Evaluating the Effects of Virtual Human Animation on Students in an Immersive VR Classroom Using Eye Movements

Hong Gao*
hong.gao@informatik.uni-tuebingen.de
Human-Computer Interaction
University of Tübingen
Tübingen, Germany

Lisa Hasenbein†
Hector Research Institute of Education Sciences and Psychology
University of Tübingen
Tübingen, Germany
lisa.hasenbein@uni-tuebingen.de

Efe Bozkir
Human-Computer Interaction
University of Tübingen
Tübingen, Germany
efe.bozkir@uni-tuebingen.de

Richard Göllner
Hector Research Institute of Education Sciences and Psychology
University of Tübingen
Tübingen, Germany
richard.goellner@uni-tuebingen.de

Enkelejda Kasneci
Human-Centered Technologies for Learning
Technical University of Munich
Munich, Germany
enkelejda.kasneci@tum.de

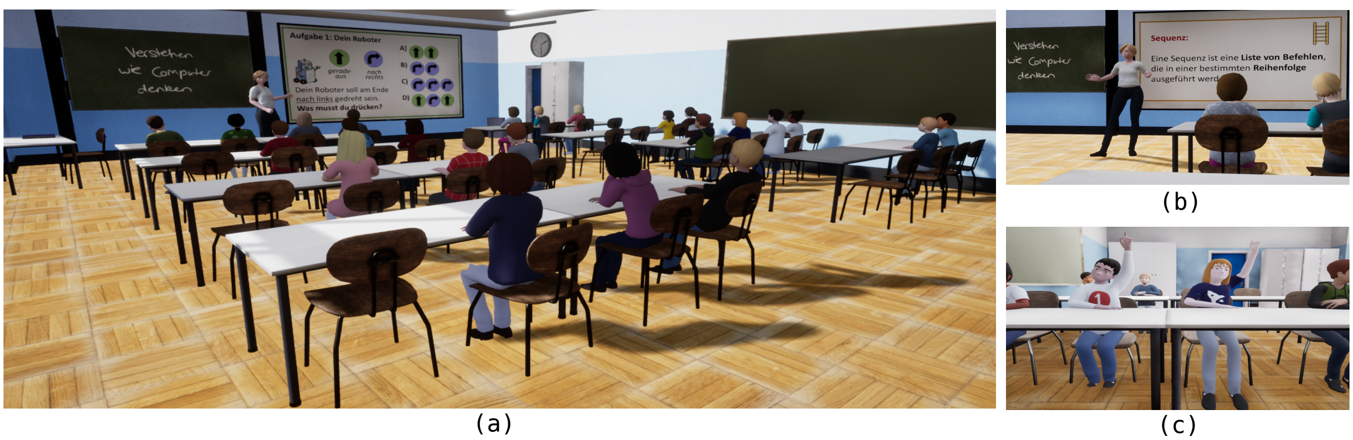


Figure 1: Immersive virtual reality classroom: (a) overall view, (b) student's view, (c) virtual peer learners with hand raising animations.

ABSTRACT

Virtual humans presented in VR learning environments have been suggested in previous research to increase immersion and further positively influence learning outcomes. However, how virtual human animations affect students' real-time behavior during VR learning has not yet been investigated. This work examines the effects of social animations (i.e., hand raising of virtual peer learners) on students' cognitive response and visual attention behavior during immersion in a VR classroom based on eye movement analysis. Our results show that animated peers that are designed to enhance immersion and provide companionship and social information elicit different responses in students (i.e., cognitive, visual attention, and visual search responses), as reflected in various eye movement metrics such as pupil diameter, fixations, saccades, and dwell times. Furthermore, our results show that the effects of animations on students differ significantly between conditions (20%, 35%, 65%, and 80% of virtual peer learners raising their hands). Our research provides a methodological foundation for investigating the effects of avatar animations on users, further suggesting that such effects should be considered by developers when implementing animated virtual humans in VR. Our findings have important implications for future works on the design of more effective, immersive, and authentic VR environments.

*Equal contributor.
†Equal contributor.

VRST '22, November 29-December 1, 2022, Tsukuba, Japan

ACM ISBN 978-1-4503-9889-3/22/11…$15.00
https://doi.org/10.1145/3562939.3565623

CCS CONCEPTS

• **Human-centered computing** → **Virtual reality**; **Empirical studies in HCI**.

KEYWORDS

immersive virtual reality, virtual human animation, eye-tracking, visual attention, education

ACM Reference Format:

Hong Gao, Lisa Hasenbein, Efe Bozkir, Richard Göllner, and Enkelejda Kasneci. 2022. Evaluating the Effects of Virtual Human Animation on Students in an Immersive VR Classroom Using Eye Movements. In *28th ACM Symposium on Virtual Reality Software and Technology (VRST '22), November 29-December 1, 2022, Tsukuba, Japan.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3562939.3565623

1 INTRODUCTION

With the increasing availability of consumer-grade head-mounted displays (HMDs), virtual reality (VR) has been successfully deployed and is gaining immense popularity in the field of education [29, 39]. Particularly, VR classrooms that mimic traditional classroom environments allow for authentic social interaction and engagement, which has been seen as crucial to increase learner motivation, persistence, and interest [32], ultimately leading to better learning outcomes [33]. However, in VR classrooms to date, virtual avatars (particularly virtual classmates) that students can interact with or that provide them with social information (i.e., hand raising) to enhance immersion have not been fully developed and envisioned yet. Instead, VR classrooms are typically designed to provide either social interaction based on real-time interaction with another real person, meaning that individual learners need to participate synchronously with others in the same class [30], or provide pre-programmed but limited interactive avatars [32, 34]. Given that synchronous interaction in a VR classroom with a comparable number of students as in a real classroom (e.g., more than twenty students) is difficult to achieve, the development of VR environments that allow offline yet authentic social interaction with (pre-programmed) avatar teachers and classmates is a promising avenue for VR-based classroom learning, as such environments offer learners an authentic degree of flexibility with regard to the social information they process during learning.

It has been recently reported in [32] that presenting virtual humans pre-programmed with interactive animations in VR classrooms improved learners' motivation and immersion. However, this poses the question of how animated avatar affect learners while they are immersed in VR. Previous work has addressed this question and found that virtual human animations elicit a wide range of affective responses in humans, such as stress [47] and self-disclosure [27], and have effects on users' emotional responses [51, 54] and visual attention [20, 53]. These works highlight the importance and necessity of exploring the effect of avatar animations on various aspects of user behavior when presented in VR.

Against the background that students perceive their virtual peer's behavior in the VR classroom similarly to in a real-world learning scenario [1, 40] , peer learners can be purposefully implemented in a pre-programmed VR classroom to promote student learning. Educational psychological research on so-called peer effects has repeatedly shown that the motivation of peer learners influence individual learning trajectories regarding motivational as well as performance outcomes [49] Hand-raising presents a highly salient behavior of students in a classroom setting, and contains important information about a student's performance and motivation that is perceived by peer learners and substantially shapes their learning experience in the classroom [7, 20]. Therefore, we consider it as one of the most important peer animations needed to be explored in VR classroom studies. In particular, studying the effect of (a crowd of) animated virtual humans, which typically convey social information (e.g., hand-raising of virtual peer learners), on students' cognitive and visual attention behaviors reflected in eye movements can provide valuable insights for the design features of VR-based learning environments (e.g., VR classroom) where social information is typically present to enhance immersion. Furthermore, a comprehensive understanding of the effects of avatar animations on learners is indispensable to creating well-functioning VR learning systems or more interactive online learning systems. Eye movements as a non-intrusive and objective measurement have been proposed in previous works as a more intuitive method to investigate subjects' conscious and unconscious temporal behavior during tasks, such as cognitive load [3, 5, 8], visual attention including classroom attention [6, 14, 18, 48], visual search behavior [14], and IQ test solving [28]. With this in mind, eye movements could provide a promising avenue to study users' cognitive and behavioral (e.g., visual attention and visual search) responses triggered by virtual human animations (e.g., social information) during VR experiences. And thanks to the state-of-the-art HMDs, eye-tracking data can be easily obtained via the built-in eye tracker, further facilitating the use of eye movements in VR studies.

