Comparing Visualizations to Help a Teacher Effectively Monitor Students in a VR Classroom

Yitoshee Rahman * Arun K Kulshreshth † Christoph W Borst ‡

University of Louisiana at Lafayette, Lafayette, LA, United States

ABSTRACT

Educational virtual reality (VR) applications are the most recent addition to the learning management tools in this modern age. Due to health concerns, financial concerns, and convenience, people are looking for alternate ways to teach and learn. An efficient VR-based teaching interface could enhance student engagement, learning outcomes, and overall educational experience. Typically, teachers in a VR classroom do not have a way to know what students are doing since students are not visible. An efficient teaching interface should include some mechanism for a teacher to monitor students and alert the teacher if a student is trying to catch the attention of the teacher. An ideal interface would be one, which helps a teacher effectively monitor students while teaching without increasing the cognitive load of the teacher. In this paper, we present a comparative study of two such student monitoring interfaces. In the first interface, the student activity related information is shown using icons near the student avatar (representing a student in the VR environment). While in the second interface, a set of centrally-arranged emoticon-like visual indicators are present in addition to the student avatar, and the student activity related information is shown near the student emoticon. We present a detailed user experiment comparing the two interfaces in terms of teaching management, student monitoring capability, cognitive load, and user preference. Participants preferred and performed better with Indicator-located interface over avatar-located interface.

Index Terms: Computing methodologies—Computer graphics—Graphics systems and interfaces—Virtual reality; Human-centered computing—Visualization—Visualization design and evaluation methods; Applied computing—Education—Distance learning

1 INTRODUCTION

Educational VR applications are becoming more popular and this area is one of the most recent topics of interest in research . Especially after the COVID-19 pandemic, the educators around the world are looking for better teaching techniques and remote classroom management tools to provide education more efficiently. During the last decade, there has been a renewed increase in virtual reality which in turn made virtual reality devices more and more affordable for consumers. This ease of access to VR devices has made educational VR applications an actual possibility. Moreover, with the recent addition of cheaper wireless VR devices (e.g., Meta Quest Pro, Meta Quest 2, etc.), both teachers and students are realizing the benefits of using VR for educational purposes.

Immersive VR can produce experiences that are vividly remembered and could improve learning. VR can be used for visiting historical sites, foreign countries or for running scientific experiences while being safely inside the user's own home [42]. A VR classroom eradicates the need for travel or being in the same time while giving the students and teachers an opportunity for an immersive teaching and learning experience [20]. VR has been shown to improve instructor teaching skills as well [21]. Nonetheless, VR-based education is subject to certain issues. In a physical classroom setting, teachers can gauge student engagement and behavior through multiple cues, including body language, eye contact, and facial expressions. However, in VR environments, teachers lack direct visibility of their students, leading to a reduced awareness of their activities. In educational VR, the most common representation of a student is an avatar [1,2]. However, a student could park their avatar at any place inside the VR classroom and such student avatars could be scattered around the environment unlike a real classroom where the students are generally seated in an organized way in the viewpoint of the teacher. Thus, in VR, there is no easy way for the teachers to know if a student is trying to attract the attention of the teacher. Additionally, teaching in itself could be a very stressful job and doing it in VR might contribute even more to that stress. To address these challenges, we need a better VR-based interface for the teachers to help them monitor students while teaching without increasing their cognitive load while teaching. The goal is to provide the teachers with an interface that can help them in monitoring the students in the best way possible while not limiting their teaching performance.

Broussard et al. [10] proposed a novel VR teaching interface with various options and information delivery mechanisms for the teachers. This work was focused on the interface design. However, they did not evaluate the effectiveness of their proposed interfaces in a classroom setting. Their preliminary study on user preference for interface elements did not involve any teaching task. Another work by Rahman et al. used a similar interface for a pilot study comparing placement of student information [34]. In this paper, we present a detailed user experiment comparing two interfaces for monitoring students. In the first interface (see Figure 1), student activity indicators such as typing, speaking, or raising a hand (see Figure 3) are shown next to the corresponding student avatar scattered in the VR classroom. In the second interface (see Figure 2), a set of centrally-arranged emoticon like icons corresponding to each student is present in addition to the avatar, and the student activity indicators are shown next to the student emoticons. Our experiment compared the two interfaces, with varying number of students seeking the teachers attention, in terms of teaching management, student monitoring capability, cognitive load, and user preferences.

2 RELATED WORK

Recently there has been an increase in the number of tools that help a teacher manage classrooms better. These tools are of various types including but not limited to ambient displays, wearables, or learning analytics dashboards [5]. More and more educators are taking the help of different learning management tools and applications on their devices to keep track of students, especially with hybrid or virtual classrooms. Several researchers have emphasized the importance of real time classroom awareness and how these interfaces are aiding in improved teaching performance [26, 27, 35, 41, 43]. Education delivery system is no longer limited to regular face to face in person classrooms, rather teachers and students are using diverse educational applications to visualize and analyze student activities and communicate with students [17, 25, 40].

*e-mail: yitoshee@gmail.com
†e-mail: arunkul@louisiana.edu
‡e-mail: cwborst@gmail.com

This is an author formatted version. Original publication in the IEEE International Symposium on Mixed and Augmented Reality (ISMAR), 2023.

Using VR in classrooms is no longer just a possibility but now an actual reality. Having VR classrooms can bring in a dramatic change in the way we teach and learn. This in turn has opened up a whole new scope for research and a recent increase in research for VR in education can be noticed. Yildirim et al. mentioned how virtual reality in classrooms can increase student motivation and creativity as well as promote individual learning [45]. However, they mentioned the need to track and improve the way users i.e., students interact with the environment. Radianti et al. [33] did a focused systematic study on the evaluation of educational VR applications for higher education and discussed how there was an increase in interest in immersive VR technologies for educational purposes. Chavez and Bayona et al. [12] looked at VR in the learning context and examined the characteristics that determine successful implementation of this technology as well as its positive effects on learning outcomes. Jensen et al. [19] provides a review of declarative knowledge, gesture skills, and emotional control while using head mounted displays for educational VR .

The existing literature describes the effectiveness of VR-based instruction and how teachers felt the need to improve the mechanism for students [18, 23, 31, 44]. Research focused more on interaction between students and teaching content and how improving the student interface can result in better learning outcomes. Belani et al. [6] in their recently published work investigated how learning outcome and user experience was affected by different spatial representations of learning content. However, there is very little research on how a VR interface can be made more favorable for the teacher.

There has been some research for using augmented reality and mixed reality technology to help teachers manage classrooms better. Holstein et al. [15] used mixed-reality smart glasses to place floating indicator symbols for actions like (misusing the software, unproductively struggling, struggling, recently doing very well, idle) over empty student seats, and relocatable class-level analytics displays. In the research done by Berque et al. [7], teachers wore google glass, an augmented reality and mixed reality glassware, and could see an aggregate of the students' statuses of their level of understanding: understands", "mostly understands" or "does not understand" in real-time on the eye display and also could see a list of scrollable thumbnails, labeled with the name of the corresponding student, allowing the teachers to see which students are on task . Alavi et al. [3] proposed using an AR lantern that would change color or change intensity of color or blink in different frequencies to indicate student state. However, none of these techniques are well suited for a general real time student monitoring with multiple student activity indicators. Parmar and Bickmore conducted a study that explored different visualizations for attention management in the context of an augmented reality (AR) classroom [30].

