Approximating eye gaze with head pose in a virtual reality microteaching scenario for pre-service teachers.[⋆]

Ivan Moser[1,*,†], Martin Hlosta[1,†], Per Bergamin[1], Umesh Ramnarain[2], Christo Van der Westhuizen[2], Mafor Penn[2], Noluthando Mdlalose[2], Koketso Pila[2] and Ayodele Ogegbo[2]

[1]*Swiss Distance University of Applied Sciences, Schinerstrasse 18, 3900 Brig, Switzerland*

[2]*University of Johannesburg, Auckland Park, 2006 Johannesburg, South Africa*

Abstract

Although immersive virtual reality (IVR) technology is becoming increasingly accessible, head-mounted displays with eye tracking capability are more costly and therefore rarely used in educational settings outside of research. This is unfortunate, since combining IVR with eye tracking can reveal crucial information about the learners' behavior and cognitive processes. To overcome this issue, we investigated whether the positional tracking of learners during a short teaching exercise in IVR (i.e., microteaching) may predict the actual fixation on a given set of classroom objects. We analyzed the positional data of pre-service teachers from 23 microlessons by means of a random forest and compared it to two baseline models. The algorithm was able to predict the correct eye fixation with an F1-score of .8637, an improvement of .5770 over inferring eye fixations based on the forward direction of the IVR headset (head gaze). The head gaze itself was a .1754 improvement compared to predicting the most frequent class (i.e., Floor). Our results indicate that the positional tracking data can successfully approximate eye gaze in an IVR teaching scenario, making it a promising candidate for investigating the pre-service teachers' ability to direct students' and their own attentional focus during a lesson.

Keywords

virtual reality, eye gaze, eye tracking, positional tracking, teacher education, microteaching, multimodal learning analytics,

1. Introduction

Immersive virtual reality (IVR) enables the delivery of educational content in situations where traditional in-person instruction would be dangerous, impossible, counterproductive, or simply

Joint Proceedings of LAK 2024 Workshops, co-located with 14th International Conference on Learning Analytics and Knowledge (LAK 2024), Kyoto, Japan, March 18-22, 2024.

*Corresponding author.

†These authors contributed equally.

ivan.moser@ffhs.ch (I. Moser); martin.hlosta@ffhs.ch (M. Hlosta); per.bergamin@ffhs.ch (P. Bergamin); uramnarain@uj.ac.za (U. Ramnarain); christovdw@uj.ac.za (C. V. d. Westhuizen); mpenn@uj.ac.za (M. Penn); nmdlalose@uj.ac.za (N. Mdlalose); kpila@uj.ac.za (K. Pila); ayodeleo@uj.ac.za (A. Ogegbo)

0000-0003-2139-2421 (I. Moser); 0000-0002-7053-7052 (M. Hlosta); 0000-0002-2551-9058 (P. Bergamin); 0000-0003-4548-5913 (U. Ramnarain); 0000-0002-4762-8538 (C. V. d. Westhuizen); 0000-0001-6217-328X (M. Penn); 0000-0002-5094-1074 (N. Mdlalose); 0000-0002-8539-0348 (K. Pila); 0000-0002-4680-6689 (A. Ogegbo)

too expensive [1]. Not surprisingly, there has been a steady increase in research interest, investigating the promise and pitfalls of VR in education [2].

Besides the situational benefits, another important strength of IVR is hardware related. Modern consumer IVR headsets are equipped with an array of various built-in sensors. Originally designed to enable and enhance the experience of immersive games, they can also be exploited for the purpose of gathering real-time user data that can be related to the learning process and outcome. For example, positional data can provide insight about learning outcome [3], cognitive load [4], and social interactions [5].

One particular sensor that has been previously hardly accessible but is finding its way into consumer devices is eye tracking. Put simply, video-based eye trackers emit infrared/near-infrared light and utilize the resulting corneal reflections and their spatial relation to the center of the pupil to estimate eye gaze vectors [6]. In combination with IVR, eye tracking offers unprecedented opportunities to study human behavior and cognition [7]. IVR allows creating highly realistic and controlled environments, and modern game engines make it relatively easy to record gaze directions and areas of interest (AOI) compared to mobile eye tracking systems that track gaze in the real world. It has also been demonstrated that eye trackers integrated into IVR headsets achieve sufficient levels to reliably identify the current fixation location, provided that the gaze targets of interest are not in close proximity [8].

