during the course of the lecture, participants are supposed to listen and pay attention to the topics told by the instructor, the visual attention we observed is normal. Briefly, this is an indication that participants focus on the lecture content or instructor instead of visually interacting with their peers when seated in the front. Further, as a supporting evidence to aforementioned result, front seated participants had spent significantly more time visually attending the instructor and the screen than the participants seated in the back. We assume that these results are due to being closer to them and having fewer occluding objects in the frontal participants' FOV. These findings confirm our **H1**. Additionally, the results from the interaction effects support this hypothesis. The differences in visual attention on their virtual peer-learners for the avatar visualization style and hand-raising depend on the sitting position. Participants located in the back of the classroom have more peer-learners in their line of sight and therefore recognize the behaviour of the virtual peer-learners more, than participants seated in the front.

Our results indicate that students visually attended for longer on the peer-learners when avatars in the classroom were presented in cartoon styles. Considering the number of peer-learners in the environment and the ages of our participants being between 10-13, we argue that participants may have felt like engaging more with their peer-learners due to the emotional reasons as cartoon-styled peers are more appropriate to their ages. Realistic-styled peer-learners may be too ordinary for student engagement with peers in our setup, which led to less amount of attention. On the contrary, participants visually spent more time on the instructor when realistic-styled avatars were used. We conceive that if the avatar styles are ordinary, then the visual attention shifts to the instructor instead of interacting with the peer-learners. Lastly, as we did not find any statistical difference in attention time on the screen between different avatar visualization styles, we conclude that visual attention on the screen is not affected by such avatar visualization styles. Realism that is provided by the avatar styles may introduce additional computational complexity as such visualizations can be computationally expensive or can require additional effort to implement in advance. If the interaction with peer-learners is the main focus of the lecture, then practitioners can opt for cartoon-styled avatars. This also decreases the effort of generating the avatars. Overall, these findings confirm our **H2**.

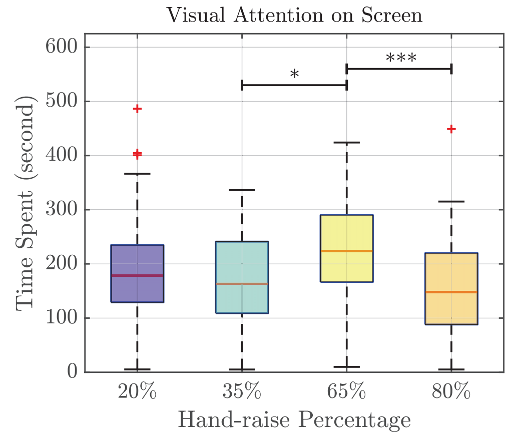
In the analysis on different hand-raising behaviors of the peer-learners, we found mixed effects. In the attention time towards peer-learners, we found a clear evidence that attention time in the extreme hand-raising conditions, namely when 80% or 20% of the virtual peer-learners raise their hands after the questions were asked by the virtual instructor is longer than in the intermediate conditions (35% and 65%). The extreme conditions may represent either more or less capable groups of peer-learners in the learning environment and participants may have a higher self-concept when surrounded by a less capable group and the other way around, which is related to the Big-fish-little-pond effect [28]. Having reasonably higher attention on peer-learners on these conditions also indicates that VR can present an opportunity to create digital environments to further study students' self-concept. On the other hand, intermediate hand-raising conditions may help students to focus more on learning related objects in the classroom instead of peer-learners such as lecture content or instructor as experimentally indicated. However,



(a) Comparison between sitting positions.



(b) Comparison between visualization types.



(c) Comparison between hand-raising behaviors.

