Internal Distraction Detection Utilizing EEG Data in an Educational VR Environment

Sarker M. Asish
University of Louisiana at Lafayette
Lafayette, Louisiana, USA
asish.sust@gmail.com

Arun K Kulshreshth
University of Louisiana at Lafayette
Lafayette, Louisiana, USA
arunkul@louisiana.edu

Christoph W Borst
University of Louisiana at Lafayette
Lafayette, Louisiana, USA
cwborst@gmail.com

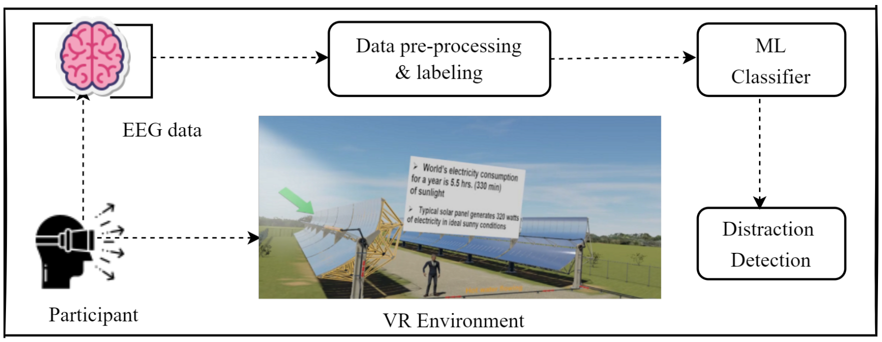


Figure 1: Overview of our approaches for detecting internal distraction.

ABSTRACT

Virtual reality (VR) makes learning more interesting for students and could help them remember what they have learned better than traditional methods. However, a student could get distracted in a VR environment because of stress, wandering thoughts, unwanted noise, outside sounds, etc. Distractions could be classified as either external (due to the environment) or internal (due to internal thoughts). To identify external distractions, previous researchers have used eye-gaze data. Eye-gaze data cannot, however, detect internal distractions because a user may be looking at the educational material in VR while also thinking about something else. We explored the usage of electroencephalogram (EEG) data to detect internal distractions. We designed an educational VR environment and trained three machine learning models: Random Forest (RF), Support Vector Machine (SVM), and k-nearest-neighbors (kNN), to detect internal distractions of students. For data labeling, we considered two window lengths (20 and 30 seconds) starting at 5 seconds after the distraction task started. We did cross-subject and cross-session tests, and our results show that kNN provides a better accuracy (64%) compared to RF and SVM. We also found that the shorter window length of 20 seconds provided a slightly better accuracy then the 30 second window. Our results are not far from such random guessing. Therefore, our contribution lies more in the fostering of ideas for future work that must employ more advanced and sophisticated techniques.

CCS CONCEPTS

• **Human centered computing** → **Machine Learning**; • **Human computer interaction (HCI)** → *Virtual Reality*; • **Education** → Internal Distraction.

KEYWORDS

Human-computer interaction (HCI), Virtual reality, Machine learning, EEG, Education

ACM Reference Format:
Sarker M. Asish, Arun K Kulshreshth, and Christoph W Borst. 2023. Internal Distraction Detection Utilizing EEG Data in an Educational VR Environment. In *ACM Symposium on Applied Perception 2023 (SAP '23), August 05–06, 2023, Los Angeles, CA, USA.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3605495.3605790

1 INTRODUCTION

Virtual reality (VR) technology and virtual environments, incorporating eye tracking with head-mounted displays (HMD) such as the HTC Vive Pro Eye, Pico Neo 2 Eye, or FOVE, have gained considerable attention in recent years. These VR headsets can be utilized for various applications, including but not limited to education [Pirker et al. 2020], training [Mikropoulos and Natsis 2011; Sinnott et al. 2019], collaboration [Cavallo et al. 2019], and business [Kim et al. 2015] (e.g., analyzing shopping trends).

Virtual reality immersion can create experiences that are vividly recalled and may enhance learning. It has also been demonstrated to enhance instructors' instructional abilities [Lamb and Etopio 2020]. However, virtual reality-based education has some challenges. 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 a VR environment, teachers lack direct visibility of their students, leading to a reduced awareness of their activities. Moreover, students may face distractions in VR due to stress, wandering thoughts, external alerts, unwanted noise, and other factors, which could have a detrimental effect on their learning. Therefore, to assist teachers in conducting effective VR-based classes, an automated system for detecting distractions is necessary.

According to real-world classroom distractions, there are two types of distractions: external (due to the environmental factors) or internal (due to internal thoughts in mind) [Tesch et al. 2011]. To identify external distractions like noise or sounds played in the VR environment, eye tracking data has previously been employed [Asish et al. 2021a]. Students may become distracted internally due to anxiety, daydreaming, or other mental activities. Unfortunately, eye-gaze data alone is incapable of gauging mental distractions [Hutt et al. 2019]. It is quite possible that the user may be looking at the educational content and may be thinking about something else. In this paper, we explore the utility of EEG data to detect internal distractions using machine learning approaches. We designed an educational VR environment (see Figure 2) and collected EEG data from 21 participants. Three machine learning models (RF, SVM, and kNN) were trained on the data set to determine whether or not there was internal distraction. Specifically, our work seeks to find answers to the following research questions:

RQ1: To what extent can we detect internal distractions in VR using EEG data?

RQ2: How long do the internal distractions stay in mind after a distraction task?

2 RELATED WORK

2.1 Virtual Reality and Education

When online learning, physical limitations, and situational constraints are considered, a VR headset with an embedded eye tracker is an excellent aid for the teaching and learning process. In the last decade, VR has been studied in many educational contexts, such as safety training [Buttussi and Chittaro 2017], medicine [Gallagher and Cates 2004], and training public security personnel [Bertram et al. 2015]. VR has been used for visualizing and interacting with abstract learning content (e.g., molecular structures [Won et al. 2019]) as well as simulation applications that would be hazardous to practice in real life (e.g., hazardous situation) [Mikropoulos and Natsis 2011].

Recent research in the fields of psychology and human-computer interaction shows that text and audio-based learning can be effective, depending on the task. According to Modality Principle, on-screen speech is superior to on-screen text for learning [Butcher 2014] in terms of complex graphic representations that include dual-channel processing in working memory. Specifically for information retention, [Baceviciute et al. 2020] investigated that reading text from a virtual book is better than listening for learning. Nevertheless, he observed no appreciable differences for knowledge transfer. [Han et al. 2022] proposed some intervention strategies to improve students' attention and their findings suggest that instructions from real world teachers can be transferred to virtual classroom. In some cases, VR leads to a higher sense of presence and keeps users engaged with educational content [Makransky et al. 2019; Meyer et al. 2019; Rucinski et al. 2018]. However, text-based presentation could lead to higher cognitive load and less learning in VR [Makransky et al. 2019]. Teachers noticed benefits in integrating immersive technologies because students were more engaged and immersed in the area of interest, according to research on VR applications in education and training [Martin et al. 2022].

