Assessment of Student Engagement in Virtual Reality Clinical Immersion Environments through Eye Tracking

Prof. Christine E King, University of California, Irvine

Dr. Christine King is an Associate Teaching Professor of Biomedical Engineering at UC Irvine. She received her BS and MS from Manhattan College in Mechanical Engineering and her PhD in Biomedical Engineering from UC Irvine, where she developed brain-computer interface systems for neurorehabilitation. She was a post-doctorate in the Wireless Health Institute at the University of California, Los Angeles, and a research manager in the Center for SMART Health, where she focused on wireless health monitoring for stroke and pediatric asthma. Her current research is on engineering education and women's health, specializing in pedagogy strategies to promote learning and innovation in design-build-test courses, including senior design, computer programming, and computer-aided-design courses, as well as pre-partum and partum medical devices.

Matthew Lo, University of California, Irvine

Matthew Lo is a graduate student pursuing his Master of Engineering in Biomedical Engineering at UC Irvine. He completed his BS in Biomedical Engineering: Premedical from UC Irvine, during which he lead and participated in a wide range of research, including (1) refining virtual reality modalities to improve medical training, (2) analyzing the mechanisms of slip and maintaining balance, (3) characterizing neuronal endolysosomal vesicle stages in Alzheimer's Disease, (4) developing a spectroscopy device to age bruises primarily for domestic abuse, (5) designing an angioplasty balloon catheter to increase the accuracy of stenting, and 6) creating a treatment device for paravalvular leak. Currently, Matthew continues most of his research from his undergraduate career, and is developing a contactless device for measuring vitals for his Master's capstone. Although mainly focused on medical device research and development, as an aspiring physician-scientist, Matthew has a deep commitment to advancing medical education to ultimately improve both patient care and outcomes through his research to better understand physiological correlations with attention and to create novel and engaging virtual clinical experiences.

Milan Das, University of California, Irvine

Milan Das is an undergraduate student pursuing his Bachelor of Science in Biomedical Engineering at UC Irvine, along with a Minor in Information and Computer Science. In his research work, Milan has developed extensive code to properly quantify electroencephalogram (EEG) and eye-tracking data underlying boredom and engagement in students for the identification of unmet clinical needs. With a passion for interdisciplinary innovation at the intersection of biomedical engineering and computer science, Milan has dedicated his undergraduate career to exploring new frontiers in medical technology. Spearheading projects in artificial intelligence for healthcare, Milan showcases his expertise in developing cutting-edge solutions to address public unmet needs. Currently, he is focused on prototyping a smart insole that utilizes machine learning to analyze gait and detect potential abnormalities that may require medical intervention.

Dalton Salvo, University of California, Irvine

Dalton Salvo is a doctoral candidate in the Dept. of English at UC Irvine. He received his BA from the Univ. of San Diego, a MS in English Literature from the Univ. of Edinburgh, a MA in Rhetoric and Writing Studies from San Diego State Univ., and a MA in English literature from UC Irvine. His current research centers on identifying mental and emotional states generated through human interaction with virtual reality and other virtual artifacts by analyzing physiological data and applying that research to create more effective virtual learning environments. Leveraging this work, he is currently creating a persistent and interactive virtual environment for hosting remote learning classes in the Dept. of Biomedical Engineering at UC Irvine.

Assessment of Student Engagement in Virtual Reality Clinical Immersion Environments through Eye Tracking

Introduction

Biomedical engineering stands at the forefront of addressing complex healthcare challenges through continuous innovation. Cultivating adept biomedical engineers crucially involves exposing them to clinical environments, enabling them to identify and address unmet needs [1-3]. Clinical immersions enable students to identify unmet needs by exposing them to diverse healthcare training and delivery contexts. These experiences equip students to make well-informed design decisions, extending beyond the classroom to the dynamic landscapes of academia and industry [4]. The skill set cultivated through clinical immersion includes proficiency in interdisciplinary communication, a nuanced understanding of the constraints that designs must address, and the capability to identify user needs and interactions that can avoid safety issues, recalls, and patient harm [3, 5]. The FDA has increasingly emphasized the importance of clinical immersion to obtain firsthand feedback from physicians, nurses, healthcare staff, and patients during ideation and initial design phases, as human factors in medical devices is a requirement to ensure safety and efficacy [6]. However, the practical integration of in-person clinical immersion for undergraduate students encounters substantial hurdles, including restricted access to medical spaces and the escalating size of student cohorts. This has led to a reduction in the number of clinical immersion programs and the number of students able to participate in such [7-9].