Consequently, the value of eye tracking in IVR could be demonstrated across a wide range of tasks. More specifically, eye tracking was shown to enhance user interactions in IVR, for example object selection [9] or typing on a virtual keyboard [10]. In the context of education and training, it is important to note that eye tracking can be used to infer cognitive load [11], joint attention of learners [12], and the distribution of teachers' visual attention in the classroom [13]. This opens up many possibilities, ranging from personalized IVR learning experiences to enhanced performance feedback for learners and teachers, respectively.

However, despite the promising research findings, eye tracking is still underrepresented in practical settings outside a scientific context. It is conceivable that the higher cost of IVR headsets with integrated eye-trackers make these devices less accessible for educational use cases. This is even more relevant in the case of collaborative learning, where a classroom would need to be equipped with a higher number of head-mounted displays.

Therefore, this study set out to investigate whether the position and orientation (i.e. pose) of an IVR headset offers a viable approximation of eye gaze. The research question was driven by the idea that, provided head pose (hereafter referred to as head gaze) and eye gaze align sufficiently well, the former could be used to substitute the latter, therefore offering a low-cost alternative to IVR headsets with integrated eye trackers.

2. Related Work

Despite the high practical relevance, little research exists to date that studied whether head gaze can sufficiently approximate eye gaze in IVR. However, there is a recent study that argued that head gaze can indeed serve as a proxy for eye gaze in the context of human-robot interaction [14] when the aim is to teach a (virtual) robot about a person's intent, i.e. what object a person is intending to interact with. Similarly, head gaze has proven useful in a scenario involving the collaboration with a virtual agent [15]. In this study, the use of bidirectional head gaze between human participants and a virtual character was shown to have a similar positive effect on task performance as bidirectional gaze using eye tracking.

In the same vein, a few studies from the field of social psychology have utilized head gaze as a proxy for social eye contact. For example, one study investigated how participants interacted with a virtual physician during a simulated clinical visit [16]. The authors reported that the emotional state of the participants influenced the amount of eye contact they made during the conversation with the physician. Another IVR study tracked nonverbal behavior of participants in a virtual classroom and found different patterns of head movement depending on the level of self-reported social anxiety. Participants with higher level of anxiety exhibited more lateral head movement, indicating increased room scanning behavior compared to participants with low levels of anxiety [17].

Both studies made the implicit assumption that users are mostly looking straight ahead when wearing an IVR headset, thus exhibiting little eye-in-head motion range. Although it has shown to be useful to approximate eye tracking with head tracking [15, 14], it is noteworthy that users' eye movements in IVR can show quite substantial deviations from the forward direction of the head pose. Sidenmark and Gellersen investigated the coordination of eye, head, and body movements during gaze shifts [18]. They found that smaller gaze shifts of 25° visual angle or less are predominantly performed with the eyes and without much contribution from the head or torso. However, they also reported large inter-individual differences between users in terms of the eyes' motion range, varying from 20° to 70° visual angle. In line with these findings, another study recently found a high correspondence between eye and head movements in IVR, leading to an accuracy of 75% for AOI with an angular size of 25°, with a substantial drop in accuracy when the AOI were smaller [19].

Taken together, the existing literature shows initial evidence that head gaze can be successfully utilized to approximate eye gaze in IVR, provided that careful attention is directed towards the design of the virtual objects (i.e., AOI). However, we are not aware of studies that investigated the practicability of these findings in applied settings of learning or training. Therefore, the aim of the study was twofold. First, we aimed to evaluate the similarity between eye and head gaze in a dynamic virtual teaching scenario. Based on the previous findings, we hypothesized that we would observe a high correspondence of head pose and eye gaze in a sparsely furnished IVR training environment (i.e., a scene with predominantly large AOI). Second, we investigated whether we could use a machine learning algorithm to successfully predict the correct eye gaze targets based on the head gaze plus additional positional tracking data recorded from the IVR headset and corresponding hand controllers.

3. Methods

3.1. Participants and context

Forty-five pre-service teachers (PSTs) at a large metropolitan university in South Africa participated in the study. The sample consisted of third-year undergraduates from the Department of Science and Technology education.