2.2 Distractions in VR

There are numerous activities that could potentially divert students in an educational virtual reality environment. According to psychological research, many students use their cellphones during class to browse the internet or purchase online [Mendoza et al. 2018]. Students may also use a cellphone for social media or other non-academic activities while learning in the classroom, likely reducing knowledge retention. Research suggests that due to multitasking, attention can be diminished by shifting from one activity to another [Dumoulin et al. 2020; Rodrigue et al. 2015; Szafir and Mutlu 2012]. Furthermore, given that a VR environment is completely accessible to the viewer and may contain a variety of irrelevant attractive objects that draw a student's attention, distractions for students are also a possibility [Gardony et al. 2013].

[Bozkir et al. 2021] investigated the seating arrangements of the students in the VR classroom, he discovered that those who were seated in the front paid more attention to the virtual teacher and lecture material. One of the limitations of their study was that VR lectures lasting longer than 45 minutes could cause extreme cybersickness. Another study discussed the feasibility of transferring teachers' instructional techniques from the real-world to immersive VR. They evaluated eye tracking based students' visual attention under different intervention strategies and found that participants sitting at the back focused more on the objects irrelevant to the learning contents [Han et al. 2022].

2.3 Detecting Distractions using Eye-Tracking and EEG Data

[Bixler and D'Mello 2016] have applied machine learning employing eye gaze attributes to detect mind wandering in a computer interface while reading text. Another study build a real-time mind-wandering detection and intervention system using eye-gaze data while reading comprehension [Mills et al. 2021]. [Vortmann et al. 2022] used feature by fusing of EEG and eye-tracking data to classify internal vs external directed attention in non-VR application. [Apicella et al. 2021] detected distraction using EEG data from a non-VR application in which participants were distracted by a PC screen-based notification of a distractor task. [Kosmyna et al. 2021] used game applications in AR to detect internal and external attention using EEG and eye tracking data.

[Rahman et al. 2020] proposed several different gaze visualizations as a method for monitoring students who were distracted. However, the accuracy of detecting distracted students was significantly lower for multiple students compared to a single student. [Asish et al. 2021a,b] used deep learning approach to detect the level of distraction using eye gaze data. However, they detect only external distractions. Although eye tracking in VR has been used successfully to measure attention, most of the previous VR research did not examine the internal level of distraction during a class environment. Many educational VR studies fail to capture run-time processes that occur during a VR educational session as they mainly focus on evaluating post-immersion learning with few isolated measures [Antonenko et al. 2010; Ayres 2006; Baceviciute et al. 2020]. These studies supported the hypothesis of an existing relationship between EEG or gaze features and distraction. However, the use of EEG and gaze features and their relation to distraction are complex due to individual variability.

In this work, we designed a system based on machine learning that identifies the internal distraction of a student based on EEG data in VR. We borrowed and modified an educational VR environment with various components (avatar, audio, text slides, and animations) to assist learning [Asish et al. 2021a]. We collected EEG data of participants using this VR environment, to train three machine learning models to detect internal distraction. The participants were presented with a variety of internal distraction events during the VR session to help with collecting data to train our machine learning models. The models' accuracy at classifying internal distraction was evaluated. Our approach is a step toward creating a real-time distraction detection system since it can gauge a student's level of distraction per session.

3 EDUCATIONAL VR ENVIRONMENT

We selected a Virtual Energy Center as our instructional VR setting for virtual solar field visits in order to simulate internal distraction (see Figure 2(b)) [Borst et al. 2016]. In order to describe how the components required for power production work, we used this application as our educational VR environment. The VR environment presented several informational cues (avatar, animations, audio, and slides) simultaneously that have been found to improve learning. Avatars have been shown to boost students' learning [Liang-Yi Chung 2011]. Our environment has a teacher avatar to point at objects and animations that help students look at the component being explained. Such animations have been used in the past to visualize the internal components of an object [Radianti et al. 2020]. An avatar explained the process and components using pre-recorded audio instructions, slides, and animations. Similarly, animations were used to visualize internal operations of solar devices. Audio cues explained several aspects of the solar panel. [Baceviciute et al. 2020] found that audio is not superior to reading text in terms of knowledge retention. Nevertheless, that study did not provide the information by combining the audio with other teaching resources like slides, avatars, or animations. In our study, key terms of a certain component and mathematical ideas/equations were presented on text slides. Our preliminary tests suggested that these slides were helpful for knowledge retention since mathematical concepts/equations are not easy to follow if just explained verbally. [Makransky et al. 2019] found that multimedia slides increases users' interest but creates less learning. In our study, all these components are executed synchronously to explain the subject matter. Additionally, relevant solar field components were highlighted to help students focus on the component being discussed and we assume that combining all educational assets may increase the engagement of students learning. Additionally, we designed another educational VR environment (see Figure 2(a)) related to biology and it was used as a practice session by our participants before starting the actual experimental tasks (described later).

4 METHODOLOGY

4.1 Overview

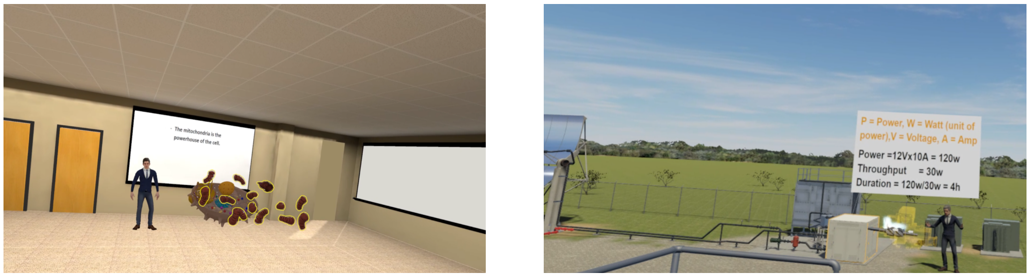
We collected EEG data from the participants while performing internal distraction tasks during an educational VR presentation (see Figure 1). We used this data to train and test three machine learning models (RF, SVM, and kNN) to see which ones work best for this type of data. The details are described in the following subsections.