In response to the above challenges, we have developed a novel virtual reality (VR) clinical immersion platform, enabling a large number of students to access healthcare spaces [10, 11]. Designed for undergraduate biomedical engineering students, this innovative platform offers the opportunity to explore and address unmet clinical needs within a simulated virtual environment. To this end, the virtual reality platform allows students to visualize various medical procedures in 360° and first-person view, enabling them to see both the full workflow of the environment and the physician’s viewpoint. The VR platform has been integrated into a course on Unmet Clinical Needs Finding, piloted during the Spring of 2022 and 2023 as an undergraduate elective in biomedical engineering at the University of California Irvine. Students were instructed in the principles of identifying unmet clinical needs and subsequently applied the platform for needs finding, culminating in the development of a proposed design plan by the end of the course [10]. Throughout the course, students received instruction in a diverse range of topics encompassing team formation, human-centered engineering, regulatory controls, intellectual property considerations, and global, cultural, social, environmental, and economic design factors. The curriculum also covered unmet needs evaluation, go-to-market strategy, and commercialization concepts, ensuring a comprehensive understanding of the multifaceted considerations inherent in medical device innovation. After developing their design plan at the conclusion of the course, students were encouraged to continue their proposed designs throughout their senior capstone course the following year [12].

Pre- and post-survey information from prior studies utilizing VR platforms suggests that the VR experiences greatly improved students’ feeling of immersion and presence, or being physically present in the operating room, compared to traditional 2-dimensional video content [10]. To further assess this from a physiological perspective, and to be able to determine whether students were engaged within the VR environments, we conducted a study in which we collected electroencephalogram (EEG), eye gaze, pupil dilation, heart rate, and heart rate variability from a VR headset (HP Reverb G2 Omnicept VR Headset, HP Inc., Palo Alto, CA) [13]. These preliminary results revealed that boredom and engagement are associated with changes in eye gaze and EEG power spectral densities, particularly in the alpha band (8-12 Hz) over the parietal lobe. The results also demonstrated that EEG can easily have excessive noise and artifacts, which can lead to poor results. However, the signal to noise ratio of the eye tracking data was much higher than the EEG data, demonstrating that the eye tracking data provided the most accurate indication of whether participants were bored or engaged while watching the VR clinical immersion video when compared to the other samples. These promising findings warrant further investigation into using eye tracking to assess boredom and engagement within virtual environments. Because eye tracking can be more easily implemented into online course environments by leveraging the webcams commonly found in most laptops and tablets, it holds tremendous potential for assessing student engagement and improving learning by detecting boredom and then adapting the learning content to each participant’s learning pace.

Given the prior findings through the physiological assessment of engagement within the VR clinical immersion platform, the overarching objectives of the current study are threefold. Firstly, we aim to augment the innovation capacity of undergraduate biomedical engineering students by providing them with virtual reality clinical immersion experiences. Secondly, this study seeks to evaluate the degree of user engagement during these VR clinical immersion sessions through eye tracking in a VR headset by first developing a gauge of when a student is bored versus engaged from eye tracking software. Lastly, it focuses on the analysis of gaze behavior through the integration of eye-tracking technology using webcams so that a more scalable approach to eye tracking can be employed in online courses. The investigation into gaze behavior is undertaken to reveal whether we can assess boredom versus engagement when a student is interacting with learning content in a virtual environment, shedding light on the intricacies of sustaining focus during clinical immersion. It aims to further discover and detail the most common eye gaze patterns that manifest when bored or engaged, respectively. By elucidating these patterns associated with both engagement and boredom, this study contributes not only to the advancement of biomedical engineering education but also resonates with the broader paradigm of virtual education solutions in response to contemporary educational demands.