As an integral part of their third-year curriculum, PSTs are practicing their teaching skills by

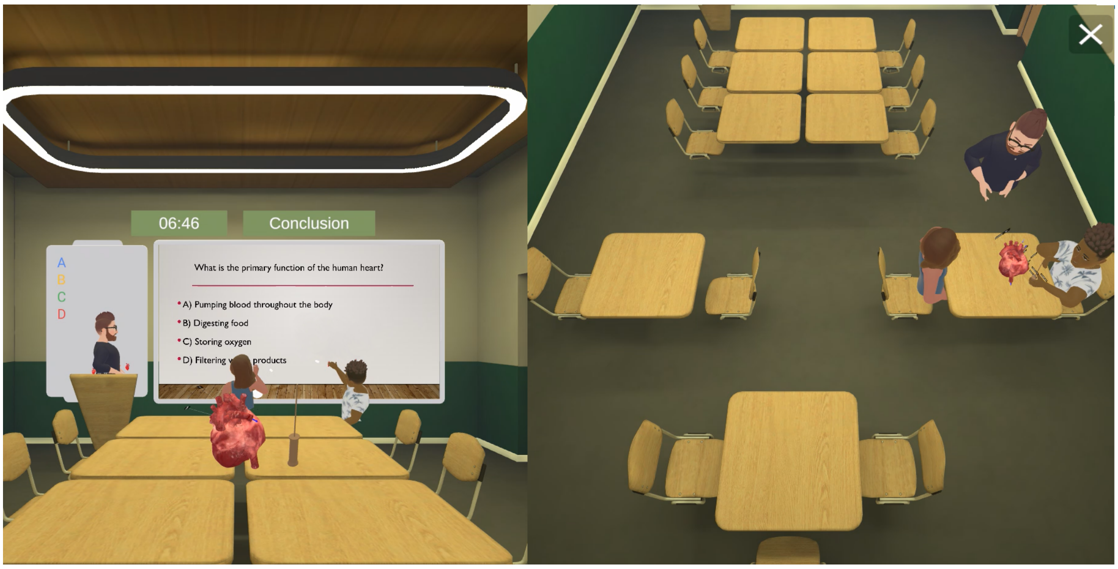


Figure 1: Screenshot from the VR Microlesson application with 2 students in the lesson with the human heart 3D model.

conducting several microlessons throughout their studies. Microlessons are defined as short lesson presentations, typically revolving around a single, tightly defined topic [20]. The goal of microlessons is to develop the PSTs' pedagogical skills in a safe environment and to teach them how to reflect on their own behavior.

In the context of this study, PSTs chose one of sixteen topics from the subjects of biology, physics and chemistry. Their task was to prepare the lesson and deliver it using a learner-centered, inquiry-based teaching strategy inside the IVR environment. In an inquiry-based science classroom, the teacher is seen as a facilitator, who provides ample opportunities for learners to actively engage in the learning process [21, 22]. The study was approved by the local ethics committee of the University of Johannesburg.

3.2. IVR Learning Environment

The IVR application was co-designed with five teacher educators to ensure the alignment with inquiry-based teaching. Hence, the IVR classroom was set up in the Unity Game Engine with two types of tables 1) a main table to accommodate students during the introductory and closing phases of the microlessons, and 2) three separate breakout tables, where groups of two to three students could collaborate on a given task. The classroom was also equipped with a whiteboard for slide presentations and drawings, and a flipchart for displaying quiz results. Furthermore, there was a teacher's podium that hosted a control panel to manipulate various classroom functions (e.g., controlling the slides, starting a quiz, etc.). Importantly, it could also be used to select, spawn, and move 3D objects as well as students between tables. For illustration, a screenshot of the IVR environment is depicted in Figure 1.

Table 1
Features extracted from the microlessons

name	description	type
head position	x, y, z coordinates of the IVR headset, relative to the environment	input feature
hands position	x, y, z coordinates from both hand controllers, relative to the head position	input feature
eye rotation	x, y, z components of the rotation vector (Euler angles) for the eye (average for left and right eye)	support
head rotation	x, y, z components of the rotation vector (Euler angles) for the head	input feature
gaze target head	object in the classroom that intersected with the forward vector (ray cast) from the user's head (computed in the game engine)	input feature
gaze target eye	object in the classroom that intersected with the eye gaze vector of the user (computed in the game engine)	target variable

3.3. Procedure

Before the participants delivered their microteaching lesson, they received a brief training about the IVR classroom including a short hands-on experience to familiarize them with the available tools and objects of the IVR classroom. Then, the PSTs carried out the teaching exercise while changing roles after each microlesson. For example, in a group of four PSTs A, B, C, and D, PST A would first take on the role of teacher, while the other three PSTs would assume the role of students. After a maximum of 15 minutes allocated for the microlesson, they would change roles and repeat the procedure until each PST had completed their lesson. Later, the PSTs received individual feedback on their teaching behavior from their educator based on a recording of the lesson and a learning analytics dashboard.