Bozkir et al. [8] did a user study that showed how placement of students and their viewpoint contributed to their learning outcome . Alhalabi et al. [4] provided qualitative evidence of the impact of VR over no VR and also the importance of student involvement to enhance student achievement in engineering students . Simeone et al. [36] highlighted the importance of the live teacher inside VR and also indicated contact teaching should improve student motivation and overall performance. This raises the importance of providing the teachers with a VR interface to help them monitor students in realtime.

From the training strategies' perspective, research has been done on the cognitive load of the students using various techniques of classic classroom settings and with using learning management systems and how those can be measured [22, 24, 37]. However not much research has been found that takes into consideration the cognitive load of the teacher while using these applications to teach a lesson. The work of Ouwehand et al. [29] discussed the challenges to quantify cognitive load through various measurement methods and suggested that numerical scales can better reflect cognitive processes underlying complex problems. Use of physiological data such as blink rate, heat flow, galvanic skin response, or heart rate addresses many of the shortcomings associated with traditional techniques like self assessment and NASA-task load index [13]. Research shows heart rate can be used as a measurement of cognitive load [14, 39] and increased heart rate in healthy individuals can be an indicator of increase cognitive load and stress [38]. Thus, for our work we used a combination of physiological sensor data (heart rate monitor), and self-reported data to measure the cognitive load of the teacher while performing the teaching task. Additionally, we also studied how the cognitive load would change as we increase the number of students seeking the teachers' attention. An ideal interface should help a teacher with student monitoring without any significant increase in cognitive load.

3 VR Teaching Interfaces

Our interface was a VR classroom with 30 student avatars placed at arbitrary locations inside the virtual room. This interface is based on the interface proposed by Broussard et al. ([9, 11]). They did a user study where they had different indicators for student representation. The pilot study asked the users if their proposed mechanisms should be present or not, and how those features can be improved. We optimized the interface by selecting the best options based on their preliminary study results and chose the visual indicator that was most liked, and also decided to use tethers to connect visual indicators and student avatars only when students were important. The classroom had a white board which worked as a slideshow presenter which the users could use to teach the class. Students had their names appearing on top of their avatars. Students in the environment would arbitrarily do nothing or present one of the three action states: raising hands, speaking or typing. These action states were represented by action icons (Fig. 3). Students needing the teachers' immediate attention (e.g., to ask a question like in a real world classroom) would raise their hand and we call these as important students (for the teacher). Based on the placement of the action icons, we had two versions of the teaching interface : avatar-located icons and indicator-located icons. These interfaces are described in the following subsections.

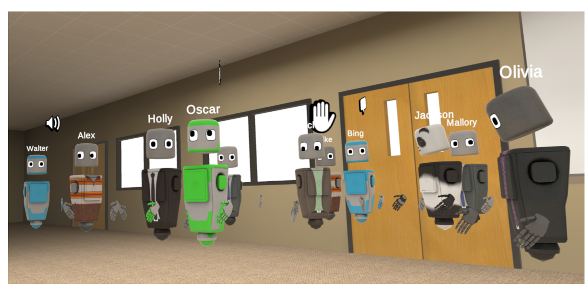

Figure 1: Avatar-located Icons Interface. Student action icons appear on the top of the avatar.

3.1 Avatar-Located (AL) Icons Interface

In this interface (see Figure 1), each student is represented by an avatar. The avatar is placed inside the virtual classroom at a random location. We made sure that no two students will occupy the same physical space. The name of the student is shown on the top of the avatar. The design was motivated by an existing work [16] which proposed a mixed-reality based wearable teacher awareness tool for real-time monitoring of students. Student action icons appeared on top of the avatar. The focus of their work was how it helped the learning outcome of students and how much time the teachers spent in the vicinity of students. But they did not take teachers' preference or performance into account. We used a similar idea for our VR teaching interfaces with both teacher and students in a VR classroom, and students represented by avatars.

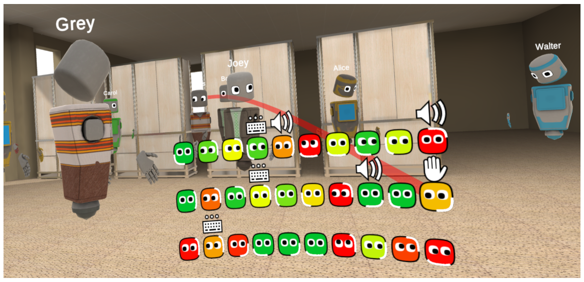

Figure 2: Indicator-located Icons Interface. A set of centrally-arranged emoticon-like visual indicators for each student. Student action icons appear on the top of the emoticon. A red tether, connecting the student emoticon with its avatar, appears when the student raises hand to seek the immediate attention of the teacher.

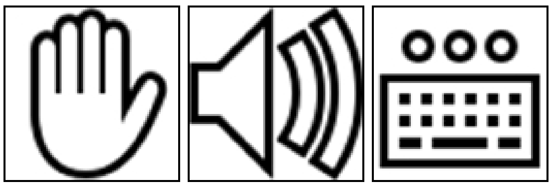

Figure 3: Student activity indicators or action icons: Raising hand, Speaking, and Typing.

Figure 4: A white edge to indicate student avatar position relative to the emoticon from the point of view of the teacher.

3.2 Indicator-located (IL) Icons Interface

In this interface (see Figure 2), all student avatars have a representative emoticon-like visual indicator arranged centrally in front of the user who is playing the role of a teacher. The visual indicators are stacked in multiple rows, 3 rows to be exact with 10 indicators in each row. The indicator set was placed in front of the teacher but slightly below so that it didn't obstruct the teachers view of the classroom. The indicators had colors ranging from red to green based on the simulated students' attention state. The green color represents attentive students and the red represents highly distracted students. A white edge around the sides of the emoticon represents on which side of the teacher the student avatar is placed (see Figure 4). Action icons appear right on the top of the emoticon indicators. A red tether connecting the emoticon and the student avatar appears when a student becomes important (by raising their hand). These indicator and action icons were designed by optimizing the interface proposed by Broussard et al. [9, 11].

4 User Evaluations

Our user study was a within-subjects comparison experiment with an increasing number of important students, needing immediate attention of the teacher, to evaluate the effectiveness of the two interfaces. Our independent variables were interface type (IT) (avatar-located and indicator-located) and the number of important students (NS) (3 options: 1,3, or 5 important student).Thus, in total, we had $2 \times 3 = 6$ conditions for each participant. Each participant had one trial for each interface type, in counterbalanced order, with varying numbers of important students in random order. Both trials took about 5-10 minutes based on how long it took the participant to complete the lesson they were asked to teach. For both trials, simulated students would randomly raise their hands, speak or type. Students with raised hands were considered as important students and the number of important students would be one, three and five during the lesson. Six other random students could be found typing or speaking at any random time to emulate a busy class.

The dependent variables were participant heart rate, lesson duration, average response time, and participants' self assessment of their cognitive load, classroom management, and teaching performance. Additionally, we also asked the participants to compare the two interfaces based on their preference. Lesson duration is the time it took the participants to complete the lesson. Average response time(ART) is defined in equation 1, where N is total number of important students, T_l is time subject looked at student i and T_i is time student i became important.

$$ART = \frac{\sum_{i=1}^{N}(T_l - T_i)}{N} \quad (1)$$

Based on prior research, and our understanding of the two interfaces, we have the following hypotheses:

Hypothesis 1 (H1): Participants would take longer to complete lessons using the avatar-located interface compared to the indicator-located interface.