Methodology

The study performed was approved under exemption approval (UCI IRB Exempt No. 2023-2678) to be able to assess the differences between engagement and boredom and the relative engagement within a VR clinical immersion environment among student participants. We examined a diverse participant pool primarily concentrated in the fields of Biomedical Engineering and medical school education, ensuring a well-rounded and comprehensive group for examination. Undergraduate Biomedical Engineers, Master of Engineering in Biomedical Engineering students, and first- and second-year medical students enrolled at the University of California Irvine were recruited and agreed to participate in the study. The students were asked to first perform an eye-tracking calibration session, followed by watching engaging, clinical immersion, and boring VR videos within the HP Reverb G2 Omnicept VR Headset. The eye tracking sensors were first calibrated using a built-in simulator software [14] that requires the participant to click a mouse to register the x and y coordinates of their gaze while attending to locations on a circle at the center of the screen. The centroid of the circle (0, 0) is placed in the direct middle of the screen, and the (x, y) coordinates are then calculated given the user input and pixel coordinate convention.

After calibrating the HP Omnicept VR Headset eye tracking sensors, we then had student participants watch our three video samples, presenting them in a specific order. Early experiments revealed that the boring video increased levels of exhaustion, negatively impacting participants' subsequent experiences with the other videos. Accordingly, participants watched our sample videos in the following order: (1) interesting, (2) clinical immersion, and (3) boring. Given the subjective nature of boredom and engagement and the study's focus on assessing such phenomena, control samples were deliberately chosen to be either boring or interesting for the broadest participant appeal. As such, we assigned our interesting sample to be a conspiracy theory video on alien reproduction vehicles and alternative energy sources, given the increased public interest in aliens and unknown aerial phenomena in recent years [15]. Our clinical immersion experience was a clip from a spinal deformation procedure in a standard operating room [10]. Lastly, our boring sample was extracted from a 1989 Microsoft Word tutorial, assuming that most students would already be familiar with the information, and the outdated content would enhance the likelihood of generating a strong sense of boredom in our participants [16]. In our pilot study, which utilized EEG and various physiological sensing technologies [11], both our interesting and boring video samples were observed to evoke distinct states of both boredom and engagement as determined by self-reported survey data.

To assess the eye tracking features associated with engagement and boredom, the eye tracking data and synchronized videos were reviewed to qualitatively determine whether any patterns or variations in visual focus were observed as different across our three video samples. To this end, the pupil dilation and gaze location for each eye (X/Y coordinates) were measured in real time at a sampling rate of 120 Hz [13]. The data was then synchronized with the videos using the open-source application programming interfaces (APIs) and developer platform provided by HP [13], Adobe Premiere Pro (Adobe Systems, San Jose, CA) [18], and Matlab for assessment of which eye tracking features correlated with boredom and engagement. Fig. 1 shows how the eye tracking and heart rate data were collected and synchronized with the video data for subsequent analyses. Since our prior research found that the heart rate and heart rate variability data did not provide any conclusive results [11], these data were subsequently removed from analysis. Using the eye tracking data collected, specific features that were further assessed included X and Y pupil location, pupil dilation, eye movement deviation from a neutral resting position (i.e. eye movement from the center to the periphery), and the variance of the eye gaze deviation from the center. The eye gaze deviation from the center was calculated as a 5-point sliding window median filter distance from the center in the X and Y directions using the Euclidean distance formula, respectively, for each pupil, where the center location of the screen is (0, 0).