In this work, we propose an authentic immersive VR classroom for offline learning and conduct an in-depth investigation of the effects of a crowd of physically animated virtual humans (particularly students' primary social counterparts: peer learners) on students' cognitive responses and visual attention behaviors while immersed in VR. Specifically, we designed an immersive VR classroom where students experience an immersive lesson that simulates a real classroom situation with a virtual teacher and a class of virtual peer learners pre-programmed with animations. The animations were created based on motion capture from real classrooms. To evaluate students' cognitive and visual attention behaviors before and after the onset of animation (i.e., hand raising of peers), we measured their eye movements and pupil measures, including pupil diameter, fixations, saccades, and dwell times. Furthermore, to examine whether the magnitude of the effects of animations was related to the amount of social information provided by the virtual peer learners, we purposefully implemented hand raising to mimic different variations of real classroom social interactions (between-subjects design: 20%, 35%, 65% or 80% of virtual peer learners are programmed with hand raising). We calculated the change scores of the eye movement variables before and after the onset of hand raising as new variables and use these new variables to represent the magnitude of the effects of animations on students' cognitive response and visual attention. A comparison was made between the four hand-raising conditions. Our study provides foundation for future work investigating the effects of a crowd of animated virtual peer learners on students' cognitive response and visual attention behavior during an immersive VR learning experience. Hence, it offers profound insights for optimizing the design of virtual learning environments from the perspective of virtual human animations. This improvement in system design will help to further improve learning, which is however beyond the scope of our current work.

2 RELATED WORK

2.1 Immersive VR Classrooms

The rapid development of affordable VR devices makes it easier to create VR learning environments for learners in various ways, such as VR labs [23, 46], VR field trips [10, 36], and VR classrooms [6, 14, 45]. Among these different VR learning contexts, VR classrooms that adopt traditional classroom settings, such as room layout and especially the presence of virtual teachers and virtual classmates exhibiting social interaction behaviors, have the advantage of providing a more authentic learning experience and higher engagement for students [30, 32, 34]. For instance, Liao et al. [32] suggested constructing virtual classmates as peer companions by synthesizing time-anchored comments from previous learners to reduce learners' sense of isolation and increase immersion and motivation during learning in the VR classroom. Learners were found to achieve better learning outcomes with the accompaniment of a few virtual classmates than without virtual classmates. Similarly, Liu et al. [34] developed a VR classroom with a virtual teacher and virtual classmates as an experimental platform to investigate the redundancy effect in learning. Interactive behaviors and animations of the virtual classmates, e.g., head movements (i.e., raising, lowering, and shaking the head), turning around, sneezing, and taking notes, were generated and found to have effects on learners and to be a decisive factor in the reverse redundancy effect. In addition to VR classrooms designed for learning, Ke et al. [30] developed a VR-based, Kinect-integrated learning environment for teaching training. Participants were teaching assistants at a university and were asked to play the role of a teacher and a student, respectively, in two sessions with corresponding avatars. With other avatars in the classroom played by peer trainees (whose body movements were projected onto their avatars in real time) and controlled by computers (with pre-programmed animations), participants maintained higher level of presence during immersion in VR-based training environments, regardless of whether they were acting as teachers or students.

The aforementioned works demonstrate the importance of the presence of animated virtual humans in VR environments, specifically classrooms, for improving learner engagement and immersion, as assessed by their subjective perception of the virtual human animations and learning outcomes. This inspired our study to develop an immersive VR classroom that mimics a real classroom scenario for offline learning, where the virtual teacher and virtual peer learners were presented with pre-programmed interactive animations to provide students with social information and companionship.

2.2 Virtual Human Animations

Researchers have studied various effects that virtual humans have on users during VR experiences. Robb et al. [41] examined the effects of the presence of a virtual human animated with verbal responses on students taking a prostate exam in a prostate exam simulation. They found that the virtual human elicited stress in students and led to increased engagement. In another study, Volonte et al. [53] examined how animation fidelity of the virtual human affected users' gaze behavior in a medical VR training simulation similar to [41]. It was found that the conversational and passive animations of the virtual human elicited visual attention responses from users, and users' visual attention shifted between the virtual human and goal directed activities. However, in these studies, users took part in VR simulations presented on a screen rather than in an immersive HMD-based VR environment, performing simple tasks and interacting with a limited number of (individual) virtual humans with limited animations. Closer to our work in terms of the number of animated virtual humans and VR environment settings, Volonte et al. [52] further investigated the effects of virtual human animations on users' affective and non-verbal behaviors while interacting with a crowd of virtual humans with animated emotions in an immersive HMD-based VR market scenario. It was found that virtual crowds with positive emotions elicited the highest scores on metrics related to interaction with the virtual agents.

However, it has not yet been investigated how animated virtual humans, which are intended to enhance immersion and serve as companions, in an immersive VR classroom affect learners' real-time and instantaneous and cognitive responses and visual attention behaviors reflected in eye movements during learning. Instead, such animated virtual humans in VR classrooms were found to influence learners in their overall experience but were not assessed by eye movements. [30, 32, 34]. A few previous works have investigated learners' eye movements during the VR experience. For instance, Gao et al. [14] investigated how different configurations of VR classrooms affect learners' gaze behavior while learning in VR. A range of eye movement features, including fixations and saccades, were extracted and analyzed with average measures. Their results showed that learners' gaze behavior differed significantly between different configurations of VR classrooms, suggesting that eye-tracking is indicative of learners' response to changes in classroom configurations. Following, Bozkir et al. [6] found that learners' visual attention switched between different virtual objects of interest (OOIs) while participating in a virtual lesson in the VR classroom. Different from the work of Gao et al. [14], which examined different classroom configurations, the study looked in more detail at learner attention to specific OOIs, such as instructional content and social information from animated peer learners, which have more to do with student learning behaviors. Although these studies examined an averaged gaze behavior across the entire duration of the VR experience, they provide evidence for our study to use eye movements as a metric to explore learners' real-time and instantaneous cognitive and visual attention responses to virtual human animations that last for a short period of time in VR classroom.

3 METHODS

3.1 Participants and Apparatus

In our study, 381 sixth-grade volunteer students (179 female, 202 male) with an average age of 11.5 years (10 to 13, $SD = 0.56$) were recruited from local schools to participate in our experiment.

All volunteer participants and their guardians provided informed consent before the experiment. Our study was IRB-approved.

The VR classroom environment was rendered using the Unreal Game Engine [1] v4.23.1. The HTC Vive Pro Eye with a refresh rate of 90 Hz and a field of view of 110° was used. The integrated Tobii eye tracker with a sampling frequency of 120 Hz and a standard calibration accuracy of $0.5° - 1.1°$ was used to record the eye tracking data.

3.2 Experimental Design

To mimic a real classroom environment for sixth graders, we used the same configuration in the virtual classroom as in real classrooms (e.g., the classroom layout). In addition, a virtual teacher and twenty-four virtual peer learners were rendered with pre-programmed animations to enhance realism and authenticity. The virtual human animations were created based on motion capture from real classrooms and therefore authentically mimic physical movements similar to those of real people in real classrooms. We used recordings and motion captures from a sixth-grade classroom to ensure that the virtual peers' behaviors correspond to their virtual representation as well as the study participants' age. We used Xsens Motion Capture suits to record the authentic movements of the teacher and of six different students during a 15-min immersive VR lesson in a regular school setting. The students were asked to behave like they usually would in their classroom. The hand-raising was embedded in natural movements that showed students' motivation to be called on (e.g., leaning forward); overall, however, to avoid any confounding effects, any other distracting behaviors (like students moving around without reason and showing off-task behavior) were not included in the animations. Participants sat in such a VR classroom and listened to a virtual lesson ($\approx$ 15 minutes) on computational thinking (including basic concepts such as sequences and loops, practical exercises applied to them) delivered by the virtual teacher. During the lesson, the virtual teacher walks around the podium, asks simple questions (twenty-one in total), and calls on virtual peer learners with body gestures (e.g., hand gestures), while the virtual peer learners turn around, think, and interact with the virtual teacher by raising their hands to answer questions. To ensure controllability, all animations were pre-programmed. To mimic different variations of real-world social interactions in the classroom, a fixed group (between-subjects design: 20%, 35%, 65% or 80%) of virtual peer learners was programmed with hand raising behavior. We chose these percentages of hand-raising students (a) for pragmatic reasons (i.e., limiting the number of conditions to four so we could reach sufficient statistical power with the target sample size) and (b) for conceptual reasons (i.e., to examine differentiated effects based on an unambiguous picture of either a minority or majority of peers raising their hands; hence, a minority of 20% or 35% or a majority of 65% and 80%).