3.4. Data collection and preparation

From a total of 51 microlessons held, we selected 23 based on the following criteria: functional version of the application able to record the eye rotation data (11 excluded sessions), minimum duration of 5 minutes (17 excluded), and 2 lessons were excluded due to the reusage of the same login for a teacher and a student in the same microlesson. This filtering resulted in excluding 6 out of 24 participants as teachers from the dataset. Eye gaze data was only collected for the teacher roles because these participants wore a Meta Quest Pro headset with integrated eye tracker as opposed to the Meta Quest 2 headsets for the student role. PSTs in the teacher role received feedback on their eye gaze behavior via the learning analytics dashboard after the lesson. Eye tracking was not available for the student roles due to the limited availability and higher cost of the eye tracking enabled IVR headsets.

Raw data was collected automatically from each device during the run of the microlessons. The collected sensors included pose data (position and rotation) from the IVR headset and both hand controllers, rotation for both eyes and the head and the objects where the user was gazing at. These are summarized in Table 1. These sensory data were collected independently for each user and sensor, with different sampling rate for each sensor - 10 Hz for positional data and 20 Hz for the eye tracking data. Hence, the sensory data needed to be synchronized first. This was done by creating 50ms time windows and taking a) the minimum value inside each window time frame for the positional and rotational data and b) the union of all objects present in the eye and head gazing data. Each collected row represents a 50ms long window from a microlesson. Concatenating data from all microlessons generated N=439,749 rows.

We modeled the problem as a multi-class classification. The target variable was the object that was detected as being gazed at by the user's eyes. Before training the model, the dataset had to be cleaned. This included the removal of eye rotation values outside the reported field-of-view of the IVR headset (representing recording failures) and saccades. Several approaches exist to identify fixations and saccades, respectively. We used an Area-of-Interest Identification algorithm, which defines a fixation as a group of consecutive gaze points that fall within the same target area [23]. Groups that did not span a minimum duration of 100 ms were regarded as saccades and excluded from the analysis. This filtering resulted in a reduced dataset of N=186,864 rows.

The gazed objects can distinguish between specific users, but since these users were different across sessions, a class User was created to represent all the users. This processing resulted in a dataset with N=338,040 rows with 13 different classes[1], recorded for 18 users in 23 microlessons, with the average duration 16 minutes (min=5, max=28) and four other students on average being present apart from the teacher (min=1, max=5).

3.5. Analysis

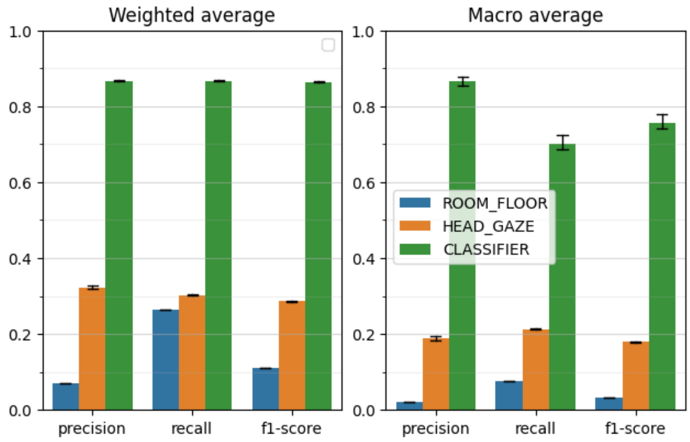
For the modeling, we utilized a Random Forest classifier. This selection was motivated by Random Forest being often one of the best classifiers in Learning Analytics data [24], but also in a recent study on VR data collected from an educational application [25]. We used the implementation from the R ranger package [26]. Due to the large dataset, we did not perform hyperparameter tuning and for the same reason, we used the version of training without replacement and training on a .632 fraction of the data, setting the seed=123, and leaving all other parameters default.

We used 5-fold cross validation (i.e. always 80% of the dataset with 20% left for testing) with the split was stratified by the class distribution. The random forest model (further referred as "CLASSIFIER") was compared to two baseline models. 1) Model "FLOOR" represents a naive classifier that classifies all the instances to the majority class. 2) Model "HEAD GAZE" represents a model that is using the gazed object as derived from the head gaze. All the metrics are reported as mean and standard deviation across the 5 testing folds.

