Research on the simulation potential of virtual reality CAVEs

Michał Mielcarek[1](✉), Miłosz Rzeźniczak[1]

[1] Gdańsk University of Technology, Gabriela Narutowicza. 11, Gdańsk 80-233, Poland
s197110@student.pg.edu.pl

Abstract. The dynamic intersection of technology and human experience, particularly in virtual environments, serves as the core focus of this study. The project, "Research on the Simulation Potential of Virtual Reality Caves", was conducted by the Department of Intelligent Interactive Systems at the Immersive 3D Visualization Lab (I3DVL) located at Gdańsk University of Technology. Led by Michał Mielcarek and Miłosz Rzeźniczak, under the supervision of Dr. Eng. Jacek Lebiedź, the project investigates the immersive capabilities of Cave Automatic Virtual Environments (CAVE) and Head-Mounted Displays (HMD). An important aspect of this research was to develop a method for quantifying user immersion in these virtual environments. The team utilized two VR applications, "Rooms Evoking Negative Emotions" and "Arachnophobia", to observe and analyze participant reactions and engagement levels. Participants were exposed to these applications in both CAVE and HMD settings, allowing for a comparative analysis of the technologies. A key element of this research was the execution of a Systematic Literature Review (SLR), which provided a solid theoretical basis for the study. This review included a thorough examination of existing works and methodologies in VR immersion research, thus informing the design of our study. The methodology incorporated surveys, observation forms, and direct interaction analysis, combining qualitative and quantitative data for a comprehensive evaluation of CAVE and HMD systems. The study involved 124 participants from varied backgrounds, enhancing the data diversity. This study fills a significant gap in existing research by utilizing the rare advanced equipment available at I3DVL, such as the six-wall CAVE. The paper presents the objectives, methodology, and findings, with a focus on comparing the immersion levels in CAVE and HMD settings. The results contribute to the academic discourse in virtual reality and human-computer interaction, offering methodological advancements in measuring immersion and guiding future research in immersive technology.

Keywords: Immersion, Virtual reality, CAVE, HMD.

1 Introduction

Virtual reality (VR) technologies have rapidly evolved recently, becoming integral tools in various fields, from entertainment to education. Numerous technologically diverse devices have been created to facilitate the creation of virtual reality. Currently, many different solutions compete in the market. Despite their differences, each aims to render reality as vividly as possible, immersing the user in the virtual world. To answer whether it is possible to compare the creation of virtual reality through different technological devices by measuring and comparing their potential, this study was conducted. For this purpose, it was decided to limit the investigation to two devices - Cave Automatic Virtual Environments (CAVE) and Head-Mounted Displays (HMD).

In summary, this study focuses on exploring the simulation potential of Cave Automatic Virtual Environments (CAVE) and Head-Mounted Displays (HMD).

The mentioned simulation potential was decided to be examined through measuring the level of user immersion. Immersion is the perception of physical presence in a non-physical, virtual reality world. The first (secondary) objective of the study was therefore to verify whether measuring immersion is possible and how it can be achieved. This objective was also chosen because a systematic review of the literature revealed that there are few research studies where immersion was measured, and there are no clearly defined ways of measuring it. The main challenge identified was that the feeling of immersion is subjective and depends on many factors.

The main hypothesis was that CAVE (CAVE Automatic Virtual Environment) is more immersive than the HMD (Head Mounted Display). This hypothesis was chosen partly because no identical study was found (i.e., one that would compare the virtual reality experience and its aspects between HMD and CAVE). Another factor was the curiosity about how divergent the level of immersion is between such different devices as HMD and CAVE. The cubic CAVE (as used in the study) is very expensive, technologically advanced, and there are very few in the world. HMDs, on the other hand, are quite common. They have become very popular in recent years due to their affordability.

The content of the article is organized as follows: first, related works will be briefly discussed, in the next section, the methods used to conduct the research and how the obtained data were processed will be thoroughly explained. Subsequently, the obtained results will be presented. This will be followed by a brief discussion. Then, the topic of threats to validity will be addressed. Finally, the conclusions and acknowledgments will be presented.

2 Related work

Before starting the research, a systematic literature review was conducted. The primary goal was to review existing studies in the field of virtual reality caves and VR sets. Additional objectives included finding information on the concept of "immersion in virtual reality" and identifying methods used in similar research.