Figure 5: Attention towards screen for different classroom manipulation configurations. $*$, $***$, and $****$ correspond to the significance levels of $p < .05$, $p < .001$, and $p < .0001$, respectively.

we expected an approximately linear increase in terms of attention time towards higher hand-raising conditions in the attention time on peer-learners. While we obtained an expected result between the 65% and 80% hand-raising conditions, the results regarding the 20% hand-raising condition do not support our hypothesis **H3**. This might be due to a moment of surprise when only a handful of peer-learners raises their hands indicating that few number of peer-learners know the answers of the questions. Furthermore, we found that attention time on the instructor tended to be longer in the intermediate levels of hand-raising than in the extreme conditions. Statistically significant results are only found for the difference between the 65% and 80% condition. While a decreasing linear trend towards the higher hand-raising percentages exists between the 65% and 80% for attention on the instructor, the overall trend is against our hypothesis, even though they are aligned with the attention time on peer-learners. Lastly, the experimental results on attention time on the screen is similar as compared to the attention time on the instructor. However, the 35% hand-raising condition drew significantly less attention than the 65% condition, which does not support our hypothesis. Overall, while some of our expectations are verified, **H3** is not confirmed. Still, the resulting behaviors should be further investigated with regard to effects on students' self-concepts during VR learning and considered when creating a classroom students are habituated to.

In summary, the three different manipulations that we studied have important effects on students' visual behavior in immersive VR classrooms in terms of social dynamics. For instance, in practice, students' self-concept can be affected by consistent hand-raising behaviors of virtual avatars over the time. While this may be less problematic in real classrooms as peer students may have different capabilities in different themes, it should carefully considered in the virtual setting, because we could present always the same behavior of the peer-learners. An adaptive strategy for hand-raising behaviors of the virtual peer-learners may be considered in practice. In addition, seating the students in the front along with realistic-styled avatars may help to increase visual attention on the lecture content. However, if a more interactive classroom environment is focused on visual interaction, practitioners can either seat students in locations where they can see their peer-learners clearly or design VR classrooms differently in terms of seating plans.

5.2 Usability of Eye Tracking Data

As eye tracking data is considered a noisy data source, we discuss our insights into the usability of this data, for particularly the immersive VR classroom setups. As aforementioned, we defined the visual attention on the different objects by using an attention threshold, which was 200 ms. In the end, in almost all conditions, the total amount of time that was spent on only the three types of objects was in the vicinity of half of the complete experiment duration despite having a relatively higher attention threshold value compared to fixation detection algorithms in the eye tracking literature. Such amount of total attention time on these three objects empirically validates our assumption of independence between them as well. We removed a significant number of samples from eye movement data due to sensory issues (e.g., lower eye tracking ratio) in order to obtain high-quality data and accurate attention mapping on the objects in the virtual classroom. While this may not be necessary for larger objects such as virtual screen in the classroom, it might cause mapping the attention wrongly for the smaller objects such as virtual avatars if the data quality is low. Considering that the participants were children in our experiments and they did not have experience with virtual reality and eye tracking, number of data removals due to such issues would be more than the experiments that are carried out with adults. In addition, unlike pre- or post-tests, eye tracking allows researchers to analyze time-dependent and temporal visual behavior changes, which can help assess students' states during virtual lectures and adapt to the environment accordingly. Therefore, despite the drawbacks, we suggest using eye movement data in such classrooms as long as an accurate calibration is applied in advance. A further iteration could take relationship of eye movement-based visual attention into consideration or analyze perceived relevance of lecture content along with eye-gaze behaviors such as in [12] and [45], respectively.

5.3 Advantages and Limitations

One of the advantages of immersive VR classroom setups is the opportunity of simulating different classroom manipulations in remote settings, which are difficult to do in real world, and evaluate students' behaviors and learning under such manipulations. Another advantage of such setups is the possibility of preserving the privacy of students since the videos that include faces are not recorded in such settings. In real world classrooms, it is troublesome to record and store videos of the class while lecturing, even though there are some efforts supporting the automated anonymization [43] of such data. In contrast, data collected from virtual classrooms can be pseudo-anonymized. However, one should be aware of the amount of personal information that can be extracted from eye movement data and how to manipulate it [7, 8, 15]. Furthermore, one should take the relationship between iris texture and biometrics into account and how to preserve privacy in case eye videos are recorded and stored [10]. In addition, we observed during experiments that some of the students intended to raise their hands when seeing the hand-raising behaviors of the virtual peer-learners. While we did not record hand tracking data in our study, it is possible to accurately assess the intentions of students towards questions asked by the virtual instructor by using a hand tracker device on the HMD, which is another advantage of VR setups compared to real classrooms. Although, hand-raising is a good indicator of children's participation during a lecture, we do not know if students interpret this behaviour of their virtual peers as a sign of competence, engagement, or motivation.