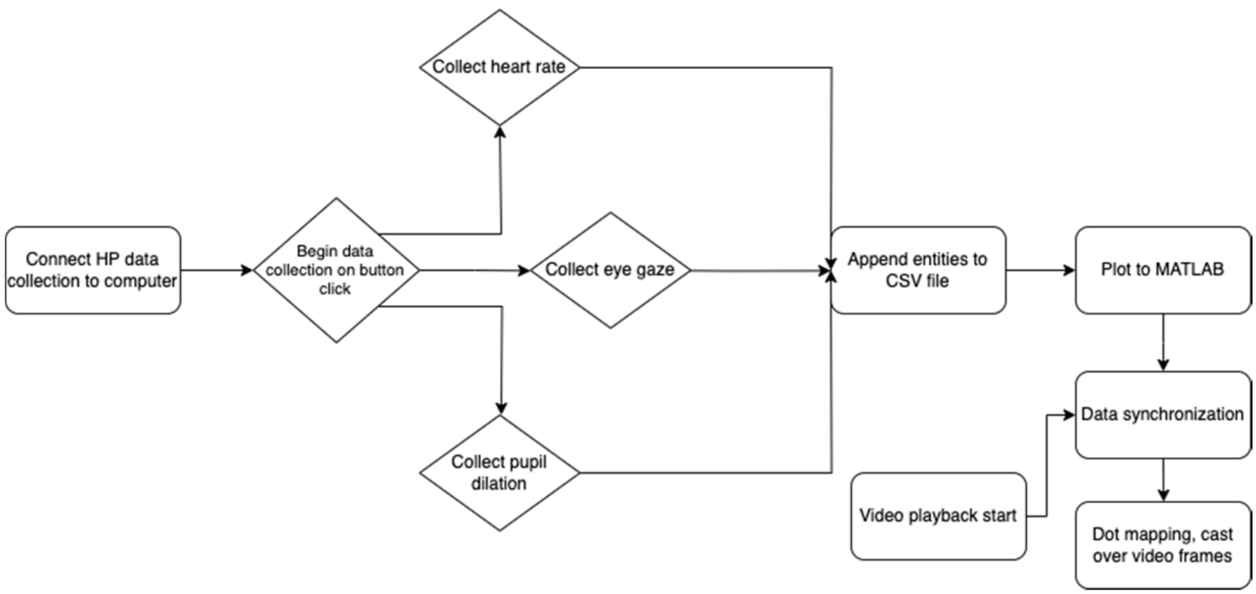

Figure 1: Schematic of the eye tracking data collection process performed from the VR headset.

Quantitative analysis was performed to determine whether there were any statistical differences between the boring, engagement, and clinical immersion videos via the above eye gaze parameters. A one-way analysis of variance (ANOVA) test was performed to compare the means of the three videos, as these were independent samples with normal distributions. The pupil dilation, median eye gaze deviation from the center of the screen, and the variance of eye gaze deviation from the center were analyzed for statistical significance at an alpha level of 0.05.

In addition to the deeper assessment of the eye tracking features associated with the detection of engagement and boredom within the HP Omnicept VR Headset, the accuracy of webcam-based eye tracking was performed to see if a more scalable approach could be utilized to assess boredom within 2D virtual environments using laptops. Since virtual environments are more accessible for use and more scalable if available through laptops over VR Headsets in Massive Open Online Courses (MOOCs) and at other institutions, personal laptop webcams were assessed in terms of eye gaze location accuracy to see if they can accurately detect gaze deviation from center and eye gaze deviation values for future use in real-time interventions within the virtual environments to increase engagement and learning. To this end, participants were asked to follow several short videos in which they were asked to follow a target (e.g. the red ball shown in Fig. 2) while their eye tracking data was recorded. This data was then synchronized with the videos using MATLAB (MathWorks, Natick, MA) software, and cross-correlation analysis was performed to assess the accuracy and time lag between the sensor data and known eye position.

Webcam eye tracking data collection involved creating windows around the face and then each eye for each frame. The algorithm first determined whether a webcam frame was detected, followed by assessing whether a face was present within the frame. Once a face was detected, it was framed, and the eyes within this frame were identified. The eye center location in the X and Y position were then found within each eye region frame. This process was then repeated for each webcam frame detected, subsequently saved as a CSV file. UTC timestamping was utilized to assist with the synchronization of dot location data in the videos for subsequent cross-correlation analysis.