4. Results

For both the machine learning model and the baselines, we report usual metrics for a multiclass classification, i.e. precision, recall and F1-score, which were averaged across all the classes

[1] 3D Object, User, Object other, Flipchart, Room LessonState, Whiteboard, Podium, Room Chair, Room Table, Room Ceiling, Room Floor, Room Wall



Weighted average

	f1-score	precision	recall
ROOM_FLOOR	0.1113±0.0000	0.0704±0.0000	0.2654±0.0000
HEAD_GAZE	0.2867±0.0017	0.3229±0.0028	0.3031±0.0016
CLASSIFIER	0.8637±0.0020	0.8672±0.0019	0.8666±0.0018

Macro average

	f1-score	precision	recall
ROOM_FLOOR	0.0323±0.0000	0.0204±0.0000	0.0769±0.0000
HEAD_GAZE	0.1793±0.0012	0.1896±0.0058	0.2131±0.0008
CLASSIFIER	0.7568±0.0161	0.8670±0.0120	0.7004±0.0160

Figure 2: Precision, recall and F1-score for all three models using weighted average (left), macro average (center) and a table with the results for both average types (right).

using a) the weighted and b) macro average. We focus more on the weighted average because we think it is important to consider the distribution of the classes.

The results of both average types are depicted in Figure 2. We see that the performance using only the baseline model ”FLOOR” is very poor, both for the weighted and macro average. The only value above .20 is a weighted-average recall, due to classifying the largest class affecting the weighted average more than the macro. The ”HEAD GAZE” baseline performs better than ”FLOOR” on all the metrics, showing promising direction. However all the values are below .35, which is still very poor performance.

On the other hand, both averages reveal a steep increase in the performance for the ”CLASSIFIER” over both of the baselines, with the F1-score for the macro-average .7568 and the weighted average .8637. The higher values of the weighted average are caused by a better performance on the larger classes. This is expected, as some of the minor classes might not have a sufficient representation in the dataset to produce good results.

A similar picture about the improvement of the machine learning model compared to the baseline appears from Figure 3, depicting the heatmaps for the ”HEAD GAZE” and ”CLASSIFIER” predictors. While the baseline model matrix is quite scattered, and full of misclassifications for almost every class, the ”CLASSIFIER” model reveals a more pronounced diagonal line indicating higher precision on all classes. For example, it is apparent that the ”HEAD GAZE”, misclassified many objects as ”User”. This would indicate that a teacher is indeed paying closer attention to students, a laudable feature of student-centered teaching. These false-positives for a ”User” are significantly reduced for the ”CLASSIFIER” model. Still, the classifier is far from perfect, especially because of the many misclassifications for the two largest classes ”Room Wall” and ”Room Floor”.

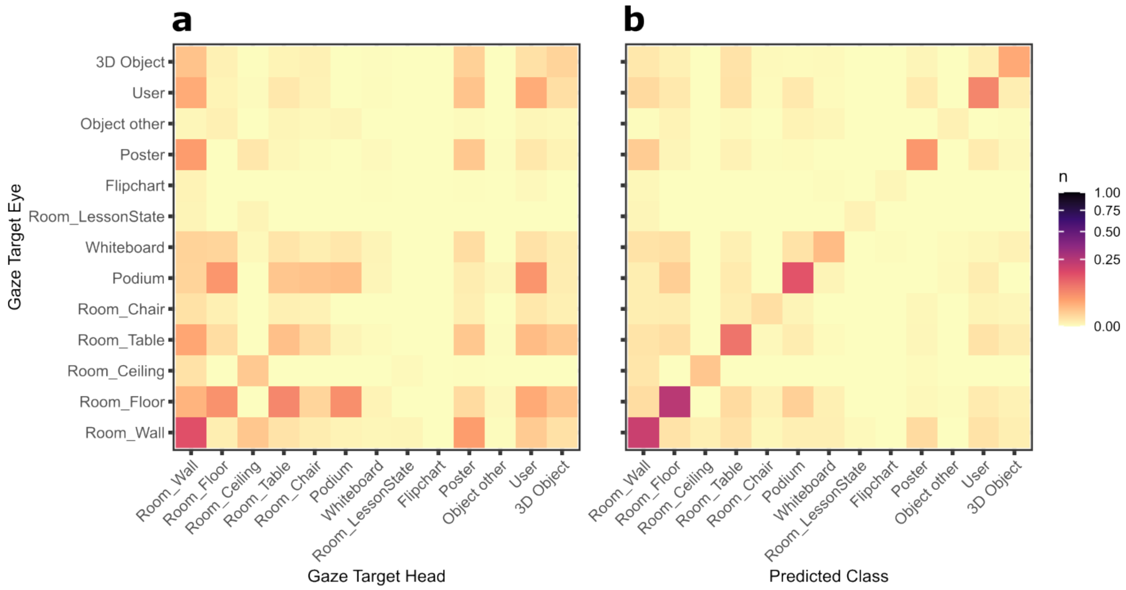


Figure 3: Two heatmaps depicting the correspondence of gaze targets as determined by the eye tracking ("Gaze Target Eye") with a) detected objects from the IVR headset's forward orientation ("Gaze Target Head") and b) the predicted gaze targets of the random forest. Darker shades on the diagonal represent higher classification performance.