Despite the advantages, there are other technical limitations regarding the use of VR classrooms. Long periods of exposure to VR lectures can lead to immense levels of cybersickness. In addition, a vast amount of HMD movement on the head may cause a drift in eye tracker calibration, leading to incorrect sensor readings. This can affect interaction experience if gaze-aware features are included in virtual environments. These should be taken into consideration when designing a virtual classroom and lecture. Particularly, the duration of the lecture should be chosen carefully to minimize these effects.

6 Conclusion

To understand the visual attention in VR classrooms in different manipulations, we analyzed object-of-interest information based on eye-gaze. We found that participants seated in the front attended more time to the virtual instructor and the screen displaying lecture content. In addition, participants focused on the cartoon-styled peer-learners more than realistic-styled ones, whereas in the realistic-styled avatar manipulation the virtual instructor drew more visual attention. The extreme conditions of hand-raising behaviors drew more attention towards virtual peer-learners, whereas in the intermediate conditions visual attention was focused more on the instructor and screen. These findings are based on the eye movements of the participants and correspond to the social dynamics of VR classrooms such as students' self-concept or peer-learner interaction; however, such manipulations may also affect learning outcomes. While our results provide primitive but fundamental cues about how to design immersive VR classrooms by taking students' visual behaviors into account for different goals in digital teaching, effects of such manipulations on the learning outcome should be further investigated.

As future work, we plan to specifically investigate the relationship between different manipulations with temporal gaze dynamics as an immediate response to asked questions and related students' performances.

Acknowledgments

This research was partly supported by a grant to Richard Göllner funded by the Ministry of Science, Research and the Arts of the state of Baden-Württemberg and the University of Tübingen as part of the Promotion Program of Junior Researchers. Lisa Hasenbein and Philipp Stark are doctoral candidates and supported by the LEAD Graduate School & Research Network, which is funded by the Ministry of Science, Research and the Arts of the state of Baden-Württemberg within the framework of the sustainability funding for the projects of the Excellence Initiative II. Authors thank Stephan Soller, Sandra Hahn, and Sophie Fink from the Hochschule der Medien Stuttgart for their work and support related to the immersive virtual reality classroom used in this study.

References

[1] I. Agtzidis, M. Startsev, and M. Dorr. 360-degree video gaze behaviour: A ground-truth data set and a classification algorithm for eye movements. In *Proceedings of the 27th ACM International Conference on Multimedia*, p. 1007–1015. ACM, New York, NY, USA, 2019. doi: 10.1145/3343031.3350947

[2] W. Alhalabi. Virtual reality systems enhance students' achievements in engineering education. *Behaviour & Information Technology*, 35(11):1–7, 2016. doi: 10.1080/0144929X.2016.1212931

[3] E. Arabadzhiyska, O. T. Tursun, K. Myszkowski, H.-P. Seidel, and P. Didyk. Saccade landing position prediction for gaze-contingent rendering. *ACM Trans. Graph.*, 36(4), 2017. doi: 10.1145/3072959.3073642

[4] J. N. Bailenson, N. Yee, J. Blascovich, A. C. Beall, N. Lundblad, and M. Jin. The use of immersive virtual reality in the learning sciences: Digital transformations of teachers, students, and social context. *Journal of the Learning Sciences*, 17(1):102–141, 2008. doi: 10.1080/10508400701793141

[5] F. Blume, R. Göllner, K. Moeller, T. Dresler, A.-C. Ehlis, and C. Gawrilow. Do students learn better when seated close to the teacher? a virtual classroom study considering individual levels of inattention and hyperactivity-impulsivity. *Learning and Instruction*, 61:138–147, 2019. doi: 10.1016/j.learninstruc.2018.10.004

[6] E. Bozkir, D. Geisler, and E. Kasneci. Assessment of driver attention during a safety critical situation in VR to generate VR-based training. In *ACM Symposium on Applied Perception 2019*. ACM, New York, NY, USA, 2019. doi: 10.1145/3343036.3343138

[7] E. Bozkir, O. Günlü, W. Fuhl, R. F. Schaefer, and E. Kasneci. Differential privacy for eye tracking with temporal correlations. https://arxiv.org/abs/2002.08972, 2020.