4.2 Experiment Design

In this experiment, we considered only internal tasks to create distractions since students are distracted due to external reasons or internal thoughts according to real-world classroom distraction studies. We divide our experiment into four sessions with educational sessions and each session ended with two educational quiz questions directly based on the educational content presented.

4.2.1 Internal Distractions. Internal distractions can occur due to many factors such as internal stress, mind-wandering, daydreaming, illness, etc. It is very challenging to simulate internal distractions in an experimental setup due to individual variability. Since we are using an educational VR class to create internal distractions, we mainly focused on relevant internal tasks for the participants. These internal tasks were secondary tasks while their primary task was to watch the educational VR class. We used four internal tasks in four sessions (see Table 1) and they appear randomly once during the educational sessions after 50 to 60 seconds from the start of the session. A doorbell sound, an external distraction, was played following the task description indicating the participants to start the internal distraction task. In our pilot studies, we found that participants in many cases missed the visual instructions to start the internal distraction task. Thus, we decided to alert them using a sound to start the internal distraction task. We made sure to clip the EEG data for our training/testing dataset to avoid any error due to this external distraction (the doorbell sound).

4.2.2 Experiment Questionnaires. Our experiment had three questionnaires: a pre-questionnaire, a post-session-questionnaire and a post-questionnaire. The pre-questionnaire and post-questionnaire were adapted from a related work by [Asish et al. 2021a]. The pre-questionnaire consisted of demographic information and distractability questions from the cognitive failure questionnaire (Table 3) to assess general distraction level in the last six months [Wallace et al. 2002], based on regular activities. Participants answered these questions as 5-point Likert scale. The post-session-questionnaire (Table 2) was filled out at the end of every session to assess if they were distracted (Yes/No) during internal and external distraction events. This helps with data labeling later for the machine learning model. Upon completion of all the sessions, participants filled out a post-questionnaire (Table 4), which was adapted from the work of [Jennett et al. 2008], to gauge their overall VR experience.



(a) (b)

Figure 2: The content is presented by teacher avatars using a combination of text slides, audio, and animations. (a) An avatar explains the nucleus cell during a practice session. (b) An avatar describing a solar panel for data collection.

Table 1: Internal distraction tasks appeared randomly in the VR scene followed by a doorbell sound indicating them to start the task.

Internal task to create distractions	
Q1	Count how many times the avatar moves his hand.
Q2	Count the total number of words starting with 'S'.
Q3	Think about the activities for the upcoming weekend.
Q4	Think about the activities during the last weekend.

Table 2: Post-Session-Questionnaire. It was filled out at the end of every session in each phase.

Post-Session-Questionnaire	
Please summarize what you did after the doorbell sound in this session.	Text type answer
Did you feel distracted during the task?	yes/no

Table 3: Pre-Questionnaire borrowed from [Asish et al. 2022]. Participants answered Q1-Q7 as 5-point Likert-like items. Q8 and Q9 were short text type.

Pre-Questionnaire Questions	
Q1	Do you say something and realize afterwards that it might be taken as insulting?
Q2	Do you fail to hear people speaking to you when you are doing something else?
Q3	Do you lose your temper and regret it?
Q4	Do you leave important letters/emails unanswered for days?
Q5	Do you find yourself suddenly wondering whether you've used a word correctly?
Q6	Do you daydream when you ought to be listening to something?
Q7	Do you start doing one thing at home and get distracted into doing something else (unintentionally)?
Q8	Do you check your mobile in a regular classroom? If yes, how often, provide an approximate time interval like every 5 or 10 minutes?
Q9	What are the common distractions for you in a regular classroom?

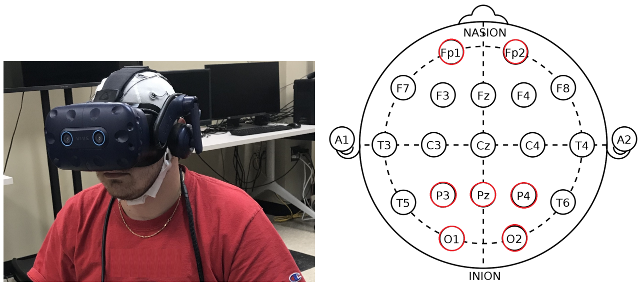
4.3 Participants and Apparatus

There were a total of 21 students (14 male and 7 female) who participated in the study. They were between the ages of 18 and 28 (mean age was 21.2, SD 2.59), and 16 of them had used VR system before. Their races include White American, Black or African American, Asian, and Native Hawaiian or other Pacific Islander. The eye-tracker calibration, consent process, and brief discussion of the subjects' VR experiences all took place during the experiment, which lasted between 35 and 50 minutes.

When watching a VR session with the HTC Vive Pro Eye, we recorded EEG data using the OpenBCI all-in-one electrode cap kit at a sampling rate of 125 Hz. The Experiment (see Figure 3(a) for user setup) was executed on a desktop computer (Core i9 11900F, Microsoft Windows 10 Pro, NVIDIA GeForce RTX 3080 Ti, 64 GB RAM) and Unity 3D v2018.2.21f1 software to implement the VR tasks. We used scikit-learn, TensorFlow, keras libraries in Python (version 3.8.8) for machine learning scripts.

Table 4: Post-Questionnaire borrowed and modified from [Asish et al. 2022]. Participants answered Q1-Q11 as 7-point Likert-like items. Q12-Q15 were multiple choice questions.

Post-Questionnaire Questions	
Q1	To what extent did the VR class hold your attention?
Q2	How much effort did you put into attending the VR class and quiz?
Q3	Did you feel you were trying your best?
Q4	To what extent did you lose attention?
Q5	Did you feel the urge to see what was happening around you?
Q6	To what extent you enjoyed the VR class and quiz exam, rather than something you were just doing?
Q7	To what extent did you find the VR class challenging?
Q8	How much knowledge you could retain after VR class over solar panels?
Q9	To what extent did you enjoy the graphics and the animation?
Q10	How much would you say you enjoyed the VR class?
Q11	Which one helped you to understand the lessons? a) audio b) slides c) avatar d) animations
Q12	Which one helped you to recall information to answer quizzes? a) audio b) slides c) avatar d) animations
Q13	Rank internal distraction events (Highest(4) to lowest (1) distractors). a) Counting words started with "S". b) Counting avatar had movements. c) Think about last weekend activities. d) Think about next weekend activities.