5. Discussion

Investigating the use of positional tracking in a microteaching scenario, we found a low correspondence of gaze targets inferred from eye tracking and the forward orientation of an IVR headset, that is, comparing eye gaze and head gaze. This result suggests that head gaze alone does not sufficiently approximate eye gaze, which is in contrast to previous reports claiming that the former can be used as proxy for the latter [15, 14]. However, our finding is consistent with the notion that people exhibit substantial eye gaze deviations from the head forward direction with large inter-individual differences regarding the eye-in-head motion range [18]. Moreover, it is noteworthy that contrary to controlled, experimental studies, we investigated positional tracking in an applied training scenario. Teaching a microlesson in the IVR environment entailed dynamic motion in terms of changing between different locations ("teleporting") and the handling of various interactive tools and objects.

Nevertheless, we could demonstrate the usefulness of positional tracking data in IVR. Although head gaze matched the eye gaze only poorly, submitting the positional data of the IVR headset and hand controllers to a Random Forest classifier, the model was able to predict the fixations of the eye tracking with high precision and recall. More specifically, the results indicated a .8637 ± .0020 F1-score for weighted-average of the random forest. Compared to the F1-score of .2867 ± .0017 for the baseline model with head gaze, this represents an improvement of .5770.

Despite the promising results regarding the usefulness of positional tracking data to predict actual eye gaze during teaching in IVR, it is important to discuss potential limitations of our approach to the data analysis. Although we trained and evaluated our classifier on different data samples, both datasets contained data from the same individuals. It is therefore conceivable that the resulting predictive performance is inflated, i.e., higher than if the model had been evaluated on new participants. This holds particularly true as people show significant inter-individual differences in eye movement behavior [18], which would make the prediction of new participants' behavior challenging. However, it is also important to emphasize that the PSTs rotate roles in the teaching exercise. Therefore, it can be considered adequate to train an algorithm on the PSTs' data in a session when they are wearing an eye tracking enabled headset, and use that model to make inferences about their gaze in the other sessions.

Another potential limitation to note is that classical random forest classifiers do not generate high-quality models on correlated data [27]. This stems from the violated assumption of independent and identically distributed when dealing with longitudinal data. Therefore, a future direction of our research is to use a more computationally intensive model (e.g. a long short-term memory (LSTM) recurrent neural network) designed to handle time-series data.

Finally, we would like to point out that, to our knowledge, no established, independent estimates of the Meta Quest Pro's eye tracking performance exist to date. A preliminary study found an accuracy of 1.652° and a precision of 0.699° standard deviation, which is comparable to other IVR devices with integrated eye tracking [28]. However, the authors of the study also point out to be careful when interpreting fixation results. Their word of caution is related to the findings that the validity and reliability of eye tracking in IVR is influenced by many interacting factors, e.g. the placement of visual targets close or far from the periphery or vision correction [8]. Generally speaking, there is always a certain uncertainty involved in eye tracking research in the absence of an external reference measurement. This is not a specific limitation of this study but rather a general problem of eye tracking research.

For future direction, we are planning to corroborate and validate our findings by a) employing a more adequate machine learning model (see above) and b) investigating how well our results transfer from the teacher to the student role. For this purpose, we are planning to equip students with eye tracking enabled devices too. This would allow us to train and test an ML algorithm on different sessions of the same PST in different roles. Showing that the good predictive performance power of the positional data generalizes across different sessions could have far-reaching practical implications for teacher education. It would equip PSTs and their educators with sophisticated, non-obtrusive ways to measure the attentional focus of teacher and students during a teaching exercise. For example, it could be used to make inferences about the PSTs' ability to distribute their attention to all students equally, and to make them aware of how their behavior compares to that of experienced teachers [13]. Generally, visualizing the attentional focus can greatly contribute to augmenting the feedback the PSTs receive from their peers and educators, therefore improving this central component of teacher training [20].