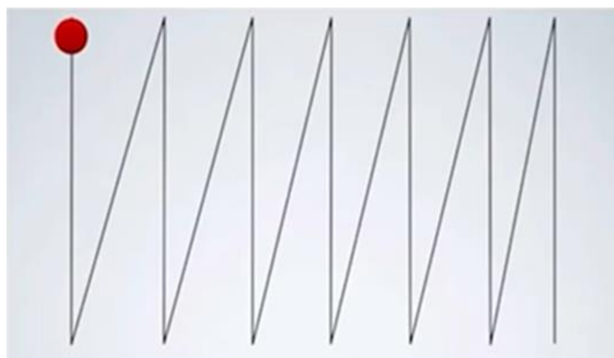

Figure 2: Sample eye tracking calibration video that was utilized to assess the accuracy of the VR headset and webcam's eye tracking sensors.

Results

The eye gaze features collected via headset revealed a qualitative difference in pupil dilation and variance of eye gaze deviation from the center of the screen (Fig. 3). Specifically, among the 12 participants, we observed that six experienced greater pupil dilation when viewing the engaging video compared to the boring video. Additionally, the variance calculated for the median eye gaze distance from the center of the screen exhibited a lower value during the engaging video compared to the boring video in 8 out of the 12 participants.

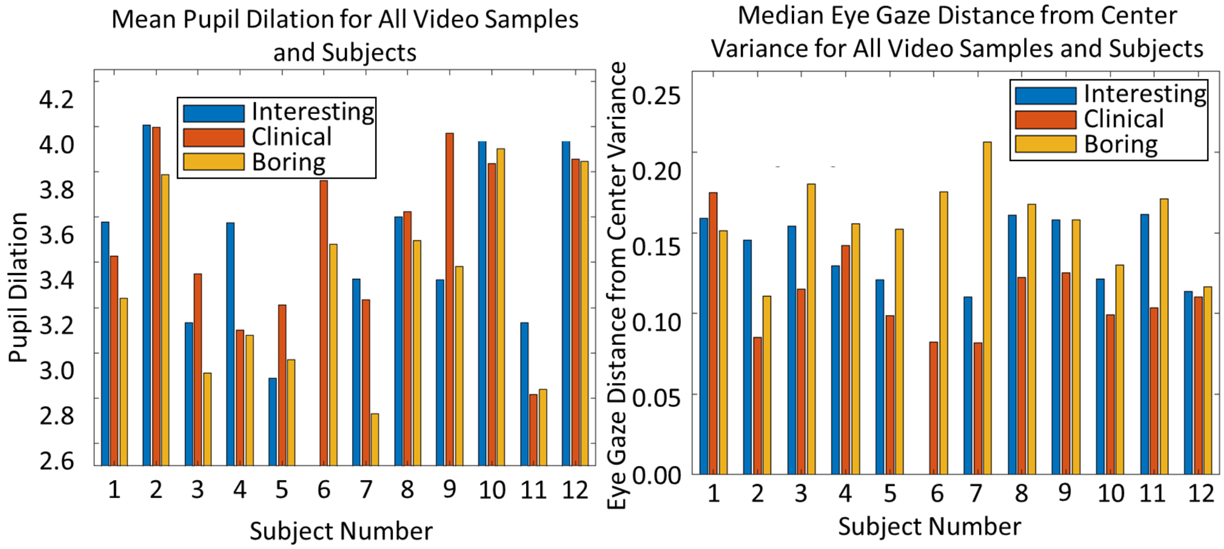

Figure 3: Left: bar graph of the pupil dilation across all subjects and videos. Right: bar graph of the calculated variance of the median filtered eye gaze distance from center.

We implemented a one-way ANOVA test to assess the statistical significance between the video types for each of these features. The evaluation suggests a statistically significant difference in the eye gaze distance from the center variance when comparing the clinical and boring videos ($p = 0.0004$, Fig. 4). All other p-values comparing each feature by video type were inconclusive ($p > 0.05$).