Hypothesis 2 (H2): The response time to notice an important student would be less for the indicator-located interface when compared with the avatar-located interface.

Hypothesis 3 (H3): Participants' increase in heart rate and cognitive load will be lower for the indicator-located interface when compared with the avatar-located interface.

Hypothesis 4 (H4): Participants will prefer the indicator-located interface over the avatar-located interface.

Figure 5: Experimental Setup.

4.1 Participants and Apparatus

The user study had 30 participants in total 1‘ (26 male and 4 female, ages ranging from 19 to 35 years with a mean age of 23). All but two participants were computer science and informatics majors in University of Louisiana at Lafayette. The other two participants were majors of the electrical engineering department at the same university. Out of the 30 participants, 27 had prior VR experience. Only one participant had previous experience of using a VR program that taught engineers about certain machines. Twelve participants had used educational applications and seven participants had prior teaching experience. The experimental setup included (Figure 5) two 27-inch Dell monitor, a HTC Vive Pro Eye with a Vive controller, a polar verity sense wrist-based heart rate monitor, and a Dell

Alienware PC (Core i9-12900KF, RTX 3080 Graphics card, 32 GB RAM) with Microsoft Windows 11. The Vive pro eye headset has a display resolution of 1440 × 1600 pixels per eye, 90 Hz refresh rate and a FOV of 110° as stated by the manufacturer. It offers calibration with five points. The teaching interface was implemented using Unity 2019.4.11f1 version.

4.2 Experiment Design and Procedure

The user experiment design and procedure were first approved by the Institutional Review Board (IRB) of University of Louisiana at Lafayette. The participants were asked to presume the role of a teacher for a basic programming course taught in Java. They were provided a set of slides containing teaching material related to conditional statements in Java. They were asked to teach a class of 30 simulated students using the slides while monitoring the students and be mindful of students if and when the students raised their hands. The participants were asked to look at student avatars when they raised their hands till the hand raised icons disappeared. The system recorded the time it took participants to notice and look at important students. The participants needed to look at a student for at least 3 seconds at a time for the student to be labeled as 'looked at' by the system. This helped in making sure that no student was labeled as 'looked at' due to unintentional eye movement of the participant. The system time, after this, 3 seconds were recorded as the time the participants noticed an important student.The 3 seconds are not included in the average time calculation and was disregarded throughout all the interface versions and number of important student cases. One, three, or five students, in-order, would become important at a time during the study. The time at which a student or a set of students became important was randomly decided by the system and subjects could not move on to the next section of slides if they had an unattended important student in the classroom. This also served as an additional indicator for the participants to make them aware of the important students (if any) in the classroom to avoid ignoring them. During this training phase, the participants were made aware that different colors of the icons meant different attention levels of the students and for this particular study they did not need to do anything regarding their attention level and these colors are just giving them additional information about the classroom.

The experiment began with the participant signing the consent form explaining the study. The participant then put on the heart rate monitor (polar verity sense) on their non-dominant hand. Next, the participant filled the demographic and prior experience questionnaire (see Table 1). The questionnaire was created using Qualtrics XM and the participant used the computer to fill the questionnaire. This was a seated VR experience and thus, the participant was seated at the station with the moderator seated to the side. The participant was seated in front of the monitor at a distance of about 3-4 feet with a Vive controller in their dominant hand. Inside the VR environment, the participants were placed in front of a power-point presentation in a VR classroom. The participants use the same set of slides for both the interface trials. The simulated students were arbitrarily placed in the classroom.

Eye calibration was done for each participant before they started the study. The participants started with a training trial where they were shown a VR classroom and were asked to look around so that they had an idea of how and where student avatars were placed. Additionally, a white sphere showed them where they were looking in the VR classroom. This option was turned off in the actual interface trials. In the training trial, the participants were also shown both the interface options. The participants were told that the students were simulated and would not be actually asking questions. Participants were told that their only job was to look at the important student when they raised their hands and to keep looking till the hand raise icon disappeared. The participants practiced looking at the important student and could see how the hand raise action icon disappeared after they looked at them for 3 seconds straight. Moreover, the participants were asked to go through the teaching slides in the training trial. This prepared them for the teaching task in the actual study. Their heart rate was measured during this training trial to get the baseline average heart rate value of each participant.

Table 1: Demographic Questionnaire. Participants answered these open-ended questions using the text box except question one which was a multiple choice question

Demographic Questionnaire	
Q1	Gender
Q2	Age?
Q3	Occupation?
Q4	Major?
Q5	Do you have any previous experience with VR applications? If yes, to what extent?
Q6	Do you have any previous experience with eye tracked VR applications? If yes, to what extent?
Q7	Do you have any previous experience with educational VR applications? If yes, to what extent?
Q8	Do you have any previous experience with educational applications? If yes, to what extent?
Q9	Do you have any previous experience of teaching? If yes, to what extent?

Table 2: Cognitive Load Questionnaire. Participants answered these questions as 7 point Likert-like items.

Cognitive Load Questionnaire	
Q1	How mentally demanding was it while teaching?
Q2	How physically demanding was the interface?
Q3	How much time pressure did you feel due to the interface or as the student became important?
Q4	How successful did you feel in what you were asked to do?
Q5	How hard did you have to work to accomplish your level of performance?
Q6	How frustrated were you with this interface?

Table 3: Self Evaluation Questionnaire. Participants gave their performance ratings as 7 point Likert-like items.

Self Evaluation Questionnaire	
Q1	I could attend to the needs of individual students.
Q2	I felt the need to stop during lessons to address important students.
Q3	I was able to attend to two events(teaching and addressing important student) simultaneously without being diverted unduly by disruptions.
Q4	I felt I had a firm grasp of the class.
Q5	I feel this interface is adequate for my classroom management techniques.

After the training phase, the participants either tried the avatar-located interface or the indicator-located interface in a counterbalanced manner. During this phase, the participants taught the students using the slides and looked at the students when they became important. By gradually changing the number of important students, we wanted to test the scalability of our interface with more and more important students. We wanted to test if the teacher could use one of these interfaces to monitor students without significantly affecting the teaching performance. Once the participants were done with the lesson, they took a break and answered questions regarding their cognitive load (see Table 2) and self-evaluated their classroom management experience (see Table 3) with the interface they tried. Subsequently, the participants tried the other interface and answered the same questionnaires for it. Lastly, they completed a comparison questionnaire (see Table 4) comparing the two interfaces along

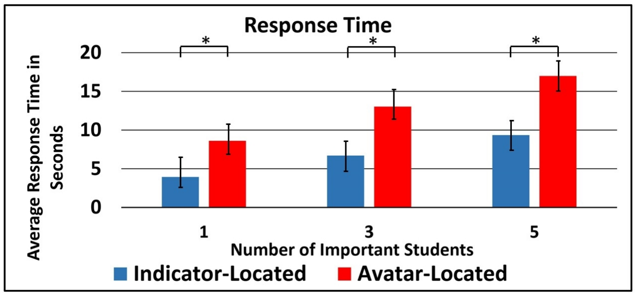

Figure 6: Average response time for each visualization (* = $p < 0.005$)

with stating their reasons for their choice. Lab discussions based on the NASA TLX questionnaire [13] and Naumann et al.'s work [28] helped us decide on the questions of the questionnaires. We only picked questions which were relevant for this study. During the trials, the system logged in the heart rate and lesson duration for each trial.