6. Conclusion

In this study, we investigated to what extent we can approximate the eye gazed objects by a) head gaze and b) a random forest classifier trained using the combination of position and rotation data. This is in the context of an IVR classroom of PSTs practicing their lesson. We found an added benefit of the machine learning model, which showed a good performance, opposed to using only the rather poor results of the pure head gaze approximation. These results are promising as they suggest that in some contexts, using cheaper devices might be sufficient to estimate the eye gaze of IVR users, and enable analytics possible currently only on expensive devices with eye tracking.

Acknowledgments

We would like to thank Lucas Martinic and Ferhan Özkan from XR Bootcamp GmbH for programming the VR application and Deian Popic for creating the 3D assets used in the study. The study was partially funded by a Higher Ed XR Innovation grant from the Tides Foundation.

References

[1] J. Bailenson, Experience on demand: What virtual reality is, how it works, and what it can do., W. W. Norton & Company, New York, NY, US, 2018. Pages: 290.

[2] R. E. Mayer, G. Makransky, J. Parong, The Promise and Pitfalls of Learning in Immersive Virtual Reality, Int. Journal of Human–Computer Interaction (2022) 1–10. doi:10.1080/10447318.2022.2108563.

[3] A. G. Moore, R. P. McMahan, H. Dong, N. Ruozzi, Extracting Velocity-Based User-Tracking Features to Predict Learning Gains in a Virtual Reality Training Application, in: 2020 IEEE Int. Symposium on Mixed and Augmented Reality (ISMAR), IEEE, Porto de Galinhas, Brazil, 2020, pp. 694–703. doi:10.1109/ISMAR50242.2020.00099.

[4] I. Moser, I.-S. Comsa, B. Parsaeifard, P. Bergamin, Work-in-Progress–Motion Tracking Data as a Proxy for Cognitive Load in Immersive Learning, in: 2022 8th International Conference of the Immersive Learning Research Network (iLRN), IEEE, Vienna, Austria, 2022, pp. 1–3. doi:10.23919/iLRN55037.2022.9815894.

[5] H. E. Yaremych, S. Persky, Tracing physical behavior in virtual reality: A narrative review of applications to social psychology, Journal of Experimental Social Psychology 85 (2019) 103845. doi:10.1016/j.jesp.2019.103845, publisher: Elsevier.

[6] V. Skaramagkas, G. Giannakakis, E. Ktistakis, D. Manousos, I. Karatzanis, N. Tachos, E. Tripoliti, K. Marias, D. I. Fotiadis, M. Tsiknakis, Review of Eye Tracking Metrics Involved in Emotional and Cognitive Processes, IEEE Reviews in Biomedical Engineering 16 (2023) 260–277. doi:10.1109/RBME.2021.3066072.

[7] V. Clay, P. König, S. U. König, Eye tracking in virtual reality, Journal of Eye Movement Research 12 (2019). doi:10.16910/jemr.12.1.3.

[8] I. Schuetz, K. Fiehler, Eye tracking in virtual reality: Vive pro eye spatial accuracy, precision, and calibration reliability, Journal of Eye Movement Research 15 (2022). doi:10.16910/jemr.15.3.3.

[9] Y. Wang, R. Kopper, Efficient and Accurate Object 3D Selection With Eye Tracking-Based Progressive Refinement, Frontiers in Virtual Reality 2 (2021) 607165. doi:10.3389/frvir.2021.607165.

[10] M. Zhao, A. M. Pierce, R. Tan, T. Zhang, T. Wang, T. R. Jonker, H. Benko, A. Gupta, Gaze

Speedup: Eye Gaze Assisted Gesture Typing in Virtual Reality, in: Proc. of the 28th Int. Conf. on Intelligent User Interfaces, ACM, Sydney NSW Australia, 2023, pp. 595–606. doi:10.1145/3581641.3584072.

[11] E. Bozkir, D. Geisler, E. Kasneci, Person Independent, Privacy Preserving, and Real Time Assessment of Cognitive Load using Eye Tracking in a Virtual Reality Setup, in: 2019 IEEE Conf. on Virtual Reality and 3D User Interfaces (VR), IEEE, Osaka, Japan, 2019, pp. 1834–1837. doi:10.1109/VR.2019.8797758.

[12] A. Jing, K. May, B. Matthews, G. Lee, M. Billinghurst, The Impact of Sharing Gaze Behaviours in Collaborative Mixed Reality, Proceedings of the ACM on Human-Computer Interaction 6 (2022) 1–27. doi:10.1145/3555564.