Note that the VR classroom had four rows of tables with virtual peer learners that showed the pre-programmed behavior without any randomizations. Participants were randomly either placed in the second (front position) or fourth row (back position). To ensure comparability across conditions, the peer learners in rows 1-2 and rows 3-4 were programmed in a similar manner to ensure that the participating students had similar experiences with regard to the position and proximity of the (hand-raising) peers. The hand-raising animations were identical across the conditions (with natural variations in authentic movements). Similarly, the audio file used was the same across conditions; one of the hand-raising peers was each called on and answered the teacher's question (based on audio recorded from the same sixth graders as the motion captures. The immersive VR classroom is shown in Figure 1.

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

3.3 Procedure

After signing the informed consent form, participants were randomly assigned to one of the experimental conditions. Each experimental session lasted approximately 45 minutes and included a paper-based pre-test for demographic and learning background information, participation in the virtual lesson, and a paper-based post-test reporting on the VR experience. The post-test used a 4-point rating scale to assess participants' experienced level of spatial and social presence in the VR classroom (9 items based on REFS [35, 44]; e.g., "I felt like I was sitting in the virtual classroom" or "I felt like the teacher in the virtual classroom really addressed me") and their perceived realism of the VR classroom (6 items; e.g., "What I experienced in the virtual classroom could also happen in a real classroom" or "The students in the virtual classroom behaved similarly to real classmates"). The results indicated an overall authentic experience with mean levels of experienced presence and perceived realism ranging from 2.82 to 2.97 ($0.52 < SD < 0.62$) in all configuration conditions, with no significant differences found between conditions. Prior to the virtual lesson, a standard 5-point eye-tracking calibration routine was performed. All participants were informed before the experiment that they could drop out the experiment at any time without consequences.

3.4 Data Preprocessing and Measures

Data including eye movements and head poses were collected from 381 participants. 40 participants who experienced hardware problems, invalid calibration, or synchronization issues in the VR environment were excluded. In addition, 61 participants with an eye-tracking ratio of lower than 90% (less than 90% of signal was recorded) were excluded. Therefore, 15 minutes of behavioral data from each of the 280 participants (140 female, 140 male) were used in our study. Since only raw tracking data, i.e., pupil size, gaze vectors, and head vectors, were collected, we performed data preprocessing that included normalization of pupil diameter and detection of eye movements for further analysis.

Pupil diameter has been suggested in previous studies as an indicator of cognitive load in human cognitive processes [3–5, 8], so we used it as a measure in our study. Since pupillometry signals were affected by noisy sensor readings and blinks, we smoothed and normalized pupil diameters using the Savitzky-Golay filter [43] and divisive baseline correction with a baseline duration of $\approx$ 1 second [37].

Fixations, i.e., the periods of time during which the eyes are stationary in the head, are generally considered an indicator of visual attention behavior [4, 11, 19, 38]. Saccades, i.e., rapid eye movement shifts between fixations, have been found to correlate strongly with visual search behavior [15, 21]. Before detecting such eye movements from the raw eye-tracking data, linear interpolation was performed for the missing gaze vectors. A modified algorithm based on the well-established and widely used velocity-threshold identification (I-VT) [22] was used for fixation detection [2, 14]. Specifically, fixations were detected under the stationary head condition (head velocity $< 7^\circ/s$) [2] with a maximum gaze velocity threshold of $30^\circ/s$. As saccade detection was not constrained by head movement, saccades were detected using the normal I-VT method, with a minimum gaze velocity threshold of $60^\circ/s$. In addition, duration thresholds were applied, with a minimum duration of 100*ms* and a maximum duration of 500*ms* for fixation detection and a minimum duration of 30*ms* and a maximum duration of 80*ms* for saccade detection.

Dwell time, i.e., the total time spent in an area-of-interest (AOI), including all fixations and saccades as well as revisits, is a metric that conveys the level of interest and attentive behavior within a certain AOI of the stimuli [21]. In addition, the time to the first fixation (TTFF) and the first fixation duration (FFD) in an OOI are interpreted as visual scene priority [21]. Therefore, we computed these measures for the main objects of interest (OOIs) in the VR classroom, namely the virtual teacher, virtual peer learners, and the screen displaying the instructional content. The number of virtual peers that participants fixated on was also used as a measure.

3.5 Hypotheses

Our aim is to investigate the effects of social interaction animations between the virtual teacher and peer learners (i.e., hand raising of peers) on participants' cognitive responses and visual attention behaviors by analyzing various eye movements and pupil measures. We tested the following hypotheses:

- **H1** We hypothesized that the animated virtual peers will influence participants' cognitive responses. We expected participants' pupil diameter will increase after the hand raising activation compared to before.
- **H2** We hypothesized that hand raising will affect participants' visual attention and visual search behavior, as reflected in fixations and saccades. We therefore expected to observe a significant effect of animation in these measures, i.e., longer fixation durations, more saccades, longer saccade durations, and larger saccade amplitudes.
- **H3** We hypothesized that participants' visual attention to the virtual objects will change after the onset of hand raising, shifting attention from instructional content (i.e., virtual teacher and screen) to peers. Therefore, we expected an increase in dwell time on peers as well as the number of peers that participants fixated on, in other words, a decrease in the dwell time on instructional content.
- **H4** We hypothesized not only that animated virtual peers will have effects on participants' behavior, but also that different numbers of peer animations (four hand-raising conditions) will affect participants differently.

4 RESULTS

To examine the effects of peer animations (i.e., hand raising) on participants' cognitive responses and visual attention, we measured eye movements and pupil diameters before and after the onset of hand raising. Previous studies have shown that task-evoked pupillary responses (TEPR), which index cognitive processing load, have a latency of several hundred milliseconds across tasks [5, 24]. Therefore, given the duration ($\approx$ 2 seconds) of the hand raising animation, we assessed participants' gaze behavior within time windows of 2.5 seconds before and after the hand raising activation. Dependent variables including pupil diameter, fixations, saccades, dwell times, and the number of virtual peers participants fixated on were examined.

First, to investigate whether animated peer learners influenced participants, we compared these dependent variables before and after the onset of hand raising based on all experimental data. Variables used for statistical analysis were described as *V_before* and *V_after*. For this purpose, we used a within-subjects design with one-tailed t-tests. Specifically, a paired t-test was performed for normally distributed data; a Wilcoxon signed-rank test was performed for non-normally distributed data. Furthermore, to verify whether such effects existed in each hand-raising condition, we performed the same statistical analysis (paired t-test) separately for each condition. The statistical significance (see below) that we found based on all experimental data was also found in each hand-raising condition, meaning that the animated peer learners in the VR classroom influenced the participants regardless of the number of animations (20%, 35%, 65% or 80% of virtual peer learners are pre-programmed with hand raising). Given the length of the paper, we have included the detailed paired t-test results for each condition in the Appendix.

Second, since peer learner animations were found to have an effect on participants, as indicated by significant differences in the paired t-tests (see below), we were interested in examining how such effects differed between conditions. We compared the magnitude of these effects across the four hand-raising conditions. Specifically, we calculated the change scores of the above dependent variables before and after activation of hand raising as new variables, described as *V_change* (= *V_after*−*V_before*). For comparison of *V_change* between four groups, we performed a one-way between-subjects ANOVA and the Bonferroni-corrected Tukey-Kramer test as a post-hoc test for the pairwise comparisons. For non-normally distributed data, the Kruskal-Wallis H test was used as a non-parametric version of ANOVA and the Bonferroni-corrected Dunn's Test as a post-hoc test for pairwise comparisons.

All statistical analyses were performed using SciPy [50], an open-source Python[2] library. The significance level was set at $\alpha = 0.05$ for all tests. Asterisks in Figure 3 and Figure 4 indicate significant differences (*, **, *** and n.s. for $p < .05$, $p < .01$, $p < .001$, and no statistical significance, respectively).