Table 4: Comparison Questionnaire. Participants chose one of two options (Avatar-Located Indicator or Centrally Arranged Indicator) for each comparison question

Comparison Questionnaire	
Q1	Which version of the interface are you most inclined to use?
Q2	Which version of the interface did you like better?
Q3	Which version of the interface did you think is more effective?

5 Results

For each dependent variable repeated-measures 2-factor ANOVA was used to analyze performance data. A post-hoc analysis was done using pairwise sample t-tests. To analyze the Likert scale data, we used Friedman's test and then a post-hoc analysis was done using Wilcoxon signed rank test. An α value of 0.05 was used for all the statistical measures.

5.1 Response Time

Table 5 contains the results of Repeated measures 2-factor ANOVA analysis on average response time. There were significant differences in response time for the two interfaces and for different numbers of important students. Our post-hoc analysis with pairwise t-test revealed that the indicator-located interface had significantly lower ($t_{29} = 4.309, p < 0.005$) response time compared to the avatar-located interface. When we further compared the interfaces for each number of important students, we found significant differences. Figure 6 shows the average response time for both the interfaces with increased numbers of important students. We found that the indicator-located interface had significantly lower response time for each case of 1 ($t_{29} = 3.959, p < 0.005$), 3 ($t_{29} = 3.993, p < 0.005$) or 5 ($t_{29} = 3.935, p < 0.005$) important students.

Table 5: Repeated measures 2-factor ANOVA results for average response time and increase in heart rate. IT is Interface Type (Avatar-located and Indicator-located) and NS is no of important students (one, three or five). Blue background shows significant values.

Source	Average Response Time	Heart Rate Increase
IT	$F_{1,29} = 20.162, p < 0.005$	$F_{1,29} = 0.080, p = 0.779$
NS	$F_{2,28} = 13.783, p < 0.005$	$F_{2,28} = 0.539, p = 0.589$
IT×NS	$F_{2,28} = 2.472, p = 0.103$	$F_{2,28} = 0.087, p = 0.917$

5.2 Heart Rate Increase

Table 5 contains the results of Repeated measures 2-factor ANOVA analysis for increase in heart rate from the baseline (measured during the training phase). We found that there was no significant difference between the two interfaces and for different number of important students. Although the average increase in heart rate was slightly higher when the subjects used the avatar-located interface (see Figure 7), it was not significantly different from the case with the indicator-located interface.

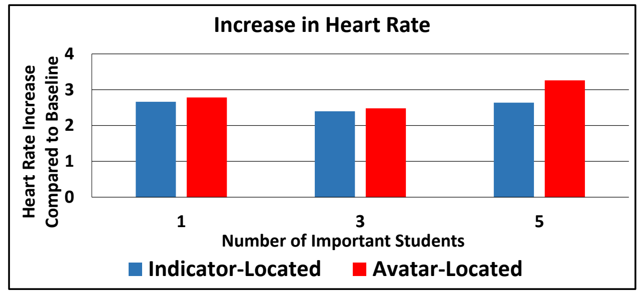

Figure 7: Increase in heart rate from the baseline for each visualization. The baseline heart rate was measured during the training scene when participants were not performing any assigned tasks.

5.3 Lesson Duration

A Paired samples t-test was conducted to determine the difference between the two interfaces. The results indicate that there was a significant difference ($t_{29} = 5.742, p < 0.005$) between the two interfaces in terms of lesson duration. Figure 8 shows the average lesson duration in seconds for both the interfaces. The lesson duration was significantly lower with the indicator-located interface (M=315.5519; SD=103.31042) compared to the avatar-located interface (M=351.9052; SD=103.83437). Only four participants took longer to finish the lesson with the indicator-located interface. However, all four of these participants tried the indicator-located interface first. So we believe they might have become more familiar with the lesson content by the time they tried the avatar-located interface and that contributed to them taking less time to finish the lesson.

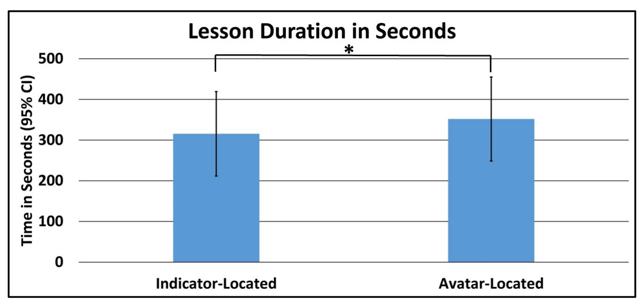

Figure 8: Average lesson duration for the two visualization techniques (* = $p < 0.005$).

5.4 Cognitive Load Questionnaire

The likert scale data collected as part of this questionnaire was analyzed using Friedman's test. The results of the Friedman's test are summarized in Table 6. There were significant differences for all the questions. The mean ratings for the questions (Q1 to Q6) are summarized in Figure 9. Table 7 summarizes the results of pairwise Wilcoxon Signed-Rank tests for each question.

Question 1 (Mental Demand): The participants felt significantly more mental demand with the avatar-located interface over the indicator-located interface. In general, the mental demand increased with the increase in number of important students. The participants found the avatar-located interface significantly more mentally demanding compared to the indicator-locator interface

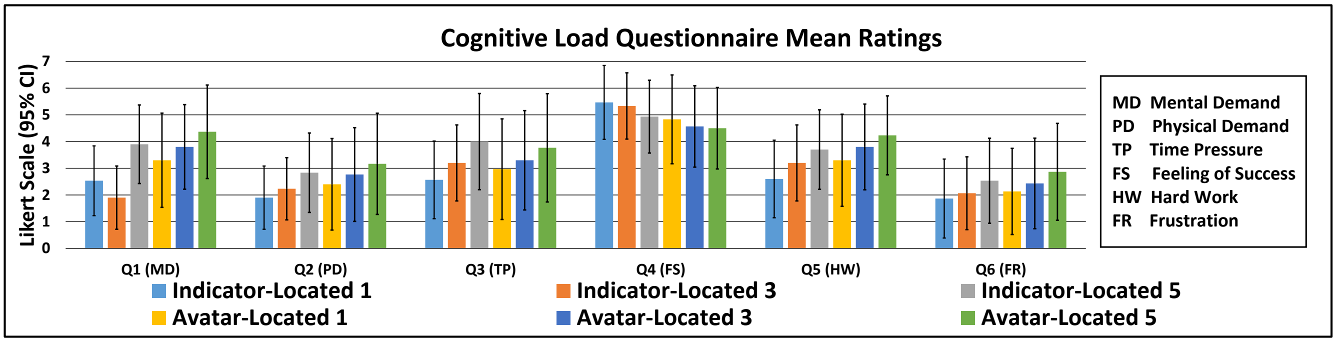

Figure 9: Mean Ratings for the cognitive load questionnaire for different numbers of students.

Table 6: Results of Friedman's test for the cognitive load questionnaire Likert scale data. Blue background shows significant values.

Question	Friedman's test
Q1 (MD)	$\chi^2(5) = 55.508, p < 0.0005$
Q2 (PD)	$\chi^2(5) = 33.356, p < 0.0005$
Q3 (TP)	$\chi^2(5) = 25.137, p < 0.0005$
Q4 (FS)	$\chi^2(5) = 17.109, p < 0.005$
Q5 (HW)	$\chi^2(5) = 43.829, p < 0.0005$
Q6 (FR)	$\chi^2(5) = 21.885, p < 0.005$

when three students were important but not for one or five important students. A significant increase in mental demand can be noticed for the avatar-located interface. For the indicator-located interface, surprisingly, the mental demand was significantly lower for 3 important students compared to when only one important student was present.