Three academic publication databases were utilized: IEEE Xplore, Scopus, and Springer. Searches in each database were conducted using the search string ("virtual reality" AND "cave automatic virtual environment" AND "immersion"). To gather the broadest and most relevant knowledge, inclusion criteria were set for articles and conference papers from 2015-2023 related to computer science, virtual reality, computer-aided instruction, human-computer interaction, augmented reality, psychology, and cognition. Only publications in Polish and English were considered. The first phase resulted in 144 publications. Based on a vote regarding the adequacy of titles and abstracts, 17 publications were selected for the second phase of the literature review. A similar voting process concerning content led to the selection of 12 publications for the third phase. For these publications, extraction was performed, and the snowballing method was applied, identifying 30 new noteworthy items. After a similar process, six additional publications deemed important were selected for analysis and extraction.

Among the publications, there were several thematically similar works found. For instance, one study investigated the ease of evacuation from a fire and the legibility of signs in a car tunnel using HMD and CAVE devices. There were also several studies measuring the level of immersion, including those that measured immersion on a single device without making comparisons. Some studies compared a three-wall cave to HMD but only in terms of the reception of displayed images (this took place at trade shows, not in a laboratory environment). Another study compared a four-wall cave with HMD in terms of ship simulation quality, but two different applications were launched on both devices. There were publications addressing the issue of immersion only theoretically. Studies were found investigating the issue of orientation or cybernetic disease on two different devices, among many others. None of the found studies compared the level of immersion between HMD and a cubic cave, considered the most immersive.

The systematic literature review uncovered various methods for measuring immersion, such as surveys and subject observation. Details were also discovered about how studies using CAVEs were conducted, including the minimum duration required for a study to allow for immersion. Examples of surveys designed to measure immersion were identified, gathering 224 example questions and challenges encountered in previous research were recognized, including differentiating between immersion and presence. The systematic literature analysis further provided information on procedures utilized in other experiments, such as the sequence of tasks performed by subjects and the duration of breaks.

In summary, the analysis of prior work on the topic concluded that such a study had not been conducted before and that it could contribute to the development of the field and bring a lot of new insights. Additionally, the analysis of prior work allowed for learning from mistakes and utilizing the achievements of already conducted research.

3 Research method

To investigate the simulation potential of CAVE and HMD, two devices were used. A cubic CAVE measuring 3.4m x 3.4m x 3.4m located at the Immersive 3D Visualization Lab (I3DVL) at the Gdańsk University of Technology. This cave utilizes 12 projectors to display images on the walls. Position tracking was done using infrared (participants wore 3D glasses with infrared reflective balls). This allowed for adjusting the perspective based on the participant's position in the CAVE. Subjects' movement tracking was done using two infrared controllers. Sound was emitted through speakers. The second device used by participants was the HMD Valve Index. The study on HMD also took place in the laboratory. The available space was approximately 2.5m x 2.5m. Subjects' position tracking was done using two torches and movement tracking was done using two controllers included in the set. Sound was emitted through headphones built into the HMD.

Two applications were selected for the study. Both were prepared for both HMD and CAVE. The first, informally called "rooms," depicted an apartment consisting of six dark, negative emotion-evoking, interactive rooms. The goal for the subject was to explore all the rooms, interact with the virtual reality, and enjoy the experience. The study duration in this application was 10 minutes. Using this application required the use of the controllers. The second selected application, related to arachnophobia informally called "spiders," aimed for the subject to watch the scenario. The next stages of the scenario were manually switched by the researcher. The study duration was a maximum of 4 minutes. In this application, users did not use controllers.

Since the feeling of immersion is highly subjective, it was decided to use two research methods to obtain the most reliable results. The first chosen method was quantitative in the form of surveys. During the study, each participant filled out three surveys, marking, among other things, their subjective feelings. The second method applied was qualitative in the form of observing the subjects. For this purpose, the participants were recorded, and the later recordings were analyzed. During the analysis, observation sheets were filled out.

The subject filled out three surveys in total. The preliminary survey collected information such as the subject's age, gender, whether they were a student, their experience with using CAVE and HMD, whether they have a fear of heights, suffer from arachnophobia, etc. Additionally, subjects filled out two "post" surveys.
The first was filled out after using the HMD and the second after using the CAVE. Both contained practically identical questions.

The questions were as follows:

1. Was participating in the study stressful for you, did you feel observed/stressed by it?

2. I was so engaged/immersed in virtual reality that I was not aware of things happening around me.
3. I felt disoriented.
4. I felt physically present in virtual reality.
5. As time passed, I felt that virtual reality was more ‘absorbing’ me.
6. I was so engaged in virtual reality that I lost track of time.
7. Do you agree with the following statement: I did not enjoy this virtual reality experience.
8. I felt sick in virtual reality.
9. Was the sense of moving in virtual reality convincing (e.g., did you feel real while walking)? In the CAVE survey: Did seeing your body interfere with your immersion in virtual reality? In the HMD survey: Did not seeing your body in virtual reality negatively affect your sense of immersion in the VR world?