Four students participated in the webcam eye tracking portion of the study using their own laptop webcams. The cross-correlation analysis of the webcam eye tracking for the four tested participants revealed an average cross-correlation of 0.874 at an average lag of 0.03 s (Fig. 5). To further assess relative accuracy of the webcam eye tracking, average correlation values were calculated across the four subjects under different conditions. This included different distances from the screen (1 ft, 2 ft, and 3 ft), the presence or absence of eyeglasses, and in both well-lit and dark room environments (Fig. 6). The correlation values ranged from 0.80 to 0.91, where many

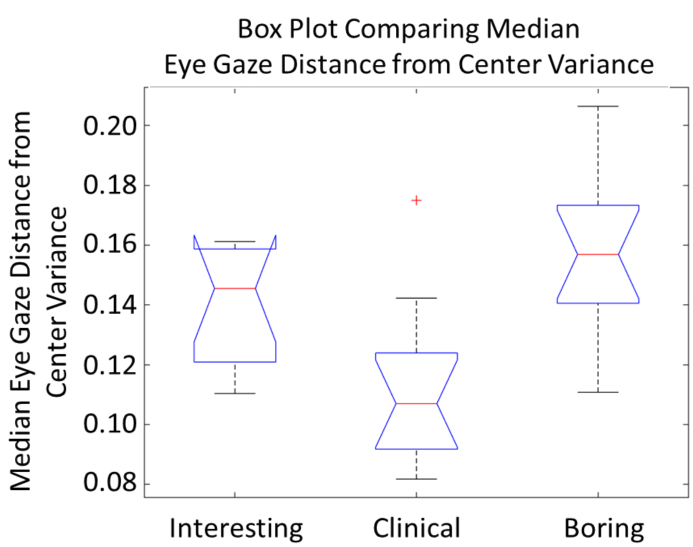

Figure 4: Box plot comparing the medians of the variance calculated for the median filtered eye gaze distance form center across video types.

participants had lower correlation values in a dark room, though this was not statistically significant.

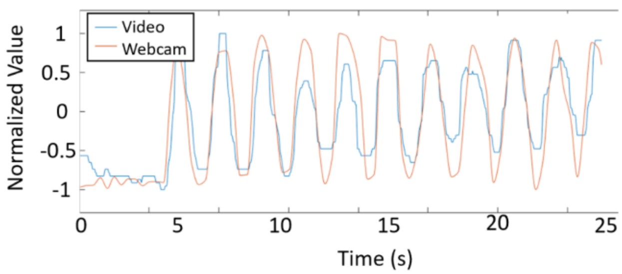

Figure 5: Sample cross-correlation results for the webcam-based eye tracking.

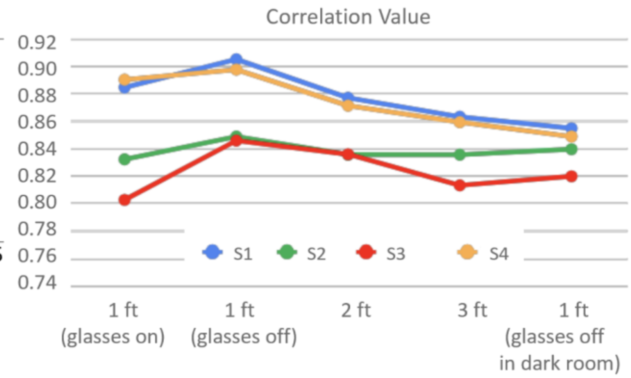

Figure 6: Correlation values under different conditions during calibration assessment of the webcam eye tracking, with differing distances from the screen (1 ft, 2 ft, 3 ft), glasses on or off, and in a low-light environment.

Discussion

Qualitative analysis revealed that the most important features for the assessment of boredom and engagement in a VR environment are pupil dilation and the calculated variance of the eye gaze deviation from the center of the screen. However, it must be noted that there are numerous variables that can influence pupil dilation, perhaps most strongly being the intensity or brightness of light. Given such variables and the relatively small number of samples in our data set, we can only offer a speculative interpretation until further data is collected and analyzed. These values showed that engaging videos have a relatively higher pupil dilation than boring videos. The boring videos had a higher variance when calculating the deviation from the center of the screen, and the clinical video was statistically different from the boring video data using this feature, more closely resembling the engaging video data. The qualitative analyses also demonstrated that the clinical immersion video is oftentimes in between the boring and engaging data, which is corroborated by the self-reports of perceived boredom by the participants who generally scored it as being more boring than the engaging sample but less than the boring sample. Lastly, we hypothesized that the eye gaze patterns associated with boredom could be represented as high arousal or low arousal. When the bored participant exhibited a high arousal state, the participant tended to look around the screen trying to find something to attend to in order to increase their engagement. This would result in a greater variance associated with the distance from center eye gaze values in the left and right eye. Low arousal boredom, or the observation that a participant tends to experience lower gaze variance (colloquially referred to as a “zombie stare”), was observed in only two participants: Subject 1 and 2.