4.1 Pupil Diameter

To gain direct insight into the changes in pupil diameter during the virtual lesson, we plotted the normalized pupil diameter over time. As shown in Figure 2, the blue dots represent the mean normalized pupil diameter of all participants at specific time points (every 100*ms*), while the red rectangles each mark the 2.5-second time window after the onset of hand raising. As shown, pupil diameter increases dramatically after the onset of hand raising compared to before peer learners raised their hands. The mean normalized pupil

[2] https://www.python.org/

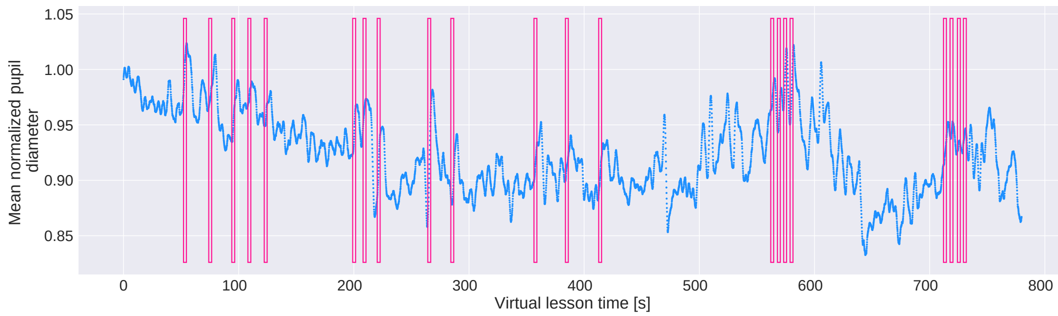


Figure 2: Pupil diameter changes throughout the virtual lesson.

diameter before and after the onset of hand raising (i.e., *Pupil_before* and *Pupil_after*) was calculated for each participant. As shown in Figure 3a, *Pupil_after* ($M = 0.95$, $SD = 0.17$) is significantly larger than *Pupil_before* ($M = 0.93$, $SD = 0.17$), with $t = 1,224$, $p < .001$.

Moreover, we calculated the change scores of mean normalized pupil diameter (i.e., *Pupil_change*=*Pupil_after*−*Pupil_before*) for each participant. A Kruskal-Wallis test revealed a statistically significant difference between groups in the change in mean normalized pupil diameter ($H(3) = 13.59$, $p < .01$). Notably, *Pupil_change* in the 80% condition ($M = 0.0243$, $SD = 0.011$) is significantly greater than in the 20% ($M = 0.0163$, $SD = 0.009$), 35% ($M = 0.0168$, $SD = 0.009$), and 65% ($M = 0.0168$, $SD = 0.011$) conditions, as shown in Figure 4a.

4.2 Fixation and Saccade

4.2.1 Fixation. The mean fixation duration before and after the onset of hand raising (i.e., *FixaDur_before* and *FixaDur_after*) was calculated for each participant. As shown in Figure 3b, *FixaDur_after* ($M = 196.38ms$, $SD = 96ms$) is significantly longer than *FixaDur_before* ($M = 192.54ms$, $SD = 95ms$), with $t = 8,127$, $p < .05$.

In addition, we calculated the change scores of mean fixation duration (i.e., *FixaDur_change*=*FixaDur_after*−*FixaDur_before*) for each participant. However, no statistical differences in the change in mean fixation duration were found between groups, as shown in Figure 4b.

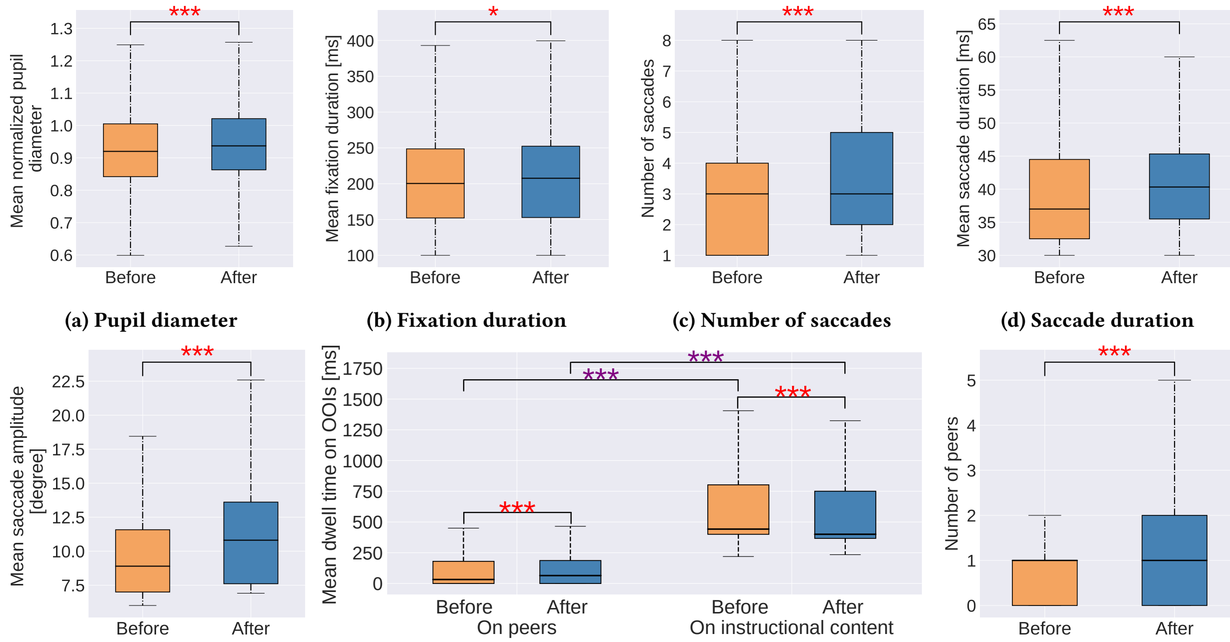
4.2.2 Saccade. We found a significant effect of hand raising on saccade measures, i.e., number of saccades (i.e., *SaccNum_before* and *SaccNum_after*), saccade duration (i.e., *SaccDur_before* and *SaccDur_after*), and saccade amplitude (i.e., *SaccAmpli_before* and *SaccAmpli_after*), as shown in Figure 3c, Figure 3d, and Figure 3e, respectively. In particular, there are significantly more saccades after the onset of hand raising ($M = 3.29$, $SD = 1.75$) than before ($M = 2.99$, $SD = 1.73$), with $t = 1,184$, $p < .001$. *SaccDur_after* ($M = 42.35ms$, $SD = 9.95ms$) is significantly longer than *SaccDur_before* ($M = 39.97ms$, $SD = 10.08ms$), with $t = 6,597$, $p < .001$. Similarly, *SaccAmpli_after* ($M = 11.67deg$, $SD = 4.63deg$) is significantly greater than *SaccAmpli_before* ($M = 10.07deg$, $SD = 3.91deg$), with $t = 5,926$, $p < .001$.

Furthermore, we calculated the change scores of saccade variables (i.e., *SaccNum_change*, *SaccDur_change*, *SaccAmpli_change*, where *Sacc_change*=*Sacc_after*-*Sacc_before*) for each participant. A Kruskal-Wallis test revealed a statistical significance between groups in the change in mean saccade amplitude ($H(3) = 19.13$, $p < .001$). Notably, *SaccAmpli_change* in the 20% condition ($M = 2.13deg$, $SD = 5.76deg$) is significantly greater than in the 35% ($M = 1.37deg$, $SD = 6.05deg$), 65% ($M = 1.53deg$, $SD = 5.76deg$), and 80% ($M = 1.36deg$, $SD = 5.56deg$) conditions. However, no statistical significance was found between groups with respect to the change in saccade number and mean saccade duration. The results are depicted in Figure 4c, Figure 4d, and Figure 4e, respectively.

4.3 Eye Movements on OOIs

4.3.1 First fixation and dwell time on OOIs. We found that after the activation of hand raising, participants initially focused their attention more on the virtual peers OOIs than on the instructional content OOIs (i.e., virtual teacher and screen), as evidenced by the fact that 76% of their first fixations occurred in peer OOIs. Moreover, the time to the first fixation (TTFF) was significantly shorter in peer OOIs ($M = 0.45s$, $SD = 0.21s$) than in instructional content OOIs ($M = 0.93s$, $SD = 0.27s$), with $t = 4,847$, $p < .001$. Although no statistical difference in the first fixation duration (FFD) was found, a longer mean FFD was observed in peer OOIs ($M = 251.32ms$, $SD = 110.55ms$) than in instructional OOIs ($M = 242.56ms$, $SD = 102.09ms$).