[13] O. Keskin, T. Seidel, K. Stürmer, A. Gegenfurtner, Eye-tracking research on teacher professional vision: A meta-analytic review, Educational Research Review 42 (2024) 100586. doi:10.1016/j.edurev.2023.100586.

[14] P. Higgins, R. Barron, C. Matuszek, Head pose as a proxy for gaze in virtual reality, in: 5th international workshop on virtual, augmented, and mixed reality for HRI, 2022. URL: https://openreview.net/forum?id=ShGeRZBcp19.

[15] S. Andrist, M. Gleicher, B. Mutlu, Looking Coordinated: Bidirectional Gaze Mechanisms for Collaborative Interaction with Virtual Characters, in: Proc. of the 2017 CHI Conf. on Human Factors in Computing Systems, ACM, Denver Colorado USA, 2017, pp. 2571–2582. doi:10.1145/3025453.3026033.

[16] S. Persky, R. A. Ferrer, W. M. P. Klein, Nonverbal and paraverbal behavior in (simulated) medical visits related to genomics and weight: a role for emotion and race, Journal of Behavioral Medicine 39 (2016) 804–814. doi:10.1007/s10865-016-9747-5.

[17] A. S. Won, B. Perone, M. Friend, J. N. Bailenson, Identifying Anxiety Through Tracked Head Movements in a Virtual Classroom, Cyberpsychology, Behavior, and Social Networking 19 (2016) 380–387. doi:10.1089/cyber.2015.0326.

[18] L. Sidenmark, H. Gellersen, Eye, Head and Torso Coordination During Gaze Shifts in Virtual Reality, ACM Trans. on Computer-Human Interaction 27 (2019) 1–40. doi:10.1145/3361218.

[19] J. Llanes-Jurado, J. Marín-Morales, M. Moghaddasi, J. Khatri, J. Guixeres, M. Alcañiz, Comparing Eye Tracking and Head Tracking During a Visual Attention Task in Immersive Virtual Reality, in: M. Kurosu (Ed.), Human-Computer Interaction. Interaction Techniques and Novel Applications, volume 12763, Springer International Publishing, Cham, 2021, pp. 32–43. doi:10.1007/978-3-030-78465-2_3.

[20] C. L. Banga, Microteaching, an efficient technique for learning effective teaching, Scholarly research journal for interdisciplinary studies 15 (2014) 2206–2211.

[21] L. B. Duran, E. Duran, The 5E instructional model: A learning cycle approach for inquiry-based science teaching., Science Education Review 3 (2004) 49–58. Publisher: ERIC.

[22] S. Turan, Pre-Service Teacher Experiences of the 5E Instructional Model: A Systematic Review of Qualitative Studies, Eurasia Journal of Mathematics, Science and Technology Education 17 (2021) em1994. doi:10.29333/ejmste/11102.

[23] D. D. Salvucci, J. H. Goldberg, Identifying fixations and saccades in eye-tracking protocols, in: Proc. of the symposium on Eye tracking research & applications - ETRA '00, ACM Press, Palm Beach Gardens, Florida, United States, 2000, pp. 71–78. doi:10.1145/355017.355028.

[24] C. Imhof, I.-S. Comsa, M. Hlosta, B. Parsaeifard, I. Moser, P. Bergamin, Prediction of dilatory behavior in elearning: A comparison of multiple machine learning models, IEEE Transactions on Learning Technologies (2022).

[25] G. Santamaría-Bonfil, M. B. Ibáñez, M. Pérez-Ramírez, G. Arroyo-Figueroa, F. Martínez-Álvarez, Learning analytics for student modeling in virtual reality training systems: Lineworkers case, Computers & Education 151 (2020) 103871.

[26] M. N. Wright, A. Ziegler, ranger: A fast implementation of random forests for high dimensional data in c++ and r, arXiv preprint arXiv:1508.04409 (2015).

[27] C. Ngufor, H. Van Houten, B. S. Caffo, N. D. Shah, R. G. McCoy, Mixed effect machine learning: A framework for predicting longitudinal change in hemoglobin a1c, Journal of Biomedical Informatics 89 (2019) 56–67. doi:10.1016/j.jbi.2018.09.001.

[28] S. Wei, D. Bloemers, A. Rovira, A Preliminary Study of the Eye Tracker in the Meta Quest Pro, in: Proceedings of the 2023 ACM International Conference on Interactive Media Experiences, ACM, Nantes France, 2023, pp. 216–221. doi:10.1145/3573381.3596467.