Question 2 (Physical Demand): Participants reported less physical demand while using the indicator-located interface compared to the avatar-located interface, and the demand increased with the increase of the numbers of important students. There were significant increases in physical demand for increasing numbers of important students for both the avatar-located icons and the indicator-located icons.

Question 3 (Time Pressure): Participants reported they felt more time pressure with the increased number of important students for both the interfaces. However, interestingly for 5 students, participants only felt more time pressure with the indicator-located interface when compared to the avatar-located interface. For the avatar-located interface, there was no significant difference in time pressure between one and three important students.

Question 4 (Feeling of Success): There was no significant difference for any of the comparison pairs based on participants' feeling of success.

Question 5 (Hard Work): Participants felt they needed to work harder with the avatar-located interface. Participants felt that they worked significantly harder with the increase of number of important students with both the interfaces. When we did pairwise tests, we found that the participants felt that they did significantly more hard work with the avatar-located interface when only one important student was present.

Question 6 (Frustration): Participants were significantly more frustrated with the avatar-located interface compared to the indicator-located interface. On average, the feeling of frustration got worse as we increased the number of important students. For the avatar-located interface, the feeling of frustration got increased significantly as we increased the number of important students. However, for the indicator-located interface, the feeling of frustration was not significantly different when we increased the number of important students from 1 to 3. The feeling of frustration increased significantly when we increased the number of important students from 3 to 5 in this case of indicator-located interface.

5.5 Self Evaluation Questionnaire

The likert scale data collected as part of this self-evaluation questionnaire was analyzed using Friedman's test. The results of Friedman's test are summarized in Table 8. There were significant differences for all the questions. The mean ratings for the questions (Q1 to Q5) are summarized in Figure 10. Table 9 summarizes the results of pairwise Wilcoxon Signed-Rank tests for each question.

Question 1 (Attention to student needs): Significantly more participants felt that they could attend to individual needs more with the avatar-located interface and their performance regarding this question decreased with increased number of important students for both the interfaces.

Question 2 (Need to stop): Significantly more participants felt less need to stop during their lessons while using the indicator-located interface and this need increased with the increase in the number of important students for both the interfaces. Other than the pair with one and three important students, there were significant differences for all the other pairs for each of the interfaces.

Question 3 (Multi-Tasking ability): There was no significant difference in participants' ability to attend dual events (teaching and monitoring students) as we increased the number of important students. For the indicator-located interface, the participants indicated that multitasking ability was reduced as we increased the number of important students from 1 to 5.

Question 4 (Firm grasp of the class): The results were not significantly different for the avatar-located interface as we increased the number of important students. For the indicator-located interface, the ratings for 1 and 3 students were similar. However, the participants felt significantly less grasp of the class as we increased the number of students from 1 to 5 or 3 to 5.

Question 5 (Adequacy of the interface): When we compared the two interfaces for each number of important students, the participants felt that the indicator-located interface was significantly better compared to the avatar-located interface when three important students were present. For each interface, it became less and less adequate as we increased the number of important students.

5.6 Comparison Questionnaire

Figure 11 shows the participants responses when asked to choose between the two interfaces. In general, 22 participants out of 30 were inclined to use the indicator-located interface. Few participants thought the indicators were distracting from the lesson and thought the avatar-located interface felt more natural to use and they could simply see the hand indicator on top of the students raising their hands. On the other hand, participants found the indicator-located interface easier to see who needed attention at the moment and felt it helped to find the students who needed help faster. While a participant agreed indicator-located interface could be distracting, they were more inclined to use indicator-located interface because monitoring students was easier and for awareness of all class actions

Table 7: Summary of results of pairwise Wilcoxon Signed-Rank test for Cognitive load questions 1 to 6 where AL is avatar-located interface and IL is indicator-located interface. The number at the end represents the number of important students. Blue background shows significant values.

Pairs	Q1 (MD)	Q2 (PD)	Q3 (TP)	Q4 (FS)	Q5 (HW)	Q6 (FR)
AL1 vs. AL3	Z = -2773, p =0.006	Z = -2.840, p =0.005	Z = -2.161, p =0.031	Z = -1.565, p =0.118	Z = -3.274, p =0.001	Z = -2.714, p =0.007
AL1 vs. AL5	Z = -3.297, p =0.001	Z = -3.660,p<0.0005	Z = -2.865, p =0.004	Z = -1687, p =0.092	Z = -3.703,p<0.0005	Z = -3.256, p =0.001
AL3 vs. AL5	Z = -3.022, p =0.003	Z = -2.972, p =0.003	Z = -2.913, p =0.004	Z = -632, p =0.527	Z = -3.127, p =0.002	Z = -2.970, p =0.003
IL1 vs. IL3	Z = -2.282, p =0.022	Z = -2.673, p =0.008	Z = -2.700, p =0.007	Z = -0.882, p =0.378	Z = -3.116, p =0.002	Z = -1.137, p =0.256
IL1 vs. IL5	Z = -4.066, p<0.0005	Z = -3.834,p<0.0005	Z = -3.644,p<0.0005	Z = -2.491, p =0.013	Z = -2.249, p =0.024	Z = -2.520, p =0.012
IL3 vs. IL5	Z = -4.063,p <0.0005	Z = -3.082, p =0.002	Z = -3.275, p =0.001	Z = -2.101, p =0.036	Z = -3.116, p =0.002	Z = -2.913, p =0.004
AL1 vs. IL1	Z = -2.203, p =0.028	Z = -1.490, p =0.136	Z = -1.226, p =0.220	Z = -1.878, p =0.060	Z = -2.249, p =0.024	Z = -0.866, p =0.386
AL3 vs. IL3	Z = -3.644,p<0.0005	Z = -1.511, p =0.131	Z = -0.203, p =0.839	Z = -2.575, p =0.010	Z = -1.770, p =0.077	Z = -0.892, p =0.373
AL5 vs. IL5	Z = -0.964, p =0.335	Z = -1.020, p =0.308	Z = -0.618, p =0.538	Z = -1.252, p =0.211	Z = -1.441, p =0.150	Z = -0.717, p =0.473

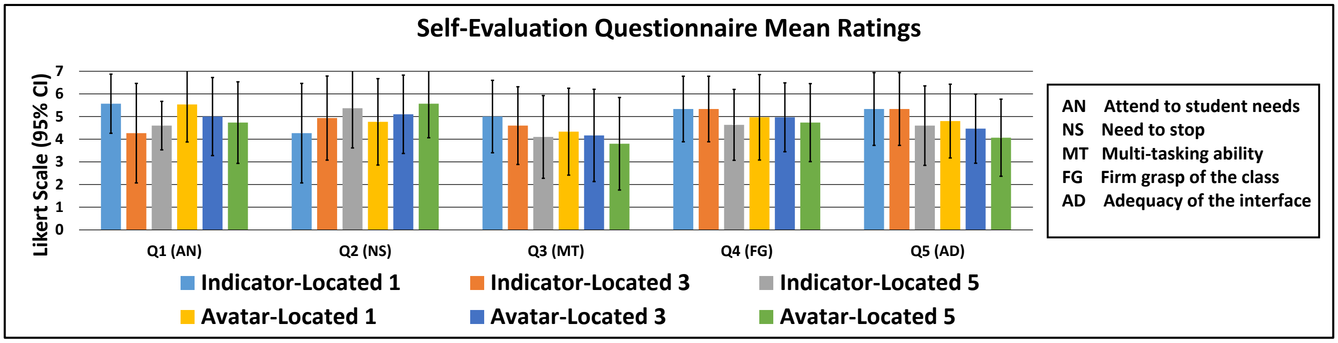

Figure 10: Mean Ratings for the self-evaluation questionnaire for different numbers of students.