[8] E. Bozkir, A. B. Ünal, M. Akgün, E. Kasneci, and N. Pfeifer. Privacy preserving gaze estimation using synthetic images via a randomized encoding based framework. In *ACM Symposium on Eye Tracking Research and Applications*. ACM, New York, NY, USA, 2020. doi: 10.1145/3379156.3391364

[9] A. Casu, L. D. Spano, F. Sorrentino, and R. Scateni. Riftart: Bringing masterpieces in the classroom through immersive virtual reality. In *Smart Tools and Apps for Graphics - Eurographics Italian Chapter Conference*, pp. 77–84. The Eurographics Association, Geneva, Switzerland, 2015. doi: 10.2312/stag.20151294

[10] A. K. Chaudhary and J. B. Pelz. Privacy-preserving eye videos using rubber sheet model. In *ACM Symposium on Eye Tracking Research and Applications*. ACM, New York, NY, USA, 2020. doi: 10.1145/3379156.3391375

[11] K. Cheng and C. Tsai. A case study of immersive virtual field trips in an elementary classroom: Students' learning experience and teacher-student interaction behaviors. *Computers & Education*, 140:103600, 2019. doi: 10.1016/j.compedu.2019.103600

[12] E. B. Cloude, D. A. Dever, M. D. Wiedbusch, and R. Azevedo. Quantifying scientific thinking using multichannel data with crystal island: Implications for individualized game-learning analytics. *Frontiers in Education*, 5:217, 2020. doi: 10.3389/feduc.2020.572546

[13] U. Díaz-Orueta, C. García-López, N. Crespo-Eguílaz, R. Sánchez-Carpintero, G. Climent, and J. Narbona. AULA virtual reality test as an attention measure: Convergent validity with conners' continuous performance test. *Child Neuropsychology*, 20(3):328–342, 2014. doi: 10.1080/09297049.2013.792332

[14] L. Freina and M. Ott. A literature review on immersive virtual reality in education: State of the art and perspectives. In *Proceedings of the 11th International Scientific Conference eLearning and Software for Education*, pp. 133–141. Carol I NDU Publishing House, Bucharest, Romania, 2015. doi: 10.12753/2066-026X-15-020

[15] W. Fuhl, E. Bozkir, and E. Kasneci. Reinforcement learning for the privacy preservation and manipulation of eye tracking data. https://arxiv.org/abs/2002.06806, 2020.

[16] W. Fuhl, M. Tonsen, A. Bulling, and E. Kasneci. Pupil detection for head-mounted eye tracking in the wild: an evaluation of the state of the art. *Machine Vision and Applications*, 27(8), 2016. doi: 10.1007/s00138-016-0776-4

[17] J. Hadnett-Hunter, G. Nicolaou, E. O'Neill, and M. Proulx. The effect of task on visual attention in interactive virtual environments. *ACM Trans. Appl. Percept.*, 16(3), 2019. doi: 10.1145/3352763

[18] E. Johnston, G. Olivas, P. Steele, C. Smith, and L. Bailey. Exploring

pedagogical foundations of existing virtual reality educational applications: A content analysis study. *Journal of Educational Technology Systems*, 46(4):414–439, 2018. doi: 10.1177/0047239517745560

[19] S. Kavanagh, A. Luxton-Reilly, B. Wuensche, and B. Plimmer. A systematic review of virtual reality in education. *Themes in Science and Technology Education*, 10(2):85–119, 2017.