The results of the comparison across the engaging, clinical, and boring videos demonstrated that eye gaze patterns can be utilized to measure engagement within a VR environment. The use of pupil dilation and the variance of eye gaze from the center of the screen can be used to determine whether students are engaged with the virtual learning content by assessing the intensity of their felt boredom. This opens the possibility to create learning content that can algorithmically adapt to each individual user's learning pace. For example, student response system questions (e.g. Clickers, online discussion boards, social media posts) or real-time sentiment feedback systems [19] can be triggered using this approach. Further research will implement these techniques and utilize machine learning methods such as Fisher's Linear Discriminant Analysis (Fisher's LDA) or other forms of discriminant analysis [20] to find which combinations of these features best model the difference between boredom and engagement.

The webcam-based eye tracking results show that there is a relatively high accuracy when utilizing laptop webcams to assess eye gaze behavior. However, few limitations were found. For example, the distance a person is from their webcam and the environment's lighting conditions can significantly affect the algorithm's ability to identify and track eye location. This is problematic in MOOCs, as students can participate in the VR clinical environment in different conditions. This may lead to misclassification of boredom if real-time learning interventions are triggered using eye gaze behaviors and should be measured in future studies to prevent student frustration.

Conclusion

This research proposes practical strategies to enhance educational experiences in academic settings by further physiologically analyzing the mental and emotional states of boredom and engagement and how such data can be leveraged in the creation of educational content. Firstly, further establishing the most common eye movement and pupil dilation patterns associated with boredom and engagement. Secondly, using that data, develop webcam-based eye tracking triggers to create adaptive virtual learning content to enhance overall student engagement and learning experience. Lastly, integrate such triggers into our immersive environments to individualize the learning experience in real time. In particular, immersive multi-user virtual environments complete with voice chat, screen sharing, and digital learning content hold tremendous pedagogical potential. We believe that, when coupled with the ability to not only assess the effectiveness of our learning content and environments but also adapt itself to the students' attention span in real time by tracking boredom and engagement, eye tracking triggered learning interventions will allow us to create pedagogical content that more effectively enhances our students educational experience.

References

[1] Y. Yazdi, "Developing Innovative Clinicians and Biomedical Engineers," *American Journal of Preventive Medicine*, vol. 44, no. 1, 2013. doi:10.1016/j.amepre.2012.09.013

[2] A. Singh, D. Ferry, and S. Mills, “Improving Biomedical Engineering Education Through Continuity in Adaptive, Experiential, and Interdisciplinary Learning Environments,” *Journal of Biomechanical Engineering*, vol. 140, no. 8, 2018. doi:10.1115/1.4040359

[3] M. Kotche, A. E. Felder, K. Wilkens, and S. Stirling, “Perspectives on Bioengineering Clinical Immersion: History, Innovation, and Impact,” *Annals of Biomedical Engineering*, vol. 48, no. 9, pp. 2301–2309, 2020. doi:10.1007/s10439-020-02508-x

[4] S. Miller and S. Higbee, "Work in Progress: A Clinical Immersion Program to Train Biomedical Engineers to Identify Unmet Health Needs in Urban Clinics," *in 2022 ASEE Annual Conference & Exposition*, June 26-29, Minneapollis, MN, 2022.

[5] V. Mittal, M. Thompson, S. M. Altman, P. Taylor, A. Summers, K. Goodwin, and A. Y. Louie, "Clinical Needs Finding: Developing the Virtual Experience— A Case Study," *Annals of Biomedical Engineering*, vol. 41, no. 9, pp. 1899–1912, 2013.