Subjects provided answers using a Likert scale.

During the examination of each participant, two observation forms were filled out: one for the "spiders" application and another for the "rooms" application. These forms captured various metrics, including the duration of time spent in each application. For the "spiders" application, the form queried whether the participant moved away from the spider, their emotional reactions (e.g., screams or other expressions of emotion), the nature of these emotions, their willingness to interact with the spider, the type of interaction (indicating a desire to engage), and whether the participant successfully completed the task without interruption. The form for the "rooms" application contained more comprehensive inquiries due to the application's larger scope and the study's longer duration. Questions included whether the participant attempted to avoid a hole in the floor, removed a kettle from the fire, exhibited any unique reactions (and what those were), the number of objects the participant 'passed through', and any issues encountered with the controllers.

To ensure that the research results were not disturbed, it was implemented so that each person surveyed used two apps once. The detailed course of the survey was as follows:

1. Two individuals were simultaneously invited to the lab, where they each completed an initial survey.
2. Participants were divided, with one beginning their session in the CAVE and the other using the HMD. They undertook specific tasks within one of the applications.
3. Researchers monitored the activities, one observing the participant in the CAVE and another monitoring the HMD user. They documented their observations on designated forms and recorded any supplementary notes.
4. Upon completing the application, participants were asked to fill out a brief survey.
5. The devices were then swapped between participants, with each engaging in the tasks of the alternate application. Observations continued as before, with researchers completing the relevant observation forms.
6. A final brief survey was filled out by participants after they finished using the second application.

The entire process for each participant lasted approximately 30 minutes. Additionally, participants were recorded to facilitate later analysis. Both the first-person perspective and their behavior were captured on video.

4 Results

During the research conducted at Gdańsk University of Technology, 124 individuals were examined over a period of 9 days. As a result, the team gathered 124 preliminary surveys, 248 observation sheets, and 248 "post" surveys. The participants included 31 women and 93 men (Fig. 1). Out of these, 111 were students, while 13 were not (Fig. 2). The sequence of applications and devices is presented in Fig. 3.

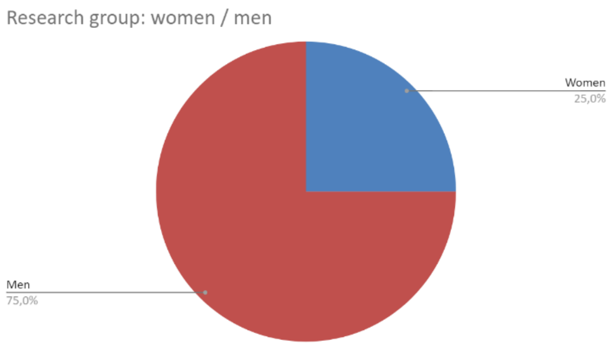

Fig. 1. Research group – woman / men.

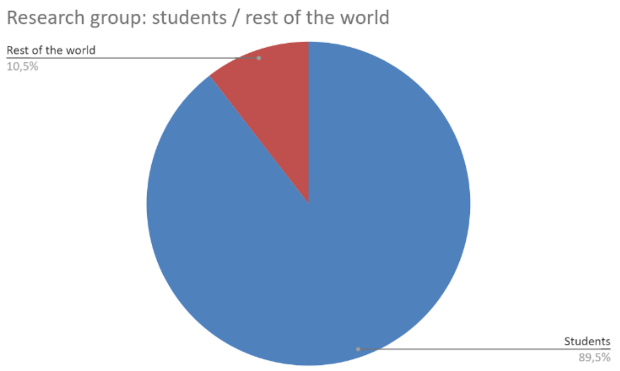

Fig. 2. Research group – students / rest of the world.

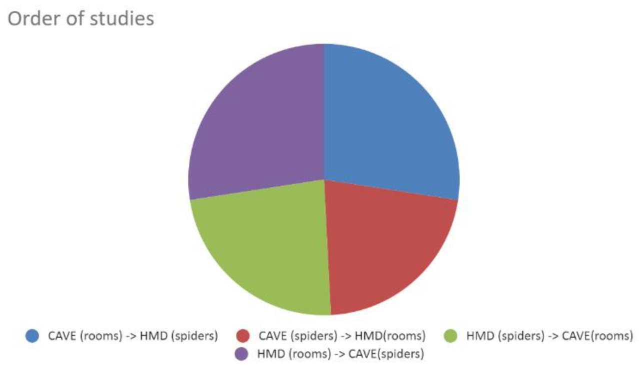

Fig. 3. Sequence of applications and devices.