[20] M. Khamis, C. Oechsner, F. Alt, and A. Bulling. VRpursuits: Interaction in virtual reality using smooth pursuit eye movements. In *Proceedings of the 2018 International Conference on Advanced Visual Interfaces*. ACM, New York, NY, USA, 2018. doi: 10.1145/3206505.3206522

[21] R. Lamb and E. A. Etopio. Virtual Reality: A Tool for Preservice Science Teachers to Put Theory into Practice. *Journal of Science Education and Technology*, 29(4):573–585, 2020. doi: 10.1007/s10956-020-09837-5

[22] Y. Lang, L. Wei, F. Xu, Y. Zhao, and L.-F. Yu. Synthesizing personalized training programs for improving driving habits via virtual reality. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 297–304. IEEE, New York, NY, USA, 2018. doi: 10.1109/VR.2018.8448290

[23] E. Langbehn, F. Steinicke, M. Lappe, G. F. Welch, and G. Bruder. In the blink of an eye: Leveraging blink-induced suppression for imperceptible position and orientation redirection in virtual reality. *ACM Trans. Graph.*, 37(4), 2018. doi: 10.1145/3197517.3201335

[24] T. M. Lee, J.-C. Yoon, and I.-K. Lee. Motion sickness prediction in stereoscopic videos using 3D convolutional neural networks. *IEEE Transactions on Visualization and Computer Graphics*, 25(5):1919–1927, 2019. doi: 10.1109/TVCG.2019.2899186

[25] M.-Y. Liao, C.-Y. Sung, H.-C. Wang, and W.-C. Lin. Virtual classmates: Embodying historical learners' messages as learning companions in a VR classroom through comment mapping. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 163–171. IEEE, New York, NY, USA, 2019. doi: 10.1109/VR.2019.8797708

[26] M. Lombard, T. Bolmarcich, and L. Weinstein. Measuring presence: The temple presence inventory. In *Proceedings of the 12th Annual International Workshop on Presence*, pp. 1–15. The International Society for Presence Research, Los Angeles, CA, USA, 2009.

[27] A. Mangalmurti, W. Kistler, B. Quarrie, W. Sharp, S. Persky, and P. Shaw. Using virtual reality to define the mechanisms linking symptoms with cognitive deficits in attention deficit hyperactivity disorder. *Scientific Reports*, 10(1), 2020. doi: 10.1038/s41598-019-56936-4

[28] H. W. Marsh and J. W. Parker. Determinants of student self-concept: Is it better to be a relatively large fish in a small pond even if you don't learn to swim as well? *Journal of Personality and Social Psychology*, 47(1):213–231, 1984. doi: 10.1037/0022-3514.47.1.213

[29] X. Meng, R. Du, and A. Varshney. Eye-dominance-guided foveated rendering. *IEEE Transactions on Visualization and Computer Graphics*, 26(5):1972–1980, 2020. doi: 10.1109/TVCG.2020.2973442

[30] C. Moro, Z. Štromberga, A. Raikos, and A. Stirling. The effectiveness of virtual and augmented reality in health sciences and medical anatomy. *Anatomical Sciences Education*, 10(6):549–559, 2017. doi: 10.1002/ase.1696

[31] P. Nolin, A. Stipanicic, M. Henry, Y. Lachapelle, D. Lussier-Desrochers, A. S. Rizzo, and P. Allain. ClinicaVR: Classroom-CPT: A virtual reality tool for assessing attention and inhibition in children and adolescents. *Computers in Human Behavior*, 59:327–333, 2016. doi: 10.1016/j.chb.2016.02.023

[32] J. Orlosky, Y. Itoh, M. Ranchet, K. Kiyokawa, J. Morgan, and H. Devos. Emulation of physician tasks in eye-tracked virtual reality for remote diagnosis of neurodegenerative disease. *IEEE Transactions on Visualization and Computer Graphics*, 23(4):1302–1311, 2017. doi: 10.1109/TVCG.2017.2657018

[33] A. A. Rizzo, T. Bowerly, J. G. Buckwalter, D. Klimchuk, R. Mitura, and T. D. Parsons. A virtual reality scenario for all seasons: The virtual classroom. *CNS Spectrums*, 11(1):35–44, 2006. doi: 10.1017/S1092852900024196

[34] A. A. Rizzo, J. G. Buckwalter, T. Bowerly, C. Van Der Zaag, L. Humphrey, U. Neumann, C. Chua, C. Kyriakakis, A. Van Rooyen, and D. Sisemore. The virtual classroom: A virtual reality environment for the assessment and rehabilitation of attention deficits. *CyberPsychology & Behavior*, 3(3):483–499, 2000. doi: 10.1089/10949310050078940