A significant effect of hand raising on participants' dwelling behavior to peers (i.e., *DwellPeer_before* and *DwellPeer_after*) and instructional content (i.e., *DwellInstr_before* and *DwellInstr_after*) was found, as shown in Figure 3f. *DwellPeer_after* ($M = 146.08ms$, $SD = 59.07ms$) is significantly longer than *DwellPeer_before* ($M = 132.96ms$, $SD = 65.85ms$), with $t = 2,927$, $p < .001$. In contrast, participants were found to pay decreased attention to the instructional content after peers raised their hands, with *DwellInstr_after* ($M = 607.65ms$, $SD = 368.39ms$) being significantly shorter than *DwellInstr_before* ($M = 638.77ms$, $SD = 383.99ms$), with $t = 5,625$, $p < .001$. Notably, dwell time on instructional content was found to be significantly longer than dwell time on peers, regardless of whether or not there were hand raising animations. As seen in Figure 3f, *DwellInstr_before* ($M = 638.77ms$, $SD = 383.99ms$) is significantly longer than *DwellPeer_before* ($M = 132.96ms$, $SD = 65.85ms$), with $t = 9,052$, $p < .001$; *DwellInstr_after* ($M = 607.65ms$, $SD = 368.39ms$) is significantly longer than *DwellPeer_after* ($M = 146.08ms$, $SD = 59.07ms$), with $t = 10,610$, $p < .001$.



(a) Pupil diameter (b) Fixation duration (c) Number of saccades (d) Saccade duration

(e) Saccade amplitude (f) Dwell on peers and instructional content (g) Number of peers

Figure 3: Statistical comparison of the dependent variables between the conditions before and after the start of the hand raising animations

. Dependent variables include (a) pupil diameter, (b) fixation duration, (c) number of saccades, (d) saccade duration, (e) saccade amplitude, (f) dwell on peers and instructional content, and (g) number of peers fixated by participants.

In addition, we calculated the change scores of dwell time on peers (*DwellPeer_change=DwellPeer_after-DwellPeer_before*) as well as on instructional content (i.e., the virtual teacher and screen) (*DwellInstr_change=DwellInstr_after-DwellInstr_before*). A Kruskal-Wallis test revealed a significant difference between groups in the change in mean dwell time on instructional content ($H(3) = 56.96$, $p < .001$), as shown in Figure 4g. Notably, *DwellInstr_change* (absolute value) is significantly greater in the 80% condition ($M = -48.44ms$, $SD = 38.93ms$) than in the 20% ($M = -22.51ms$, $SD = 47.74ms$), 35% ($M = -26.98ms$, $SD = 36.02ms$), and 65% ($M = -24.22ms$, $SD = 45.56ms$) conditions. However, no statistical differences were observed in the change in mean dwell time on peers between groups (see Figure 4f).

4.3.2 Number of Peers Fixated by Participants. We calculated the number of virtual peers that participants fixated on before and after the onset of hand raising (i.e., *NumPeer_before* and *NumPeer_after*). As shown in Figure 3g, we found a statistically significant increase in the number of peers that participants fixated on, with *NumPeer_after* ($M = 1.40$, $SD = 0.91$) being significantly larger than *NumPeer_before* ($M = 0.92$, $SD = 0.61$), with $t = 1,344$, $p < .001$.

Moreover, we calculated the change scores of the number of peers fixated by participants (*NumPeer_change=NumPeer_after-NumPeer_before*). A Kruskal-Wallis test revealed a statistically significant difference between groups in the change in the number of peers fixated by participants ($H(3) = 187.09$, $p < .001$). The *NumPeer_change* is significantly greater in the 80% condition ($M = 1.25$, $SD = 2.39$) than in 20% ($M = 0.11$, $SD = 1.51$), 35% ($M = 0.11$, $SD = 1.38$), and 65% ($M = 0.32$, $SD = 1.47$) conditions, as shown in Figure 4h.

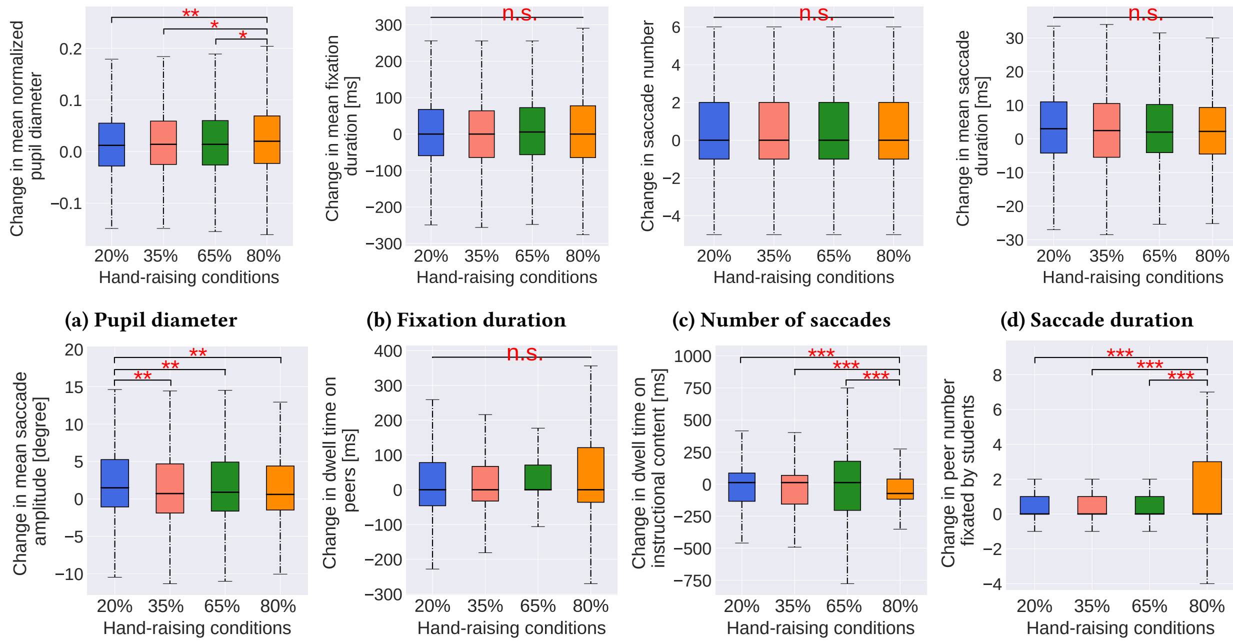
5 DISCUSSION

5.1 Results Discussion

In this section, we discuss the results based on the previously postulated hypotheses **H1-H4**. Overall, the results show that the preprogrammed social animation of virtual peers (i.e., hand raising in response to the virtual teacher's questions) had an effect on participants' cognitive responses and visual attention behaviors during the VR classroom experience, as indicated by various eye movements and pupil measures (see Figure 3). The magnitude of this effect differed between four hand-raising groups (see Figure 4).

Specifically, it was found that participants' pupil diameter increased significantly after the onset of hand raising (see Figure 2 and Figure 3a). Since pupil diameter is often reported as an indicator of cognitive load [9, 31], this significant increase in pupil diameter suggests that participants have increased cognitive processing load and exert more cognitive effort when exposed to the social information that needs to be processed [13], which can be further confirmed by the results of other eye movement measures. Thus, our first hypothesis **H1** is confirmed.

Analysis of fixation and saccade measures allowed investigation of the effects of peer animations on participants' visual attention



(a) Pupil diameter (b) Fixation duration (c) Number of saccades (d) Saccade duration

(e) Saccade amplitude (f) Dwell on peers (g) Dwell on teacher and screen (h) Number of peers

Figure 4: Statistical comparison of the change scores of pupil and eye movement variables before and after the onset of hand raising between four hand-raising groups, including change scores of (a) pupil diameter, (b) fixation duration, (c) number of saccades, (d) saccade duration, (e) saccade amplitude, (f) dwell on peers, (g) dwell on teacher and screen, and (h) number of peers fixated by participants.

and visual search behavior. First, we found that participants exhibited longer fixation duration after hand raising was activated (see Figure 3b), suggesting that they had longer cognitive processing time and more attentive behavior [12, 26] while immersed in a VR classroom with animated virtual peers that were animated with more hand raising. This could be due to participants directing their attention to the raised hands around them. And this social information conveyed by their peers requires more processing time, which consequently implies additional cognitive effort [25, 42]. The increase in information processing time is consistent with the increase in cognitive processing load reflected in pupil diameter. Second, it was found that participants showed significantly different visual search behavior when exposed to hand raising animations, which was reflected in saccade measures (see Figure 3c, Figure 3d, Figure 3e). Participants were found to exhibit more saccades after peers raised their hands, suggesting that their visual spatial attention was influenced and that they showed more visual search behavior [16]. In addition, saccades with greater amplitude were detected after the onset of hand raising, which is another indication that participants' attention might be drawn to more salient attraction cues from a distance around them [17]. These results support our second hypothesis **H2**.