Table 8: Results of Friedman's test for the self-evaluation questionnaire Likert scale data. Blue background shows significant values.

Question	Friedman's test
Q1 (AN)	$\chi^2(5) = 18.076, p < 0.005$
Q2 (NS)	$\chi^2(5) = 23.046, p < 0.0005$
Q3 (MT)	$\chi^2(5) = 15.225, p < 0.005$
Q4 (FG)	$\chi^2(5) = 11.700, p < 0.05$
Q5 (AD)	$\chi^2(5) = 25.653, p < 0.0005$

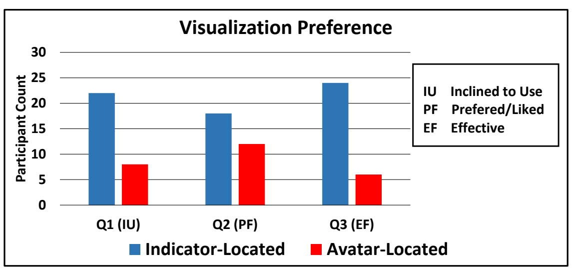

Figure 11: Visualization preference.

as a whole. Participants liked that the indicator-located interface allowed for the teacher to have a faster response time to attend to a student's needs and to have a better understanding of how the students are behaving at a given time.

Out of 30, 18 participants liked the indicator-located interface more as it showed exactly where the student was instead of having to search for a student who may have a question. The participants appreciated that the indicator-located interface provided the participants the exact count of students that raised their hands and that they did not have to look around to know if a student was raising their hand. One user felt stressed to keep looking back to check if students had questions in the avatar-located interface. Participants felt it was easier to plan lessons and when to address important students with the indicator-located interface. However, a few participants found the overall experience easier with the avatar-located interface. The avatar-located interface was less distracting as this interface allowed them to teach more naturally, without having any icons right in front in their field of view during the teaching session. One participant felt that without indicator-located icons, they could teach at their own pace while being able to look at the class instead of just relying on indicators to know when to look at them.

Out of 30, 24 participants thought indicator-located interface was more effective. It made it easier for the teacher to keep track of the students without having to look away from the board and stop teaching. The indicator-located interface allowed the teachers to multitask by minimizing the visual needs of the teachers for student awareness. Participants liked the quicker response time of the indicator-located interface to address students' needs. They felt when students' questions were unaddressed for long periods, they might lower their hand and withdraw their question. For a few participants, the avatar-located interface felt more organic to manage a classroom. A participant felt they could keep their train of thought and speech while looking at the slides.

As described above, some participants did seem to like the avatar-located interface better even if they did not think this interface was more effective (6 participants) or they were more inclined to use it (4 participants). When asked if the participants had any general comments about the interfaces, participants mentioned teaching in a virtual environment is more difficult, but there is also less pressure. One participant acknowledged how the sporadic and non-uniform standing of the students around the room can be a little distracting. While participants liked indicator-located interface more than the avatar-located interface, they suggested the interface given should be simplified and not take up as much screen space as it did. Having an icon for every single student is a bit much and based on participants' feedback we think that showing indicators for only important students would be a better idea.

6 Discussion

Based on our results, we were able accept our hypothesis 1 (H1) since participants took longer to teach with the avatar-located interface. Since the average response time was lower with the indicator-located interface, we were able to accept our hypothesis 2 (H2). We had to reject our hypothesis 3 (H3) since we did not find any significant difference in heart rate between the two interfaces. Based on our results from the comparison questionnaire, we were able to accept

Table 9: Summary of results of pairwise Wilcoxon Signed-Rank test for the Self Evaluation questions from 1 to 5 where AL is avatar-located interface and IL is indicator-located interface. The number at the end represents the number of important students. Blue background shows significant values.

Pairs	Q1 (AN)	Q2 (NS)	Q3 (MT)	Q4 (FG)	Q5 (AD)
AL1 vs. AL3	Z = -2.719,p = 0.007	Z = -2.153,p = 0.031	Z = -1.343,p = 0.179	Z = -0.696,p = 0.484	Z = -2.178,p = 0.029
AL1 vs. AL5	Z = -3.019,p<0.005	Z = -3.223,p<0.005	Z = -2.088,p = 0.037	Z = -0.960,p = 0.337	Z = -2.429,p = 0.015
AL3 vs. AL5	Z = -1.721,p = 0.085	Z = -3.276p<0.005	Z = -2.484,p = 0.013	Z = -1.393,p = 0.163	Z = -364,p = 0.018
IL1 vs. IL3	Z = -2.298,p = 0.007	Z = -2.153,p = 0.031	Z = -2.144,p = 0.032	Z = 000,p = 1.000	Z = 0.000,p = 1.000
IL1 vs. IL5	Z = -3.372p<0.005	Z = -3.293p<0.005	Z = -3.056p<0.005	Z = -2.914p<0.005	Z = -3.115,p<0.005
IL3 vs. IL5	Z = -0.694,p = 0.488	Z = -2.667,p = 0.008	Z = -1.966,p = 0.049	Z = -2.914,p<0.005	Z = -3.115,p<0.005
AL1 vs. IL1	Z = -0.021,p = 0.983	Z = -1.568,p = 0.117	Z = -1.646,p = 0.100	Z = -0.999,p = 0.318	Z = -1.783,p = 0.75
AL3 vs. IL3	Z = -1.661,p = 0.097	Z = -0.233,p = 0.816	Z = -0.836,p = 0.403	Z = -1.327,p = 0.185	Z = -2.416,p = 0.016
AL5 vs. IL5	Z = -0.736,p = 0.462	Z = -0.219,p = 0.827	Z = -0.434,p = 0.664	Z = -0.514,p = 0.607	Z = -1.392,p = 0.164

our hypothesis 4 since more people preferred the indicator-located interface.

We found that based on lesson duration and average response time participants performed better with Indicator-located interface over avatar-located interface. Although the increase of students resulted in an increase of average response time, we believe the increase is within reasonable limits. Also, the increase in response time during indicator-located was decidedly less compared to the increase in response time during avatar-located interface which indicates the indicator-interface can handle increased student count better.

There was no difference in cognitive load for both the interfaces which is supported by the analysis of cognitive load questions and also by no significant difference in heart rate. While existing research showed a connection between heart rate and cognitive load, our findings did not show any significant difference in heart rate for the increasing number of students although we could find some significant difference in cognitive load based on the participant response of the questionnaire. This opens up the question that heart rate alone may never be a good indicator of cognitive load.

A lot of the participants were teaching for the first time and they seemed to have a varied range of reactions due to this fact. Some treated it as a new adventure to teach while some participants felt uncomfortable or conscious to teach. The participants did not clearly know when they would consider themselves successful while teaching and we believe that contributed to us not getting significant difference data for cognitive load question no 4 of participants' feeling of success. Although the interface with indicator-located was more liked, participants reported they got distracted by the indicator and we assume that is the reason there is a difference between the number of participants who thought the indicator-located interface was more effective and participants who liked the indicator-located icons more.