[6] D. Sawyer, K.J. Aziz, C.L. Backinger, E.T. Beers, A. Lowery, S.M. Sykes, A. Thomas, K.A. Trautman, "An Introduction to Human Factors in Medical Devices," Washington, DC: US Department of Health and Human Services, Public Health Service, Food and Drug Administration, Center for Devices and Radiological Health, 1996.

[7] J. Stephens, S. Rooney, E. Arch, and J. Higginson, “Bridging Courses: Unmet Clinical Needs to Capstone Design (Work in Progress),” *in 2016 ASEE Annual Conference & Exposition Proceedings*, June 26, 2016, New Orleans, Louisiana. [Online]. Available: ASEE PEER, Doi: 10.18260/p.26393.

[8] J. Kadlowec, T. Merrill, S. Sood, J. Greene Ryan, A. Attaluri, and R. Hirsh, “Clinical Immersion and Team-based Design: Into a Third Year,” in *2017 ASEE Annual Conference & Exposition Proceedings*, June 24-28, 2017, Columbus, Ohio. [Online]. Available: ASEE PEER, Doi: 10.18260/1-2—28040.

[9] W. H. Guilford, M. Keeley, B. P. Helmke, and T. E. Allen. "Work in Progress: A Clinical Immersion Program for Broad Curricular Impact," in *2019 ASEE Annual Conference & Exposition*, June 15, 2019, Tampa, Florida. [Online]. Available: ASEE PEER, Doi: 10.18260/1-2—33581.

[10] C. E. King and D. Salvo, "Work in Progress: Development of virtual reality platform for unmet clinical needs finding in undergraduate biomedical engineering design programs," in *2022 ASEE Annual Conference & Exposition*, June 26, Minneapolis, MN, 2022.

[11] C. E. King, K. R. Feeney, Q. Tang, M. Das, and D. Salvo, "Work in Progress: Physiological Assessment of Learning in a Virtual Reality Clinical Immersion Environment," in *2023 ASEE Annual Conference & Exposition*, June 25, 2023, Baltimore, MD, 2023.

[12] C. E. King, C. M. Hoo, W. C. Tang, and M. Khine, "Introducing entrepreneurship into a biomedical engineering capstone course at the University of California, Irvine," *Technology & Innovation*, vol. 20, no. 3, pp. 179-195, 2019.

[13] HP Developers. "Fundamentals - HP Omnicept." HP Developer Portal, HP Inc., [Accessed Jan 16, 2024], https://developers.hp.com/omnicept/docs/fundamentals.

[14] HP Developers. “Omnicept Simulator.” HP Developer Portal, HP Inc., [Accessed Jan 25, 2024], https://developers.hp.com/omnicept/docs/simulator.

[15] The Why Files. “How to Build a Working UFO | Alien Reproduction Vehicles (ARVs),” YouTube, Dec. 8, 2022 [Accessed Jan 16, 2024], https://youtu.be/yUFYnVXbLoY?si=oH33R4M0-jnq_ldl.

[16] Whamtan. “THE MOST BORING VIDEO EVER MADE (Microsoft Word tutorial, 1989),” YouTube, Apr. 20, 2014 [Accessed Jan 16, 2024], https://youtu.be/Jk71bPz5VLo?si=GOQh7550N48RT7SF.

[17] Fahlman, Shelley A., Kimberley B. Mercer-Lynn, David B. Flora, and John D. Eastwood. “Development and Validation of the Multidimensional State Boredom Scale.” Assessment, 20:1, pp. 68-85, Feb., 2013.

[18] Adobe Premiere Pro. “Change Duration of Speed Clips” Adobe Systems, [Accessed Jan 25, 2024], https://helpx.adobe.com/premiere-pro/using/duration-speed.html

[19] N. Altrabsheh, M. Cocea, and S. Fallahkhair, "Learning sentiment from students’ feedback for real-time interventions in classrooms," in *International Conference on Adaptive and Intelligent Systems*, pp. 40-49, Springer, 2014.

[20] R. A. Fisher, "The use of multiple measurements in taxonomic problems," *Annals of Eugenics*, vol. 7, no. 2, pp. 179-188, 1936.