[35] S. D. Roth. Ray casting for modeling solids. *Computer Graphics and Image Processing*, 18(2):109–144, 1982. doi: 10.1016/0146-664X(82)90169-1

[36] D. D. Salvucci and J. H. Goldberg. Identifying fixations and saccades in eye-tracking protocols. In *Proceedings of the 2000 Symposium on Eye Tracking Research & Applications*, p. 71–78. ACM, New York, NY, USA, 2000. doi: 10.1145/355017.355028

[37] A. Schmitz, A. MacQuarrie, S. Julier, N. Binetti, and A. Steed. Directing versus attracting attention: Exploring the effectiveness of central and peripheral cues in panoramic videos. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 63–72. IEEE, New York, NY, USA, 2020. doi: 10.1109/VR46266.2020.00024

[38] T. Schubert, F. Friedmann, and H. Regenbrecht. The experience of presence: Factor analytic insights. *Presence*, 10(3):266–281, 2001. doi: 10.1162/105474601300343603

[39] S.-h. Seo, E. Kim, P. Mundy, J. Heo, and K. Kim. Joint attention virtual classroom: A preliminary study. *Psychiatry Investigation*, 16(4):292–299, 2019. doi: 10.30773/pi.2019.02.08

[40] S. Sharma, R. Agada, and J. Ruffin. Virtual reality classroom as an constructivist approach. In *2013 Proceedings of IEEE Southeastcon*, pp. 1–5. IEEE, New York, NY, USA, 2013. doi: 10.1109/SECON.2013.6567441

[41] L. Sidenmark and A. Lundström. Gaze behaviour on interacted objects during hand interaction in virtual reality for eye tracking calibration. In *Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications*. ACM, New York, NY, USA, 2019. doi: 10.1145/3314111.3319815

[42] A. L. Simeone, M. Speicher, A. Molnar, A. Wilde, and F. Daiber. Live: The human role in learning in immersive virtual environments. In *Symposium on Spatial User Interaction*. ACM, New York, NY, USA, 2019. doi: 10.1145/3357251.3357590

[43] Ö. Sümer, P. Gerjets, U. Trautwein, and E. Kasneci. Automated anonymisation of visual and audio data in classroom studies. In *The Workshops of the Thirty-Forth AAAI Conference on Artificial Intelligence*. AAAI Press, Palo Alto, CA, USA, 2020.

[44] D. Weintrop, E. Beheshti, M. Horn, O. Kai, K. Jona, L. Trouille, and U. Wilensky. Defining computational thinking for mathematics and science classrooms. *Journal of Science Education and Technology*, 25(1):127–147, 2016. doi: 10.1007/s10956-015-9581-5

[45] M. D. Wiedbusch and R. Azevedo. Modeling metacomprehension monitoring accuracy with eye gaze on informational content in a multimedia learning environment. In *ACM Symposium on Eye Tracking Research and Applications*. ACM, New York, NY, USA, 2020. doi: 10.1145/3379155.3391329

[46] J. O. Wobbrock, L. Findlater, D. Gergle, and J. J. Higgins. The aligned rank transform for nonparametric factorial analyses using only anova procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, p. 143–146. ACM, New York, NY, USA, 2011. doi: 10.1145/1978942.1978963

[47] X. Zhang, S. Park, T. Beeler, D. Bradley, S. Tang, and O. Hilliges. ETH-XGaze: A large scale dataset for gaze estimation under extreme head pose and gaze variation. In A. Vedaldi, H. Bischof, T. Brox, and J.-M. Frahm, eds., *Computer Vision – ECCV 2020*, pp. 365–381. Springer International Publishing, Cham, Switzerland, 2020. doi: 10.1007/978-3-030-58558-7_22

[48] Y. Zhang, W. Hu, W. Xu, C. T. Chou, and J. Hu. Continuous authentication using eye movement response of implicit visual stimuli. *ACM Interact. Mob. Wearable Ubiquitous Technology*, 1(4):177:1–177:22, 2018. doi: 10.1145/3161410