Even with more availability and cheaper VR options, not as many teachers are inclined to use VR teaching interfaces. It could be due to age and their disinterest in learning new technology or just due to lack of proper exposure. If teachers are given adequate training in VR to the point where they can confidently be familiar with VR and also shown how VR can eradicate the boundary of time, space and safety to teach lessons, we believe a lot more teachers will be interested in using VR to teach their class. Having a teaching interface that feels intuitive to the teacher, can go a long way in removing the gap that seems to be present between teachers being interested in teaching in VR and students being interested in learning in VR.

There are a few factors that could have affected our results. The students in the experiment were simulated students and this could have some effect on our results. However, it still gave us a chance to evaluate the two interfaces on a teaching task. In the future, we would like to test these interfaces with real students. Additionally, the teaching task was for a short duration. In a real VR classroom, the lessons would be much longer. We need further testing to evaluate the two interfaces with longer duration lessons. One limitation of our user study is that most of our participants did not have real experience as a teacher. However, we kept the java lesson plan to beginner level so that it was easy even for the first time teachers to teach. Also during the training phase, they could take as much time as they wanted to prepare their lesson and ask questions in case they had any. To avoid a bias because of different lesson plans we fixed the lesson all participants had to teach but that limited the generalizations of the teaching interface across all fields of education. Furthermore, only 13% of our participants were females and this could have created a potential gender bias [32]. For our future studies, we would like to include equal number of male and female participants to avoid this bias.

Maintaining student privacy is an important concern when sharing physiological sensor data of students with the teacher. Educational VR applications are an emerging field but educators and learners need to be conscious of the privacy concerns while using VR. In VR, it is possible to track user movements, user interaction between themselves and with objects and even biometric data like facial expression or eye tracking data. Safe and secure VR learning environments should be designed with keeping privacy in mind and also proper security measures are needed to be taken. In our study, eye-tracking and heart rate data was collected from the participants who gave permission to use their data within a standard informed consent model and the recorded data was anonymized.

7 Conclusion and Future Work

In this paper, we present a comparative study of two such student monitoring interfaces: avatar-located and indicator-located. In case of the avatar-located interface, the student activity related information is shown using icons near the student avatar (representing a student in the VR environment). While in case of indicator-located interface, a set of centrally-arranged emoticon-like visual indicators are present in addition to the student avatar, and the student activity related information is shown near the student emoticon. We did a detailed user experiment to compare the two interfaces in terms of teaching management, student monitoring capability, cognitive load, and user preference. Our results show that the indicator-located interface is better since it makes it easy to monitor students as indicated by lower response time and lesson duration. Additionally, this interface does not significantly increase the cognitive load of the participant.

In the future, it would be interesting to see if using other EEG sensor data can help us understand teachers' cognitive load better. Additionally, seeing how using the interfaces affect the way a teacher addresses and responds to a student when they are asking a question could be interesting. Furthermore, it will be interesting to see how the interface performs with real students.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant No. 1815976 and by the Louisiana Board of Regents under contract No. LEQSF(2022-25)-RD-A-24.

References

[1] Hubs by mozilla. https://hubs.mozilla.com/.

[2] Virbela. https://www.virbela.com/.

[3] H. S. Alavi and P. Dillenbourg. An ambient awareness tool for supporting supervised collaborative problem solving. *IEEE Transactions on Learning Technologies*, 5(3):264–274, jul 2012. doi: 10.1109/tlt.2012.7

[4] W. Alhalabi. Virtual reality systems enhance students' achievements in engineering education. *Behaviour & Information Technology*, 35(11):919–925, jul 2016. doi: 10.1080/0144929x.2016.1212931

[5] P. An, K. Holstein, B. d'Anjou, B. Eggen, and S. Bakker. The TA Framework: Designing Real-time Teaching Augmentation for K-12 Classrooms. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Apr. 2020. doi: 10.1145/3313831.3376277

[6] M. Belani, H. V. Singh, A. Parnami, and P. Singh. Investigating spatial representation of learning content in virtual reality learning environments. In *2023 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*, pp. 33–43, 2023. doi: 10.1109/VR55154.2023.00019

[7] D. A. Berque and J. T. Newman. GlassClass: Exploring the Design, Implementation, and Acceptance of Google Glass in the Classroom. In *Lecture Notes in Computer Science*, pp. 243–250. Springer International Publishing, 2015. doi: 10.1007/978-3-319-21067-4_25

[8] E. Bozkir, P. Stark, H. Gao, L. Hasenbein, J.-U. Hahn, E. Kasneci, and R. Gollner. Exploiting Object-of-Interest Information to Understand Attention in VR Classrooms. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*. IEEE, Mar. 2021. doi: 10.1109/vr50410.2021.00085

[9] D. M. Broussard, Y. Rahman, A. K. Kulshreshth, and C. W. Borst. An Interface for Enhanced Teacher Awareness of Student Actions and Attention in a VR Classroom. In *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 284–290, 2021. doi: 10.1109/VRW52623.2021.00058

[10] D. M. Broussard, Y. Rahman, A. K. Kulshreshth, and C. W. Borst. An Interface for Enhancing Teacher Awareness of Student Actions and Attention in a VR Classroom. In *2021 IEEE 7th Workshop on Everyday Virtual Reality (WEVR)*, 2021.

[11] D. M. Broussard, Y. Rahman, A. K. Kulshreshth, and C. W. Borst. Visual Indicators for Monitoring Students in a VR class. In *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 502–503, 2021. doi: 10.1109/VRW52623.2021.00133

[12] B. Chavez and S. Bayona. Virtual reality in the learning process. In *Advances in Intelligent Systems and Computing*, pp. 1345–1356. Springer International Publishing, 2018. doi: 10.1007/978-3-319-77712-2_129

[13] K. N. Cranford, J. M. Tiettmeyer, B. C. Chuprinko, S. Jordan, and N. P. Grove. Measuring Load on Working Memory: The Use of Heart Rate as a Means of Measuring Chemistry Students' Cognitive Load. *Journal of Chemical Education*, 91(5):641–647, Apr. 2014. Publisher: American Chemical Society (ACS). doi: 10.1021/ed400576n

[14] S. Digiesi, V. M. Manghisi, F. Facchini, E. M. Klose, M. M. Foglia, and C. Mummolo. Heart rate variability based assessment of cognitive workload in smart operators. *Management and Production Engineering Review*, 11(3):56–64, 2020. Publisher: de Gruyter. doi: 10.24425/mper.2020.134932

[15] K. Holstein, G. Hong, M. Tegene, B. M. McLaren, and V. Aleven. The classroom as a dashboard. In *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*. ACM, mar 2018. doi: 10.1145/3170358.3170377

[16] K. Holstein, B. M. McLaren, and V. Aleven. Student learning benefits of a mixed-reality teacher awareness tool in AI-enhanced classrooms. In *Lecture Notes in Computer Science*, pp. 154–168. Springer International Publishing, 2018. doi: 10.1007/978-3-319-93843-1_12

[17] K. Holstein, B. M. McLaren, and V. Aleven. Co-Designing a Real-Time Classroom Orchestration Tool to Support Teacher–AI Complementarity. *Journal of Learning Analytics*, 6(2), July 2019. Publisher: Society for Learning Analytics Research. doi: 10.18608/jla.2019.62.3

[18] E. Hu-Au and J. Lee. Virtual reality in education: a tool for learning in the experience age. *International Journal of Innovation in Education*, 4, Apr. 2018. doi: 10.1504/IJIIE.2017.10012691

[19] L. Jensen and F. Konradsen. A review of the use of virtual reality head-mounted displays in education and training. *Education and Information Technologies*, 23(4):1515–1529, nov 2017. doi: 10.1007/s10639-017-9676-0

[20] D. Kalkofen, S. Mori, T. Ladinig, L. Daling, A. Abdelrazeq, M. Ebner, M. Ortega, S. Feiel, S. Gabl, T. Shepel, J. Tibbett, T. H. Laine, M. Hitch, C. Drebenstedt, and P. Moser. Tools for teaching mining students in virtual reality based on 360°video experiences. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, mar 2020. doi: 10.1109/vrw50115.2020.00096

[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.