(a) (b)

Figure 3: VR device and EEG equipment used in our study. (a) User Setup for the experiment with the OpenBCI all-in-one electrode cap and the HTC Vive Pro Eye headset. (b) The positions of electrodes in the International 10-20 system used for EEG recording. The red circles indicate the EEG channels selected for our analysis. This image was taken from Wikipedia and has a free image license [Wikipedia 2023].

4.4 Data Collection Process

Upon arrival, participants were provided with information regarding the protocol of the study and informed that they would be seeing educational virtual reality content that consisted of various sessions that ranged in length from 1.5 to 4 minutes. After each session, they were asked to report their distraction level. Subsequently, the participant provided signed consent and was seated two meters away from the moderator. Participants then filled out the pre-questionnaire. Following that, they were asked to wear EEG cap and they put VR headset (HTC VIVE Pro Eye) on the top of the EEG cap. Once the VR headset and EEG device were properly setup, participants were also asked to avoid excessive hand/leg movements as far as possible since the muscle activity in the leg/arms could influence the EEG signal. Participants then went through two practice sessions (using the biology environment shown in Figure 2(a)) to familiarize themselves with the internal distraction tasks. In the first practice session, they just watched the educational VR content. In the second practice session, they were instructed to count how many times the avatar moves his hand after a doorbell sound while watching the educational VR content related to biology. No data was collected during the practice sessions. Following the practice session, we calibrated the eye tracker and checked the EEG cap's functionality. Four sessions of the experiment were then completed by the participants in a random order. After each session, they answered quiz questions (related to the educational content presented) and post-session questions (Table 2). After the completion of the four sessions, they filled out the post-questionnaire (see Table 4) about their experience. Additionally, we asked our participants if they have any feedback about our educational VR experience and which components of the presentation distracted them or helped them with learning the content. Our experimental workflow is summarized in Fig 4.

We recorded EEG data during the four sessions of the experiment. Raw EEG data was collected from the frontal, central, occipital and parietal regions throughout the sessions using the OpenBCI all-in-one electrode cap kit. It had 16 channels and we placed these channels for each participant with electrode locations according to the international 10-20 system [Homan 1988]. The notch filter was applied at the power frequency of 60 Hz, bandpass filter was 1 to 50 Hz and smoothing was turned on from the OpenBCI tool settings while recording the EEG data. The sampling rate for EEG data was 125 Hz. We down-sampled it to 120Hz to match with the eye-tracker sampling rate of 120 Hz.

4.5 Data Pre-Processing

Simple pre-processing steps were applied to EEG data to remove the noisy data/channels. We removed all the outliers by performing a z-score analysis on the data that were three standard deviations away from the mean value.

4.5.1 EEG Data Pre-Processing. In this study, we used 16 channels of EEG cap but EEG data contains lots of artifacts which are hard to process and many of these channels are not related to attention/distraction. Thus, we had just seven channels for recording data for each participant with electrode locations (see Figure 3(b)) of FP1, FP2, P3, P4, Pz, O1 and O2 according to 10-20 system [Homan 1988]. We select these channels because the occipital lobe (O1, O2) is responsible for vision processing [Malach et al. 1995], the parietal lobe (P3, P4, Pz) provides information about attentional demands [Klimesch 1997], and the prefrontal cortex (FP1, FP2) is in charge of decision making, cognitive state, and problem solving [Miller and Cohen 2001]. We used a notch filter at the power frequency of 60 Hz, a bandpass filter from 1 to 50 Hz, and smoothing was turned on while recording EEG using the OpenBCI tool. Due to the nature of the measurement technique, the EEG data contains a significant amount of artifacts caused by eye movements, hand/leg movements, and heart beats [Urigüen and Garcia-Zapirain 2015]. These artifacts are very challenging to remove from the EEG wave analysis. We used the MNE-Python library's independent component analysis (ICA) to eliminate artifacts from the EEG data. In addition, we focused on the experiment design to reduce the occurrence of these abnormalities (asking the user to reduce leg/arm movements as much as feasible) rather than completely eliminating artifacts using independent component analysis (ICA) [Jeong et al. 2019; Urigüen and Garcia-Zapirain 2015]. Since it is hard to control user's unintentional movement, we noticed that 6 participants did leg/arm movements most of the sessions, and hard to remove all those movements. We used the least number of channels (only seven channels) for our classification task so that artifacts can't significantly impact our results analysis. The raw EEG data has been proven to be suitable for classification tasks [Schirrmeister et al. 2017] in the past.

4.5.2 Data Standardization. Data normalization to scale the data was the last step in the pre-processing phase because it is necessary for the ML models [Goodfellow et al. 2016]. We standardized the EEG data with the following equation: $X_t = (Xt - \mu)/\sigma$, where μ represents the mean value of the corresponding data, and σ represents the standard deviation.

4.6 Data Labeling

For the distraction classification task, we took into account two classes: internally distracted (ID) and not distracted (ND). Since VR helps a student to engage more with the contents and distraction does not stay for a long time. Thus, our initial research query was

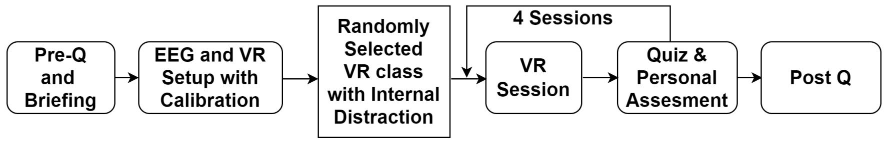


Figure 4: Experiment Workflow.

how long the simulated distraction stays. After the distraction task began, we might take into account a 10, 15, or 30-second distraction window. Our aim is to inform the student in real-time while they are being distracted, though. As a result, if we notify pupils more frequently, the system would itself distract the students. To find a more appropriate window length, we considered 20 second and 30 second window lengths for detecting internal distraction to evaluate which window length provides better results for detecting distraction in our experiment.

For the internal tasks, we sliced and labeled the data as Internally-distracted (ID) for the next 20 second window starting at 5 seconds after the doorbell sound if the participant reported that s/he was distracted (through the post-session-questionnaire). The 20 second window data was marked as Not-distracted prior to the doorbell sound. Similarly, we sliced 30 second window length data for both classes. The remaining data were discarded from the training dataset.

We discovered that the number of data points associated with each class of distraction (ID and ND labels) was significantly different. The data was more skewed in favor of ND class. We used an equal number of slices for the two classes to train the ML models in order to prevent bias. To avoid bias and overfitting with the training data, we applied a combination of under-sampling and Synthetic Minority Over-sampling (SMOTE) [Chawla et al. 2002] so that classifiers can learn from the dataset perturbed by "SMOTING" the minority class and under-sampling the majority class.