Before the peer learners raise their hands, the virtual teacher and the screen displaying the learning content are the main focus of the participants' attention. After being confronted with social information from peers (i.e., hand raising), participants gave higher visual priority to the animated peers than to the instructional content to which they paid attention before hand raising, as indicated by their first-fixation behavior [21]. This visual priority is further evidenced by the time to the first fixation (TTFF) within OOIs. Furthermore, participants were found to shift their attention from instructional content to their peers (see Figure 3f). As a consequence, the number of peers participants fixated on was found to increase significantly after hand raising (see Figure 3g), further supporting our attention shift hypothesis. The change in visual priority is consistent with the shift in visual attention reflected in dwell time measures, and they can be parsed together. Taken together, all these results support our third hypothesis **H3**. Surprisingly, it is noteworthy that although participants' attention was captured by their peers' social animations, their attention remained primarily focused on the learning content (i.e., virtual teacher and screen) while the hand raising was occurring (see Figure 3f). This indicates that regardless of the animations, participants exhibited significantly longer attention time on the instructional content than on peers, which in turn implies that their learning activities were still ongoing when they were exposed to social information.

The attention shift results evidencing hypothesis **H3** (attention shifts from instructional content to peer learners) further support our hypothesis **H1** (cognitive load increases) and hypothesis **H2** (visual attention and visual search behavior change). When participants were exposed to peer animation (i.e., hand raising) intended to increase immersion and authenticity, their attention inevitably shifted to such animations, resulting in more focused attentional behavior and stronger visual search behavior, which ultimately also led to a change in cognitive processing load. Overall, our results support the first three hypotheses (**H1**-**H3**) that participants respond cognitively and attentionally to animated virtual companions while participating in a VR lesson, as reflected in various eye movements and pupil measures.

Furthermore, we measured the change in all variables after hand raising compared to before and considered it as a measure of the extent to which the animation affected participants. Results show that the hand raising animation affected participants differently across groups (see Figure 4), and this difference is mainly between groups with lower (20%) and higher (80%) number of animations. Specifically, when participants were surrounded by a large number of animated peers (80%) from different directions, they were able to perceive the most animations, and thus this large amount of social information caused participants' cognitive load to increase the most [13]. This is evidenced by significant increase in pupil diameter and number of peers fixated by participants in the 80% condition than in the other conditions. Consequently, participants in groups with a higher (80%) number of animations showed a significantly greater decrease in their visual attention to the lesson content than in groups with a lower (20%) number of animations. This is because that their attention was more likely to shift from the instructional content to virtual peers when they were surrounded by a larger number of animated virtual peers. Therefore, participants in the 20% condition showed a significantly greater increase in saccade amplitude than in the other conditions. Since when participants were surrounded by a small number of animated peers (20%), they had to exert more effort to search for the raised hands, which were slightly more difficult to catch, than when they were surrounded by more animations, resulting in a greater increase in visual search behavior [16, 17]. However, no significant difference were found in the increase in fixation duration, saccade number, saccade duration, and dwell time on peers between groups. Thus, our fourth hypothesis **H4** is partially validated by these results. Our results indicate that participants' cognitive and visual attention responses elicited by virtual avatar animations are related to the number of animations, which could provide insights for designing VR environment with optimal virtual avatar animations.

In this study, it was found that animated peer learners significantly affect learners from several aspects, and that these effects are related to the number of animations provided. Such a finding further contributes to how and to what extent virtual peer learners' animations can be implemented in immersive VR classrooms to both enhance authenticity and control the distractions and additional cognitive load that the VR systems impose on learners to an acceptable level. Deciding on the number of social animations depends on the purpose of the designed VR applications, i.e., whether instructional content or social information is more important to learners. Our study does not aim to tell to what extent avatar animations should be provided, but is an exploratory study that provides bases for further studies that need to decide on the animations provided in their VR applications. Such findings provide important insights specifically for designing educational VR applications by implementing certain avatar animations related to users' learning behaviors (e.g., hand raising of peer learners), but also more generally for designing avatar animations in other VR-based systems, such as animated avatars in contexts like medical training simulation [52, 53]. Eye-tracking technology offers a viable way to investigate this research question by extracting users' temporal eye movements that are indicative of various human behaviors such as visual attention, visual search, as well as cognitive processing load.

5.2 Limitations and Future Work

Our immersive VR classroom was built with animated virtual teacher and peer learners to provide students with a high level of immersion. However, since we aim to explore how students respond to the animations of their primary social counterparts (i.e., peer learners), all animations were pre-programmed in the spirit of controllability. However, this may reduce students' sense of immersion since the teacher does not call on them when they raise their hand. Although avatar animations were found to have an impact on students' cognitive and visual attention behaviors, and this impact was related to the number of animations, it is unclear whether students' learning outcome is also associated with and affected by avatar animations. In our future work, we aim to investigate this question as it will provide guidance on the optimal design of VR environments for educational purposes and thus maximize the efficiency of student learning in such VR environments.

6 CONCLUSION

In this paper, we designed an immersive VR classroom with pre-programmed animated virtual avatars to investigate how social interactions between the virtual teacher and peers (i.e., peer hand raising in response to the teacher's questions) affect students' cognitive and visual attention behaviors during a virtual lesson. To this end, eye movements and pupil measures were analyzed. We found that peers' hand raising had an effect on students' behavior in several aspects, including increased cognitive load, attentional shift from instructional contents (i.e., virtual teacher and screen) to peers, and increased information processing time and visual search behavior. The effects of hand raising on various aspects of students reflected in different measures are interrelated. In addition, we found that the magnitude of such effects is related to the number of animations (number of animated peer learners).

In our study, we developed an immersive VR classroom that closely resembles a real classroom by creating not only a virtual teacher but also a set of animated virtual peer learners with social information (i.e., hand raising) to enhance immersion and authenticity. Our research provides a methodological foundation for investigating students' instantaneous and intuitive responses to virtual human animations (particularly virtual peers) during VR experience using eye movements. Our results imply that the effects of avatar animation should be considered by developers when presenting animated virtual humans to enhance immersion or for interaction purposes in VR environments. Overall, these findings have important implications for future studies aiming to create more effective, interactive, and authentic VR-based (learning) environments.

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 was a doctoral candidate and supported by the LEAD Graduate School and 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. Hong Gao is a doctoral candidate and is supported by the University of Tübingen. This research was also partly funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC number 2064/1 - Project number 390727645. The authors thank Jens-Uwe Hahn, 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] Rebecca Adams, Paul Finn, Elisabeth Moes, Kathleen Flannery, and Albert "Skip" Rizzo. 2009. Distractibility in Attention/Deficit/ Hyperactivity Disorder (ADHD): The Virtual Reality Classroom. *Child Neuropsychology* 15, 2 (2009), 120–135. https://doi.org/10.1080/09297040802169077

[2] Ioannis Agtzidis, Mikhail Startsev, and Michael Dorr. 2019. 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* (Nice, France). ACM, New York, NY, USA, 1007–1015. https://doi.org/10.1145/3343031.3350947

[3] Tobias Appel, Christian Scharinger, Peter Gerjets, and Enkelejda Kasneci. 2018. Cross-subject workload classification using pupil-related measures. In *Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications* (Warsaw, Poland). ACM, New York, NY, USA, Article 4, 8 pages. https://doi.org/10.1145/3204493.3204531

[4] Tobias Appel, Natalia Sevcenko, Franz Wortha, Katerina Tsarava, Korbinian Moeller, Manuel Ninaus, Enkelejda Kasneci, and Peter Gerjets. 2019. Predicting Cognitive Load in an Emergency Simulation Based on Behavioral and Physiological Measures. In *2019 International Conference on Multimodal Interaction* (Suzhou, China) *(ICMI '19)*. ACM, New York, NY, USA, 154–163. https://doi.org/10.1145/3340555.3353735

[5] Jackson Beatty. 1982. Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological bulletin* 91, 2 (1982), 276.