[22] J. Leppink, F. Paas, C. P. M. V. d. Vleuten, T. V. Gog, and J. J. G. V. Merriënboer. Development of an instrument for measuring different types of cognitive load. *Behavior Research Methods*, 45(4):1058–1072, Apr. 2013. Publisher: Springer Science and Business Media LLC. doi: 10.3758/s13428-013-0334-1

[23] H. Luo, G. Li, Q. Feng, Y. Yang, and M. Zuo. Virtual reality in k-12 and higher education: A systematic review of the literature from 2000 to 2019. *Journal of Computer Assisted Learning*, mar 2021. doi: 10.1111/jcal.12538

[24] S. Martin. Measuring cognitive load and cognition: metrics for technology-enhanced learning. *Educational Research and Evaluation*, 20(7-8):592–621, Nov. 2014. Publisher: Informa UK Limited. doi: 10.1080/13803611.2014.997140

[25] M. Mavrikis, S. Gutierrez-Santos, and A. Poulovassilis. Design and evaluation of teacher assistance tools for exploratory learning environments. In *Proceedings of the Sixth International Conference on Learning Analytics &amp$\mathsemicolon$ Knowledge - LAK '16*. ACM Press, 2016. doi: 10.1145/2883851.2883909

[26] I. Molenaar and C. A. N. K.-v. Campen. How Teachers Make Dashboard Information Actionable. *IEEE Transactions on Learning Technologies*, 12(3):347–355, July 2019. Publisher: Institute of Electrical and Electronics Engineers (IEEE). doi: 10.1109/tlt.2018.2851585

[27] A. Mottus, Kinshuk, S. Graf, and N.-S. Chen. Use of Dashboards and Visualization Techniques to Support Teacher Decision Making. In *Ubiquitous Learning Environments and Technologies*, pp. 181–199. Springer Berlin Heidelberg, Sept. 2014. doi: 10.1007/978-3-662-44659-1_10

[28] A. Naumann and J. Hurtienne. Benchmarks for intuitive interaction with mobile devices. In *Proceedings of the 12th international conference on Human computer interaction with mobile devices and services*. ACM, sep 2010. doi: 10.1145/1851600.1851685

[29] K. Ouwehand, A. van der Kroef, J. Wong, and F. Paas. Measuring cognitive load: Are there more valid alternatives to likert rating scales? *Frontiers in Education*, 6, sep 2021. doi: 10.3389/feduc.2021.702616

[30] D. Parmar and T. Bickmore. Making it personal : Addressing individual audience members in oral presentations using augmented reality. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 4(2):1–22, jun 2020. doi: 10.1145/3397336

[31] T. Patterson and I. Han. Learning to Teach with Virtual Reality: Lessons from One Elementary Teacher. *TechTrends*, 63(4):463–469, July 2019. doi: 10.1007/s11528-019-00401-6

[32] T. C. Peck, L. E. Sockol, and S. M. Hancock. Mind the gap: The underrepresentation of female participants and authors in virtual reality research. *IEEE transactions on visualization and computer graphics*, 26(5):1945–1954, 2020. Publisher: IEEE.

[33] J. Radianti, T. A. Majchrzak, J. Fromm, and I. Wohlgenannt. A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers & Education*, 147:103778, 2020. Publisher: Elsevier.

[34] Y. Rahman, A. K. Kulshreshth, and C. W. Borst. Exploring the usefulness of visual indicators for monitoring students in a vr-based teaching interface. In *2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 699–700, 2023. doi: 10.1109/VRW58643.2023.00192

[35] B. A. Schwendimann, M. J. Rodriguez-Triana, A. Vozniuk, L. P. Prieto, M. S. Boroujeni, A. Holzer, D. Gillet, and P. Dillenbourg. Perceiving

Learning at a Glance: A Systematic Literature Review of Learning Dashboard Research. *IEEE Transactions on Learning Technologies*, 10(1):30–41, Jan. 2017. Publisher: Institute of Electrical and Electronics Engineers (IEEE). doi: 10.1109/tlt.2016.2599522

[36] 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, oct 2019. doi: 10.1145/3357251.3357590

[37] A. Skulmowski and G. D. Rey. Measuring Cognitive Load in Embodied Learning Settings. *Frontiers in Psychology*, 8, Aug. 2017. Publisher: Frontiers Media SA. doi: 10.3389/fpsyg.2017.01191

[38] S. Solhjoo, M. C. Haigney, E. McBee, J. J. G. v. Merrienboer, L. Schuwirth, A. R. Artino, A. Battista, T. A. Ratcliffe, H. D. Lee, and S. J. Durning. Heart Rate and Heart Rate Variability Correlate with Clinical Reasoning Performance and Self-Reported Measures of Cognitive Load. *Scientific Reports*, 9(1), Oct. 2019. Publisher: Springer Science and Business Media LLC. doi: 10.1038/s41598-019-50280-3

[39] N. Urrestilla and D. St-Onge. Measuring Cognitive Load: Heart-rate Variability and Pupillometry Assessment. In *Companion Publication of the 2020 International Conference on Multimodal Interaction*. ACM, Oct. 2020. doi: 10.1145/3395035.3425203

[40] K. VanLehn, H. Burkhardt, S. Cheema, S. Kang, D. Pead, A. Schoenfeld, and J. Wetzel. Can an orchestration system increase collaborative, productive struggle in teaching-by-eliciting classrooms? *Interactive Learning Environments*, 29(6):987–1005, May 2019. Publisher: Informa UK Limited. doi: 10.1080/10494820.2019.1616567

[41] K. Verbert, S. Govaerts, E. Duval, J. L. Santos, F. V. Assche, G. Parra, and J. Klerkx. Learning dashboards: an overview and future research opportunities. *Personal and Ubiquitous Computing*, Nov. 2013. Publisher: Springer Science and Business Media LLC. doi: 10.1007/s00779-013-0751-2

[42] I. Verhulst, A. Woods, L. Whittaker, J. Bennett, and P. Dalton. Do VR and AR versions of an immersive cultural experience engender different user experiences? *Computers in Human Behavior*, 125:106951, dec 2021. doi: 10.1016/j.chb.2021.106951

[43] L. Vermette, J. McGrenere, C. Birge, A. Kelly, and P. K. Chilana. Freedom to Personalize My Digital Classroom. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, May 2019. doi: 10.1145/3290605.3300548

[44] P. Wang, P. Wu, J. Wang, H.-L. Chi, and X. Wang. A critical review of the use of virtual reality in construction engineering education and training. *International Journal of Environmental Research and Public Health*, 15(6):1204, jun 2018. doi: 10.3390/ijerph15061204

[45] B. YILDIRIM, E. S. TOPALCENGİZ, G. ARIKAN, and S. TİMUR. Using virtual reality in the classroom: Reflections of STEM teachers on the use of teaching and learning tools. *Journal of Education in Science, Environment and Health*, jul 2020. doi: 10.21891/jeseh.711779