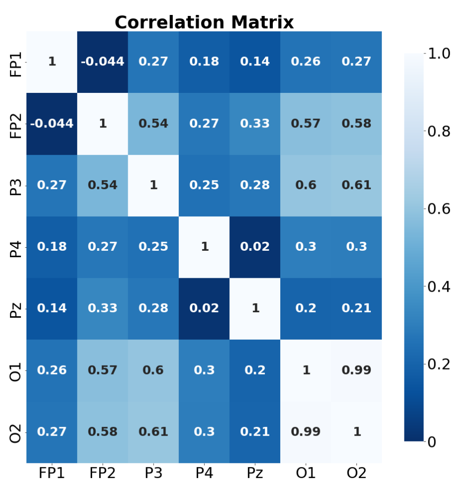


Figure 5: Which features are most related to one another is shown in the correlation matrix with heatmap.

4.7 Feature Selection

A correlation matrix for our features is shown in Figure 5. We observed that most of the channels are positively correlated with other channels except the FP1 and FP2. The FP1 and FP2 are negatively correlated with each other but positively correlated with other channels. Then, we applied a sliding window of 1 second without overlapping to extract signal features. The features used in our ML models include statistical features (mean, standard deviation, min, max) for each feature/column of the current window. We used 7 channels/features and extracted 7*4 = 28 features for each window and then appended all features into the dataset to feed into the ML model.

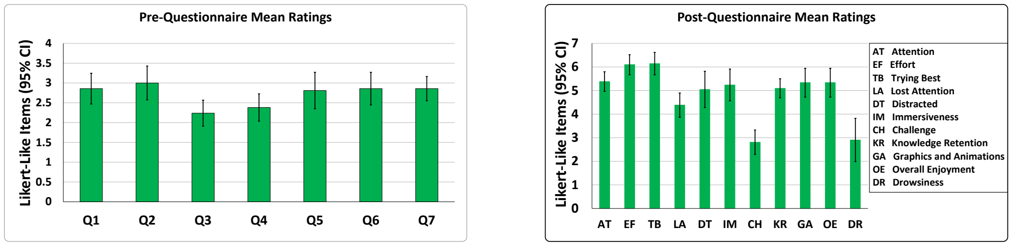
4.8 Classification Models and Evaluation Metrics

Based on previous research [Rodrigue et al. 2015; Vortmann and Putze 2021; Zheng et al. 2020], Random Forest (RF), Support Vector Machine (SVM), and k-nearest-neighbors(kNN) have worked well for classification tasks using the EEG data in non-VR and VR applications. Thus, we chose to use these models for our person-independent classification task.

Due to the EEG signal's nonstationarity, there is more inter-subject variation in the data. The literature has long acknowledged this issue, which makes the cross-subject approach difficult [Apicella et al. 2022]. Since we had 21 participants, we performed 7-fold cross-validation to evaluate our models where in each fold of our 7-fold cross-validation procedure, we used the data from 18 people to train the model and the data from 3 participants (apart from the training participants) to test the model. As recommended by earlier studies for time series classification [Asish et al. 2022], we employed accuracy, precision, recall, and F1-score to assess the models' performance on the task of identifying distractions. We used default parameters to implement these ML classifiers using Python (version 3.8.8) with sklearn library.

Random Forest (RF). : RF is an ensemble learning method which construct multiple decision trees through different data subsets, and voting on the results of multiple decision trees to get the prediction as output of the model. We used "RandomizedSearchCV" library from sklearn to optimize our hyperparameters for RF and we found the optimized parameter where estimator=100, max depth = 150, and max features = 'sqrt'. For two class classification, we changed hyper-parameters to fit the model such as estimator=10 and max depth = 30. We plugged these into the model and reported the results.

Support Vector Machine (SVM). : SVMs are a popular machine learning algorithm used for classification tasks. The basic idea behind SVMs is to find the best hyperplane that separates the different classes of data points in a high-dimensional space. The hyperplane is chosen such that the distance between the hyperplane and the closest data points from each class is maximized, and this distance is called the margin. The SVM model hyperparameters such as



(a) Pre-questionnaire mean ratings.

(b) Post-questionnaire mean ratings.

Figure 6: Pre and post-questionnaire Mean ratings.

kernel function (kernel=rbf), regularization parameter (c=1), and kernel coefficient (gamma=scale) are tuned using Grid Search to optimize the model's performance. These are similar to default hyperparameters.

K-Nearest-Neighbors (kNN). : Utilizing the k nearest neighbors, the kNN classifier executes learning. The data would determine the value of k. Overfitting of the data variance occurs at low k values (low training error and high test error). It works well for k=6 and the parameter metric is "Minkowski" by default. We tested from 1 to 10 to determine the k value.

5 RESULTS

In Figure 6 (a), we have depicted the mean ratings for the pre-questionnaire (Table 3). We observed that the majority of participants reported being easily distracted in social situations. Similarly, in Figure 6 (b), we have summarized the mean ratings for the post-questionnaire (Table 4). Most participants mentioned that they made an effort to remain attentive during the VR experience, but they experienced some internal distractions. The majority of participants expressed enjoyment with the VR content and the graphics/animations presented to them. We asked participants for comments or suggestions (see Table 4, Q12 to Q18) about the educational VR content, and which component(s) distracted them, and which component(s) helped them learn. Out of 21 participants, 16 indicated that audio helped them learn, 16 indicated slides as helpful, 14 indicated animations as helpful, and only 7 indicated the avatar was helpful. Surprisingly, 3 participants mentioned that the avatar distracted them, even though most participants mentioned that all these components work in sync and helped them to learn. Participants also ranked the distractors based on their ability to distract them. For internal distractions, Q2 (see Table 1) was ranked as the highest, Q1 was the second highest, and Q3 and Q4 both as the lowest.

Table 5: Precision, recall and F1-score of the models with 20 second window length.

Name	Class	precision %	recall %	F1-score %
RF	ND	**0.66**	0.49	0.58
	ID	**0.65**	**0.73**	**0.69**
SVM	ND	0.55	0.57	0.56
	ID	0.57	0.61	0.59
kNN	ND	**0.65**	**0.57**	**0.61**
	ID	0.61	**0.73**	0.67

Table 6: Precision, recall and F1-score of the models with 30 second window length.