The obtained results were consolidated. The data from the "post" surveys and observation sheets required special processing. Different columns were transformed in various ways. For the "post" surveys, which were based on the Likert scale, it was decided to first count the number of each type of response for every question, and then calculate the weighted average. The following weights were assigned:

- Strongly disagree: -2
- Disagree: -1
- Don't know: 0
- Agree: 1
- Strongly agree: 2

This was done separately for the "post-HMD study" survey and the "post-CAVE study" survey, yielding separate statistics for both devices. Results are presented in Fig. 4.

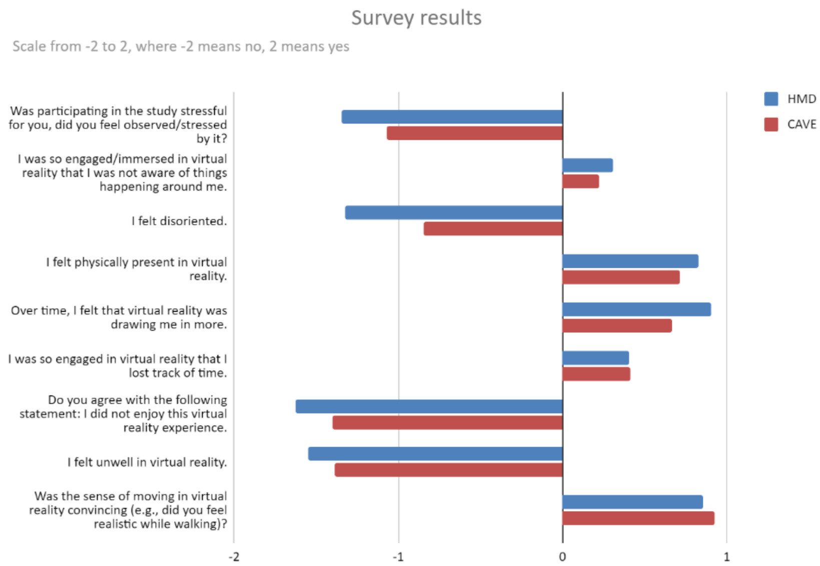

Fig. 4. Post study survey results.

Using the observation sheets, the easiness of interaction was calculated. For the observation sheets, columns "How many balls did the person find" and whether “has visited all rooms" were also selected. The average response to both questions was calculated for each of the two devices. Results are presented in Fig. 5

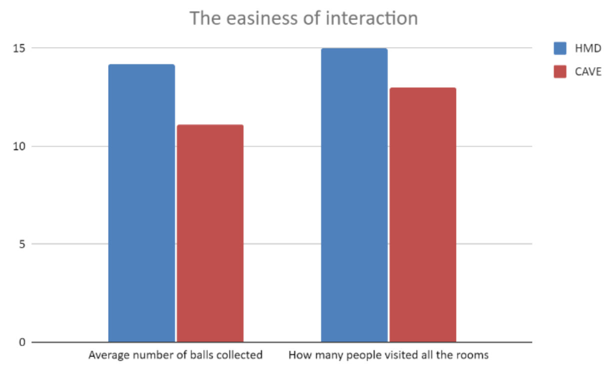

Fig. 5. The easiness of interaction

For the observation sheets, certain questions with the response scale "No", "None", "Yes" were selected (questions deemed not useful for measuring immersion by the researchers were omitted). Three columns were chosen for the "spiders" application, and 15 for the "rooms" application. Then, it was decided to first count the number of each type of response for every question, and then calculate the weighted average. The following weights were assigned:

- No: -1
- None: 0
- Yes: 1.

During this process, the device used for the application was taken into account, ultimately providing separate statistics for both devices. Results are presented in Fig. 6 and in Fig. 7.

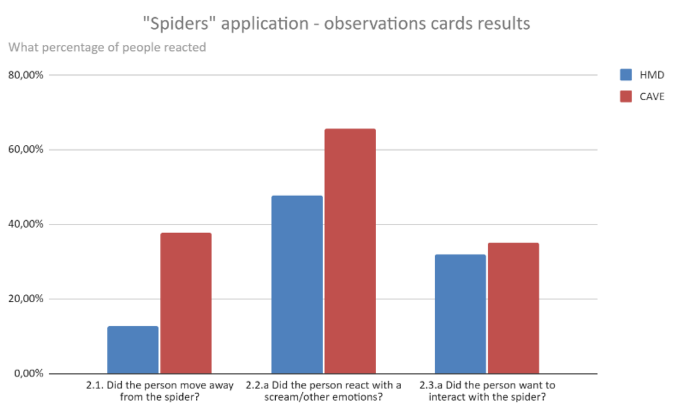

Fig. 6. Observation card results for “spiders” application.