[6] Efe Bozkir, Philipp Stark, Hong Gao, Lisa Hasenbein, Jens-Uwe Hahn, Enkelejda Kasneci, and Richard Göllner. 2021. Exploiting Object-of-Interest Information to Understand Attention in VR Classrooms. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*. IEEE, New York, NY, USA, 597–605. https://doi.org/10.1109/VR50410.2021.00085

[7] Ricardo Böheim, Tim Urdan, Maximilian Knogler, and Tina Seidel. 2020. Student hand-raising as an indicator of behavioral engagement and its role in classroom learning. *Contemporary Educational Psychology* 62 (2020), 101894. https://doi.org/10.1016/j.cedpsych.2020.101894

[8] Nora Castner, Tobias Appel, Thérése Eder, Juliane Richter, Katharina Scheiter, Constanze Keutel, Fabian Hüttig, Andrew Duchowski, and Enkelejda Kasneci. 2020. Pupil diameter differentiates expertise in dental radiography visual search. *PloS one* 15, 5 (2020), 1–19. https://doi.org/10.1371/journal.pone.0223941

[9] Siyuan Chen and Julien Epps. 2014. Using Task-Induced Pupil Diameter and Blink Rate to Infer Cognitive Load. *Human–Computer Interaction* 29, 4 (2014), 390–413. https://doi.org/10.1080/07370024.2014.892428 arXiv:https://doi.org/10.1080/07370024.2014.892428

[10] Kun-Hung Cheng and Chin-Chung Tsai. 2019. A case study of immersive virtual field trips in an elementary classroom: Students' learning experience and teacher-student interaction behaviors. *Computers & Education* 140 (2019), 103600. https://doi.org/10.1016/j.compedu.2019.103600

[11] Kuei-Pin Chien, Cheng-Yue Tsai, Hsiu-Ling Chen, Wen-Hua Chang, and Sufen Chen. 2015. Learning differences and eye fixation patterns in virtual and physical science laboratories. *Computers & Education* 82 (2015), 191–201. https://doi.org/10.1016/j.compedu.2014.11.023

[12] Kuei-Pin Chien, Cheng-Yue Tsai, Hsiu-Ling Chen, Wen-Hua Chang, and Sufen Chen. 2015. Learning differences and eye fixation patterns in virtual and physical science laboratories. *Computers & Education* 82 (2015), 191–201. https://doi.org/10.1016/j.compedu.2014.11.023

[13] Pieternel Dijkstra, Hans Kuyper, Greetje van der Werf, Abraham P. Buunk, and Yvonne G. van der Zee. 2008. Social Comparison in the Classroom: A Review. *Review of Educational Research* 78, 4 (2008), 828–879. https://doi.org/10.3102/0034654308321210 arXiv:https://doi.org/10.3102/0034654308321210

[14] Hong Gao, Efe Bozkir, Lisa Hasenbein, Jens-Uwe Hahn, Richard Göllner, and Enkelejda Kasneci. 2021. Digital Transformations of Classrooms in Virtual Reality. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. ACM, New York, NY, USA, Article 483, 10 pages. https://doi.org/10.1145/3411764.3445596

[15] Joseph H. Goldberg and Xerxes P. Kotval. 1999. Computer interface evaluation using eye movements: methods and constructs. *International Journal of Industrial Ergonomics* 24, 6 (1999), 631–645. https://doi.org/10.1016/S0169-8141(98)00068-7

[16] Joseph H Goldberg and Xerxes P Kotval. 1999. Computer interface evaluation using eye movements: methods and constructs. *International Journal of Industrial Ergonomics* 24, 6 (1999), 631–645. https://doi.org/10.1016/S0169-8141(98)00068-7

[17] Joseph H. Goldberg, Mark J. Stimson, Marion Lewenstein, Neil Scott, and Anna M. Wichansky. 2002. Eye Tracking in Web Search Tasks: Design Implications. In *Proceedings of the 2002 Symposium on Eye Tracking Research & amp; Applications* (New Orleans, Louisiana) *(ETRA '02)*. ACM, New York, NY, USA, 51–58. https://doi.org/10.1145/507072.507082

[18] Patricia Goldberg, Ömer Sümer, Kathleen Stürmer, Wolfgang Wagner, Richard Göllner, Peter Gerjets, Enkelejda Kasneci, and Ulrich Trautwein. 2021. Attentive or not? Toward a machine learning approach to assessing students' visible engagement in classroom instruction. *Educational Psychology Review* 33, 1 (2021), 27–49. https://doi.org/10.1007/s10648-019-09514-z

[19] Esteban Gutierrez Mlot, Hamed Bahmani, Siegfried Wahl, and Enkelejda Kasneci. 2016. 3D Gaze Estimation Using Eye Vergence. In *Proceedings of the International Joint Conference on Biomedical Engineering Systems and Technologies* (Rome, Italy) *(BIOSTEC 2016)*. SCITEPRESS - Science and Technology Publications, Lda, Setubal, PRT, 125–131. https://doi.org/10.5220/0005821201250131

[20] Lisa Hasenbein, Philipp Stark, Ulrich Trautwein, Anna Carolina Muller Queiroz, Jeremy Bailenson, Jens-Uwe Hahn, and Richard Göllner. 2022. Learning with simulated virtual classmates: Effects of social-related configurations on students' visual attention and learning experiences in an immersive virtual reality classroom. *Computers in Human Behavior* 133 (2022), 107282. https://doi.org/10.1016/j.chb.2022.107282

[21] Kenneth Holmqvist, Marcus Nyström, Richard Andersson, Richard Dewhurst, Jarodzka Halszka, and Joost van de Weijer. 2011. *Eye Tracking : A Comprehensive Guide to Methods and Measures*. Oxford University Press, United Kingdom.

[22] Kenneth Holmqvist, Saga Lee Örbom, Ignace TC Hooge, Diederick C Niehorster, Robert G Alexander, Richard Andersson, Jeroen S Benjamins, Pieter Blignaut, Anne-Marie Brouwer, Lewis L Chuang, et al. 2022. Eye tracking: empirical foundations for a minimal reporting guideline. *Behavior research methods* (2022), 1–53. https://doi.org/10.3758/s13428-021-01762-8

[23] Polina Häfner, Victor Häfner, and Jivka Ovtcharova. 2013. Teaching Methodology for Virtual Reality Practical Course in Engineering Education. *Procedia Computer Science* 25 (2013), 251–260. https://doi.org/10.1016/j.procs.2013.11.031 2013 International Conference on Virtual and Augmented Reality in Education.

[24] Silvia L. Isabella, Charline Urbain, J. Allan Cheyne, and Douglas Cheyne. 2019. Pupillary responses and reaction times index different cognitive processes in a combined Go/Switch incidental learning task. *Neuropsychologia* 127 (2019), 48–56. https://doi.org/10.1016/j.neuropsychologia.2019.02.007

[25] Arthur M Jacobs. 1986. Eye-movement control in visual search: How direct is visual span control? *Perception & Psychophysics* 39, 1 (1986), 47–58.