Name	Class	precision %	recall %	F1-score %
RF	ND	**0.65**	0.48	0.57
	ID	**0.65**	**0.69**	**0.67**
SVM	ND	0.55	0.50	0.52
	ID	0.57	0.61	0.59
kNN	ND	**0.65**	**0.55**	**0.60**
	ID	0.61	**0.72**	0.66

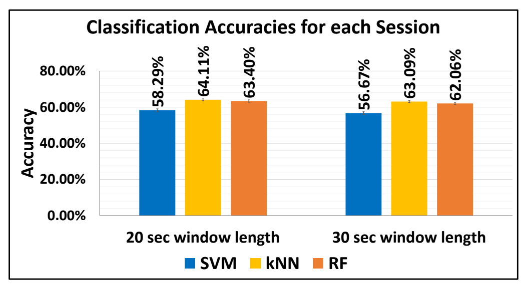


Figure 7: Mean accuracy (best = 64.11%) of the models.

To evaluate our models' performance on the distraction detection task, we used accuracy for cross-subject and cross-session test performance, and precision, recall, and F1-score for individual class performance. Figure 7 reports the mean accuracy value and Table 5 and 6 provide precision, recall and F1-score of 7-fold cross-validations (data from 18 subjects' for training and the data from the remaining 3 subjects' for testing) for our ML models for the EEG features of 20 second and 30 second window length from cross-subject test. Our cross-subject test results show that kNN provides the best results compared to SVM and RF for both window lengths. The shorter window length of 20 seconds provided a slightly better accuracy. Thus, we believe that the distraction does not last long (this answers our research questions). The accuracy for the 30-second window length dropped a little bit. However, this drop was not significant. RF also provides similar results for 20-second and 30-second window lengths. However, SVM performs worse for both window lengths. Our results are very similar to previous EEG-based classification task [Kosmyna et al. 2021; Vortmann et al. 2022], though our internal distractions tasks and VR environment are different. Since EEG data contains artifacts and distractions are complex due to individual variability, it is very challenging to obtain high accuracy.

Table 7: Precision, recall and F1-score for kNN model when 3 sessions are used for training and the remaining session is used for testing. The session used for testing is shown in column 1. The best accuracy achieved from session 4 (65%).

Session	Accuracy %	Class	precision %	recall %	F1-score %
1	**0.64**	ND	**0.64**	**0.58**	**0.61**
		D	**0.63**	**0.71**	**0.67**
2	0.59	ND	0.57	0.54	0.56
		D	0.58	0.67	0.63
3	0.61	ND	0.64	0.53	0.58
		D	0.60	0.71	0.65
4	**0.65**	ND	**0.66**	**0.59**	**0.62**
		D	**0.63**	**0.72**	**0.68**

Since kNN provided the best results, we used this model for cross-session testing to evaluate generalizability on new data. To test this, we trained the kNN model using the data from three sessions and then tested the model's accuracy using the data from the fourth session. Because each session had a different duration and contents, the test set was different for each case. The accuracy, precision, recall, and F1-scores for each session as test case, as provided by kNN models, is shown in Table 7). We see that the best accuracy was over 65% provided for session-I and sessions-IV, and the other sessions have slightly lower results. According to our results (see Table 7) and Figure 7, our ML models provided reasonably good results, using EEG features, for classifying into two distraction classes (ID and ND).

6 DISCUSSION

Our experiment investigated the accuracy of our ML models for classifying the type of distraction (internal distraction and no distraction) using statistical features from EEG data. According to our results (see Table 7 and Figure 7) with two classes, we see that the overall accuracy for kNN model was over 64%. However, when we tested the models for generalizability (using three sessions for training and using the remaining fourth session for testing), the accuracy is similar to previous results [Kosmyna et al. 2021; Vortmann et al. 2022]. This answers our first research question (RQ1) that EEG data could potentially be used to detect internal distractions with a reasonable accuracy. However, further research is required to improve the accuracy of the detection system. Furthermore, our results show that the internal distractions (using the tasks assigned in this experiment) don't last for a long time after they start. We found that a 20 second window was sufficient to detect internal distractions. The accuracy was not significantly better with a 30 second window indicating that the distractions didn't last that long. This answers our second research question (RQ2).

In an educational VR system, our goal is to detect when a student is distracted and alert the students to remind him/her to focus back. For a real VR-based class, we should be able to train the system once and use it on multiple days (ideally, the whole semester) without re-training. Our work is a step toward an educational virtual reality system that detects distractions automatically and in real-time. We believe that such a system could aid in the management of a large guided class (30 to 40 students). Without manual intervention from the teacher, the system could trigger an action (such as pointing towards the object of interest [Yoshimura et al. 2019]) to refocus the attention of pupils who have become distracted. The duration of VR class should not be longer to avoid cybersickness and distraction due to long-time exposure.

Our experiment was not without its flaws. Our dataset and the types of people who participated are both relatively small. Due to the small sample size (n=21), our results may be skewed toward either males or females (7 females vs. 14 males; [Peck et al. 2020]). Simulating internal distractions and getting effective outcomes is particularly difficult due to individual diversity and gender imbalance. In addition, internal distractions may occur even when participants are not instructed to and they are not fully controlled. The test results of our classification models might also be affected by the features we decide to include. Moreover, unintentional muscle activity in legs/arms can affect EEG data and test results. The alpha, beta, gamma, and theta bands of the collected EEG could be taken into account in the future. EEG and eye-tracking data can be combined to enhance internal distraction detection tasks since eye-tracking data can identify attentional states with a respectable degree of accuracy. This hypothesis needs to be tested in future studies.

Maintaining student privacy is an important concern when sharing physiological sensor data of students with the teacher. In our study, EEG data was collected from the participants who gave permission to use their data within a standard informed consent model. The recorded data was anonymized. If such a VR-based system is used for a real classroom, one must ensure that the students understand how the EEG data would be used and get permission from the students (and their parents, for minors) to track or record their EEG data. Special care has to be taken for any longer-term storage to provide security, address legal requirements, and avoid any misuse of EEG data.

7 CONCLUSIONS AND FUTURE WORK

We designed an educational VR environment and created multiple internal distractions for the participants during the educational presentation in VR. We collected EEG data and applied three ML models (RF, SVM, and kNN). We considered two class (internally-distracted and not-distracted) classifications, and two window lengths (20 and 30 seconds) beginning after the 5 seconds of distraction task were taken into account for data labeling. Our findings demonstrate that for both window lengths, kNN performs better than SVM and RF. A little higher accuracy was offered by the 20-second window length. So, we think the internal distraction don't last long in the simulated VR environment.