[26] Marcel A. Just and Patricia A. Carpenter. 1976. Eye fixations and cognitive processes. *Cognitive Psychology* 8, 4 (1976), 441–480. https://doi.org/10.1016/0010-0285(76)90015-3

[27] Sin-Hwa Kang and Jonathan Gratch. 2010. The Effect of Avatar Realism of Virtual Humans on Self-Disclosure in Anonymous Social Interactions. In *CHI '10 Extended Abstracts on Human Factors in Computing Systems* (Atlanta, Georgia, USA) *(CHI EA '10)*. ACM, New York, NY, USA, 3781–3786. https://doi.org/10.1145/1753846.1754056

[28] Enkelejda Kasneci, Gjergji Kasneci, Ulrich Trautwein, Tobias Appel, Maike Tibus, Susanne M. Jaeggi, and Peter Gerjets. 2022. Do your eye movements reveal your performance on an IQ test? A study linking eye movements and socio-demographic information to fluid intelligence. *PLOS ONE* 17, 3 (03 2022), 1–23. https://doi.org/10.1371/journal.pone.0264316

[29] Sam Kavanagh, Andrew Luxton-Reilly, Burkhard Wuensche, and Beryl Plimmer. 2017. A systematic review of Virtual Reality in education. *Themes in Science and Technology Education* 10, 2 (December 2017), 85–119. https://www.learntechlib.org/p/182115

[30] Fengfeng Ke, Sungwoong Lee, and Xinhao Xu. 2016. Teaching training in a mixed-reality integrated learning environment. *Computers in Human Behavior* 62 (2016), 212–220. https://doi.org/10.1016/j.chb.2016.03.094

[31] Krzysztof Krejtz, Andrew T. Duchowski, Anna Niedzielska, Cezary Biele, and Izabela Krejtz. 2018. Eye tracking cognitive load using pupil diameter and microsaccades with fixed gaze. *PLOS ONE* 13, 9 (09 2018), 1–23. https://doi.org/10.1371/journal.pone.0203629

[32] Meng-Yun Liao, Ching-Ying Sung, Hao-Chuan Wang, and Wen-Chieh Lin. 2019. 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) (Osaka, Japan). IEEE, New York, NY, USA, 163–171. https://doi.org/10.1109/VR.2019.8797708

[33] Wei-Kai Liou and Chun-Yen Chang. 2018. Virtual reality classroom applied to science education. In *2018 23rd International Scientific-Professional Conference on Information Technology (IT)*. IEEE, New York, NY, USA, 1–4. https://doi.org/10.1109/SPIT.2018.8350861

[34] Tzu Chien Liu, Yi Chun Lin, Tzu Ning Wang, Shih Ching Yeh, and Slava Kalyuga. 2021. Studying the effect of redundancy in a virtual reality classroom. *Educational Technology Research and Development* 69, 2 (2021), 1183–1200. https://doi.org/10.1007/s11423-021-09991-6

[35] Matthew Lombard, Theresa B Ditton, and Lisa Weinstein. 2009. Measuring presence: the temple presence inventory. In *Proceedings of the 12th annual international workshop on presence*. 1–15.

[36] David M. Markowitz, Rob Laha, Brian P. Perone, Roy D. Pea, and Jeremy N. Bailenson. 2018. Immersive Virtual Reality Field Trips Facilitate Learning About Climate Change. *Frontiers in Psychology* 9 (2018), 2364. https://doi.org/10.3389/fpsyg.2018.02364

[37] Sebastiaan Mathôt, Jasper Fabius, Elle Van Heusden, and Stefan Van der Stigchel. 2018. Safe and sensible preprocessing and baseline correction of pupil-size data. *Behavior Research Methods* 50, 1 (2018), 94–106. https://doi.org/10.3758/s13428-017-1007-2

[38] Shivsevak Negi and Ritayan Mitra. 2020. Fixation duration and the learning process: an eye tracking study with subtitled videos. *Journal of Eye Movement Research* 13, 6 (2020). https://doi.org/10.16910/jemr.13.6.1

[39] Jaziar Radianti, Tim A. Majchrzak, Jennifer Fromm, and Isabell Wohlgenannt. 2020. A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers & Education* 147 (2020), 103778. https://doi.org/10.1016/j.compedu.2019.103778

[40] 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. 2000. The Virtual Classroom: A Virtual Reality Environment for the Assessment and Rehabilitation of Attention Deficits. *CyberPsychology & Behavior* 3, 3 (2000), 483–499. https://doi.org/10.1089/10949310050078940

[41] Andrew Robb, Regis Kopper, Ravi Ambani, Farda Qayyum, David Lind, Li-Ming Su, and Benjamin Lok. 2013. Leveraging Virtual Humans to Effectively Prepare Learners for Stressful Interpersonal Experiences. *IEEE Transactions on Visualization and Computer Graphics* 19, 4 (2013), 662–670. https://doi.org/10.1109/TVCG.2013.35

[42] Timothy A Salthouse, Cecil L Ellis, Donald C Diener, and Benjamin L Somberg. 1981. Stimulus processing during eye fixations. *Journal of Experimental Psychology: Human Perception and Performance* 7, 3 (1981), 611.

[43] Abraham Savitzky and Marcel J. E. Golay. 1964. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry* 36 (1964), 1627–1639. https://doi.org/10.1021/ac60214a047

[44] Thomas Schubert, Frank Friedmann, and Holger Regenbrecht. 2001. The Experience of Presence: Factor Analytic Insights. *Presence: Teleoperators and Virtual Environments* 10, 3 (06 2001), 266–281. https://doi.org/10.1162/105474601300343603

[45] Sharad Sharma, Ruth Agada, and Jeff Ruffin. 2013. Virtual reality classroom as an constructivist approach. In *2013 Proceedings of IEEE Southeastcon* (Jacksonville, FL, USA). IEEE, New York, NY, USA, 1–5. https://doi.org/10.1109/SECON.2013.6567441

[46] Gurjinder Singh, Archana Mantri, Ojaswa Sharma, and Rashpinder Kaur. 2021. Virtual reality learning environment for enhancing electronics engineering laboratory experience. *Computer Applications in Engineering Education* 29, 1 (2021), 229–243. https://doi.org/10.1002/cae.22333 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/cae.22333

[47] Mel Slater, Angus Antley, Adam Davison, David Swapp, Christoph Guger, Chris Barker, Nancy Pistrang, and Maria V. Sanchez-Vives. 2006. A Virtual Reprise of the Stanley Milgram Obedience Experiments. *PLOS ONE* 1, 1 (12 2006), 1–10. https://doi.org/10.1371/journal.pone.0000039

[48] Ömer Sümer, Patricia Goldberg, Sidney D'Mello, Peter Gerjets, Ulrich Trautwein, and Enkelejda Kasneci. 2021. Multimodal Engagement Analysis from Facial Videos in the Classroom. *IEEE Transactions on Affective Computing* (2021), 1–1. https://doi.org/10.1109/TAFFC.2021.3127692

[49] Ulrich Trautwein, Hanna Dumont, and Anna–Lena Dicke. 2015. Schooling: Impact on Cognitive and Motivational Development. In *International Encyclopedia of the Social & Behavioral Sciences (Second Edition)* (second edition ed.), James D. Wright (Ed.). Elsevier, Oxford, 119–124. https://doi.org/10.1016/B978-0-08-097086-8.26056-X

[50] Pauli Virtanen, Ralf Gommers, Travis E Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, et al. 2020. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature methods* 17, 3 (2020), 261–272.

[51] M. Volante, S. V. Babu, H. Chaturvedi, N. Newsome, E. Ebrahimi, T. Roy, S. B. Daily, and T. Fasolino. 2016. Effects of Virtual Human Appearance Fidelity on Emotion Contagion in Affective Inter-Personal Simulations. *IEEE Transactions on Visualization & Computer Graphics* 22, 04 (04 2016), 1326–1335. https://doi.org/10.1109/TVCG.2016.2518158

[52] Matias Volonte, Yu-Chun Hsu, Kuan-Yu Liu, Joe P. Mazer, Sai-Keung Wong, and Sabarish V. Babu. 2020. Effects of Interacting with a Crowd of Emotional Virtual Humans on Users' Affective and Non-Verbal Behaviors. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, New York, NY, USA, 293–302. https://doi.org/10.1109/VR46266.2020.00049

[53] Matias Volonte, Andrew Robb, Andrew T. Duchowski, and Sabarish V. Babu. 2018. Empirical Evaluation of Virtual Human Conversational and Affective Animations on Visual Attention in Inter-Personal Simulations. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, New York, NY, USA, 25–32. https://doi.org/10.1109/VR.2018.8446364

[54] Yanxiang Wu, Sabarish V. Babu, Rowan Armstrong, Jeffrey W. Bertrand, Jun Luo, Tania Roy, Shaundra B. Daily, Lauren Cairco Dukes, Larry F. Hodges, and Tracy Fasolino. 2014. Effects of Virtual Human Animation on Emotion Contagion in Simulated Inter-Personal Experiences. *IEEE Transactions on Visualization and Computer Graphics* 20, 4 (2014), 626–635. https://doi.org/10.1109/TVCG.2014.19