In the future, we would like to consider more metrics and sensor data (EEG, heart rate, skin conductance, etc.) for detecting distractions (both internal and external). Additionally, we would like to test our machine learning models by extracting the common EEG features such as alpha, theta, delta, beta, and gamma wave. Furthermore, we need to test our approach for real-time distraction detection for a wider range of VR environments using a larger participant pool.

ACKNOWLEDGMENTS

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

REFERENCES

Pavlo Antonenko, Fred Paas, Roland Grabner, and Tamara Van Gog. 2010. Using electroencephalography to measure cognitive load. *Educational Psychology Review* 22, 4 (2010), 425–438.

Andrea Apicella, Pasquale Arpaia, Mirco Frosolone, Giovanni Improta, Nicola Moccaldi, and Andrea Pollastro. 2022. EEG-based measurement system for monitoring student engagement in learning 4.0. *Scientific Reports* 12, 1 (2022), 1–13.

Andrea Apicella, Pasquale Arpaia, Mirco Frosolone, and Nicola Moccaldi. 2021. High-wearable EEG-based distraction detection in motor rehabilitation. *Scientific Reports* 11, 1 (2021), 1–9.

Sarker Monojit Asish, Ekram Hossain, Arun K. Kulshreshth, and Christoph W. Borst. 2021a. Deep Learning on Eye Gaze Data to Classify Student Distraction Level in an Educational VR Environment. In *ICAT-EGVE 2021 - International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments*, Jason Orlosky, Dirk Reiners, and Benjamin Weyers (Eds.). The Eurographics Association. https://doi.org/10.2312/egve.20211326

Sarker Monojit Asish, Arun K Kulshreshth, and Christoph W Borst. 2021b. Supervised vs Unsupervised Learning on Gaze Data to Classify Student Distraction Level in an Educational VR Environment. In *Symposium on Spatial User Interaction*. 1–2.

Sarker Monojit Asish, Arun K Kulshreshth, and Christoph W Borst. 2022. Detecting distracted students in educational VR environments using machine learning on eye gaze data. *Computers & Graphics* 109 (2022), 75–87.

Paul Ayres. 2006. Using subjective measures to detect variations of intrinsic cognitive load within problems. *Learning and instruction* 16, 5 (2006), 389–400.

Sarune Baceviciute, Aske Mottelson, Thomas Terkildsen, and Guido Makransky. 2020. Investigating representation of text and audio in educational VR using learning outcomes and EEG. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–13.

Johanna Bertram, Johannes Moskaliuk, and Ulrike Cress. 2015. Virtual training: Making reality work? *Computers in Human Behavior* 43 (2015), 284–292.

Robert Bixler and Sidney D'Mello. 2016. Automatic gaze-based user-independent detection of mind wandering during computerized reading. *User Modeling and User-Adapted Interaction* 26, 1 (2016), 33–68.

Christoph W Borst, Kenneth A Ritter, and Terrence L Chambers. 2016. Virtual energy center for teaching alternative energy technologies. In *2016 IEEE Virtual Reality (VR)*. IEEE, 157–158.

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, 597–605.

Kirsten R Butcher. 2014. The multimedia principle. *The Cambridge handbook of multimedia learning* 2 (2014), 174–205.

Fabio Buttussi and Luca Chittaro. 2017. Effects of different types of virtual reality display on presence and learning in a safety training scenario. *IEEE transactions on visualization and computer graphics* 24, 2 (2017), 1063–1076.

Marco Cavallo, Mishal Dolakia, Matous Havlena, Kenneth Ocheltree, and Mark Podlaseck. 2019. Immersive insights: A hybrid analytics system forcollaborative exploratory data analysis. In *25th ACM Symposium on Virtual Reality Software and Technology*. 1–12.

Nitesh V Chawla, Kevin W Bowyer, Lawrence O Hall, and W Philip Kegelmeyer. 2002. SMOTE: synthetic minority over-sampling technique. *Journal of artificial intelligence research* 16 (2002), 321–357.

Stéphanie Dumoulin, Stéphane Bouchard, Claudie Loranger, Pamela Quintana, Véronique Gougeon, and Kim L Lavoie. 2020. Are cognitive load and focus of attention differentially involved in pain management: an experimental study using a cold pressor test and virtual reality. *Journal of Pain Research* 13 (2020), 2213.

Anthony G Gallagher and Christopher U Cates. 2004. Virtual reality training for the operating room and cardiac catheterisation laboratory. *The Lancet* 364, 9444 (2004), 1538–1540.

Aaron L Gardony, Tad T Brunyé, Caroline R Mahoney, and Holly A Taylor. 2013. How navigational aids impair spatial memory: Evidence for divided attention. *Spatial Cognition & Computation* 13, 4 (2013), 319–350.

Ian Goodfellow, Yoshua Bengio, and Aaron Courville. 2016. *Deep learning*. MIT press.

Yu Han, Yu Miao, Jie Lu, Mei Guo, and Yi Xiao. 2022. Exploring Intervention Strategies for Distracted Students in VR Classrooms. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts*. 1–7.

Richard W Homan. 1988. The 10-20 electrode system and cerebral location. *American Journal of EEG Technology* 28, 4 (1988), 269–279.

Stephen Hutt, Kristina Krasich, Caitlin Mills, Nigel Bosch, Shelby White, James R Brockmole, and Sidney K D'Mello. 2019. Automated gaze-based mind wandering detection during computerized learning in classrooms. *User Modeling and User-Adapted Interaction* 29 (2019), 821–867.

C. Jennett, A. L. Cox, P. Cairns, S. Dhoparee, A. Epps, T. Tijs, and A. Walton. 2008. Measuring and defining the experience of immersion in games. *International Journal of Human-Computer Studies* 66,9 (2008), 641–661.

Daekyo Jeong, Sangbong Yoo, and Jang Yun. 2019. Cybersickness analysis with eeg using deep learning algorithms. In *2019 IEEE conference on virtual reality and 3D user interfaces (VR)*. IEEE, 827–835.

Mirae Kim, Min Kyung Lee, and Laura Dabbish. 2015. Shop-i: Gaze based interaction in the physical world for in-store social shopping experience. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*. 1253–1258.

Wolfgang Klimesch. 1997. EEG-alpha rhythms and memory processes. *International Journal of psychophysiology* 26, 1-3 (1997), 319–340.

Nataliya Kosmyna, Qiuxuan Wu, Chi-Yun Hu, Yujie Wang, Cassandra Scheirer, and Pattie Maes. 2021. Assessing Internal and External Attention in AR using Brain Computer Interfaces: A Pilot Study. In *2021 IEEE 17th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*. IEEE, 1–6.

Richard Lamb and Elisabeth A Etopio. 2020. Virtual reality: A tool for preservice science teachers to put theory into practice. *Journal of Science Education and Technology* 29, 4 (2020), 573–585.

Liang-Yi Chung. 2011. Using avatars to enhance active learning: Integration of virtual reality tools into college English curriculum. In *The 16th North-East Asia Symposium on Nano, Information Technology and Reliability*. 29–33. https://doi.org/10.1109/NASNIT.2011.6111116

Guido Makransky, Thomas S. Terkildsen, and Richard E. Mayer. 2019. Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and Instruction* 60 (apr 2019), 225–236. https://doi.org/10.1016/j.learninstruc.2017.12.007

Rafael Malach, JB Reppas, RR Benson, KK Kwong, H Jiang, WA Kennedy, PJ Ledden, TJ Brady, BR Rosen, and RB Tootell. 1995. Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proceedings of the National Academy of Sciences* 92, 18 (1995), 8135–8139.

Daniel Martin, Sandra Malpica, Diego Gutierrez, Belen Masia, and Ana Serrano. 2022. Multimodality in VR: A survey. *ACM Computing Surveys (CSUR)* 54, 10s (2022), 1–36.

Jessica S Mendoza, Benjamin C Pody, Seungyeon Lee, Minsung Kim, and Ian M McDonough. 2018. The effect of cellphones on attention and learning: The influences of time, distraction, and nomophobia. *Computers in Human Behavior* 86 (2018), 52–60.

Oliver A. Meyer, Magnus K. Omdahl, and Guido Makransky. 2019. Investigating the effect of pre-training when learning through immersive virtual reality and video: A media and methods experiment. *Computers & Education* 140 (oct 2019), 103603. https://doi.org/10.1016/j.compedu.2019.103603

Tassos A Mikropoulos and Antonis Natsis. 2011. Educational virtual environments: A ten-year review of empirical research (1999–2009). *Computers & Education* 56, 3 (2011), 769–780.

Earl K Miller and Jonathan D Cohen. 2001. An integrative theory of prefrontal cortex function. *Annual review of neuroscience* 24, 1 (2001), 167–202.

Caitlin Mills, Julie Gregg, Robert Bixler, and Sidney K D'Mello. 2021. Eye-mind reader: An intelligent reading interface that promotes long-term comprehension by detecting and responding to mind wandering. *Human–Computer Interaction* 36, 4 (2021), 306–332.

Tabitha C Peck, Laura E Sockol, and Sarah M Hancock. 2020. Mind the gap: The underrepresentation of female participants and authors in virtual reality research. *IEEE transactions on visualization and computer graphics* 26, 5 (2020), 1945–1954.

Johanna Pirker, Andreas Dengel, Michael Holly, and Saeed Safikhani. 2020. Virtual reality in computer science education: A systematic review. In *26th ACM Symposium on Virtual Reality Software and Technology*. 1–8.

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.

Yitoshee Rahman, Sarker M Asish, Nicholas P Fisher, Ethan C Bruce, Arun K Kulshreshth, and Christoph W Borst. 2020. Exploring Eye Gaze Visualization Techniques for Identifying Distracted Students in Educational VR. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 868–877.

Mathieu Rodrigue, Jungah Son, Barry Giesbrecht, Matthew Turk, and Tobias Höllerer. 2015. Spatio-temporal detection of divided attention in reading applications using EEG and eye tracking. In *Proceedings of the 20th international conference on intelligent user interfaces*. 121–125.

Christina L Rucinski, Joshua L Brown, and Jason T Downer. 2018. Teacher–child relationships, classroom climate, and children's social-emotional and academic development. *Journal of Educational Psychology* 110, 7 (2018), 992.

Robin Tibor Schirrmeister, Jost Tobias Springenberg, Lukas Dominique Josef Fiederer, Martin Glasstetter, Katharina Eggensperger, Michael Tangermann, Frank Hutter, Wolfram Burgard, and Tonio Ball. 2017. Deep learning with convolutional neural networks for EEG decoding and visualization. *Human brain mapping* 38, 11 (2017), 5391–5420.

Christian Sinnott, James Liu, Courtney Matera, Savannah Halow, Ann Jones, Matthew Moroz, Jeffrey Mulligan, Michael Crognale, Eelke Folmer, and Paul MacNeilage. 2019. Underwater Virtual Reality System for Neutral Buoyancy Training: Development and Evaluation. In *25th ACM Symposium on Virtual Reality Software and Technology*. 1–9.

Daniel Szafir and Bilge Mutlu. 2012. Pay attention! Designing adaptive agents that monitor and improve user engagement. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 11–20.

Fred Tesch, Donna Coelho, and Ronald Drozdenko. 2011. The relative potency of classroom distracters on student concentration: We have met the enemy and he is us*. In *ASBBS Annual Conference*.

Jose Antonio Urigüen and Begoña Garcia-Zapirain. 2015. EEG artifact removal—state-of-the-art and guidelines. *Journal of neural engineering* 12, 3 (2015), 031001.

Lisa-Marie Vortmann, Simon Ceh, and Felix Putze. 2022. Multimodal EEG and Eye Tracking Feature Fusion Approaches for Attention Classification in Hybrid BCIs. *Recognizing the State of Emotion, Cognition and Action from Physiological and Behavioural Signals* (2022).

Lisa-Marie Vortmann and Felix Putze. 2021. Exploration of person-independent bcis for internal and external attention-detection in augmented reality. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 2 (2021), 1–27.

J Craig Wallace, Steven J Kass, and Claudia J Stanny. 2002. The cognitive failures questionnaire revisited: dimensions and correlates. *The Journal of general psychology* 129, 3 (2002), 238–256.

Wikipedia. 2023. 21 electrodes of International 10-20 system for EEG. https://commons.wikimedia.org/wiki/File:21_electrodes_of_International_10-20_system_for_EEG.svg.

Mihye Won, Mauro Mocerino, Kok-Sing Tang, David F Treagust, and Roy Tasker. 2019. Interactive immersive virtual reality to enhance students' visualisation of complex molecules. In *Research and practice in chemistry education*. Springer, 51–64.

Andrew Yoshimura, Adil Khokhar, and Christoph W. Borst. 2019. Eye-gaze-triggered Visual Cues to Restore Attention in Educational VR. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE. https://doi.org/10.1109/vr.2019.8798327

Lim Jia Zheng, James Mountstephens, and Jason Teo. 2020. Four-class emotion classification in virtual reality using pupillometry. *Journal of Big Data* 7 (2020), 1–9.