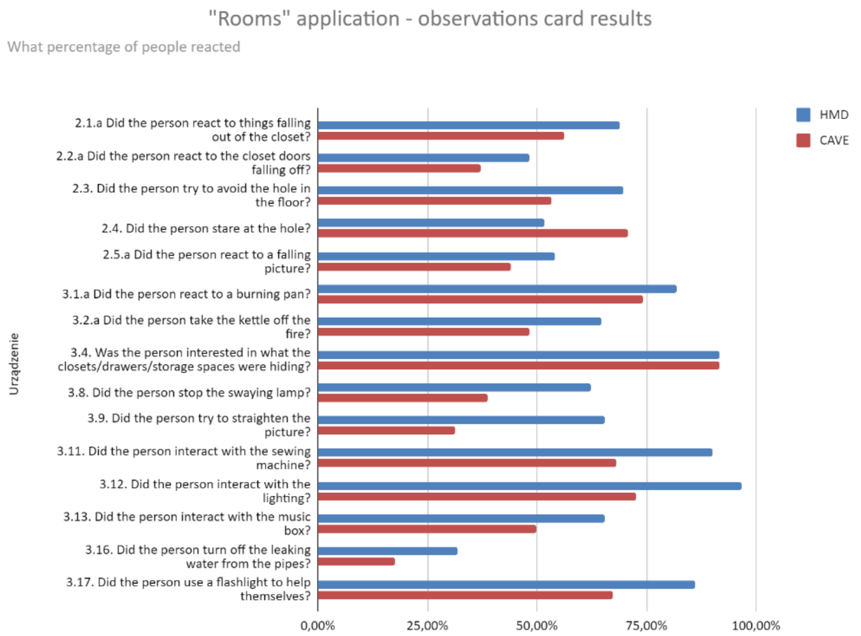

Fig. 7. Observation card results for “rooms” application.

The level of immersion was determined by summing the results obtained from observation sheets and surveys for both devices and applications. Each device could achieve a score ranging from 0 to 2 from:

- observation sheets - "spiders" application
- observation sheets - "rooms" application
- post-study survey

The immersion value calculated in this way could range from 0 to 6, where 0 represents the lowest level of immersion and 6 the highest. CAVE achieved a lower level of immersion (2.87) compared to HMD (3.01). The results are presented in Fig. 8.

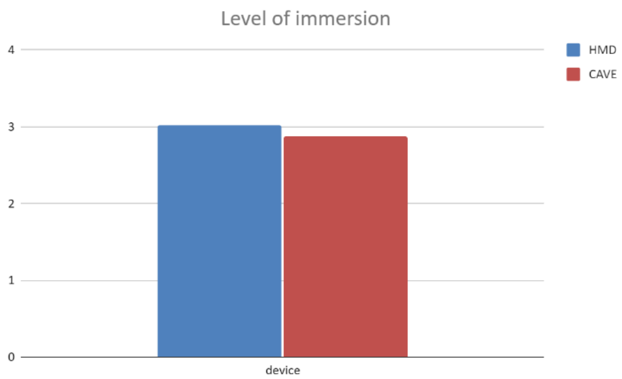

Fig. 8. Calculated level of immersion.

5 Discussion

The main hypothesis was the assertion that the CAVE (CAVE Automatic Virtual Environment) is more immersive than the HMD (Head Mounted Display). Based on the results obtained, it can be stated that CAVE is less immersive than HMD. This refutes the thesis set at the beginning. However, the observed difference is not significant.

Additionally, comparing Fig. 6, which shows the level of immersion in the "spiders" application, with Fig. 7, which shows the level of immersion in the "rooms" application, it can be observed that CAVE achieved a higher level of immersion in the "spiders" application, i.e., the application where subjects did not use controllers and did not have to interact with the virtual world surrounding them. Based on this, it can be argued that if a user only watches the virtual world without interacting with it, the level of immersion is higher in CAVE than in HMD. This is likely because the level of immersion in the "rooms" application could have been lowered by problems related to interaction (which will be discussed more broadly in the Threats to Validity). In CAVE, nine people had problems using controllers, whereas in HMD, not a single person had issues with controllers. This is also evident in Fig. 5. Subjects using HMD more often reached all the rooms in the "rooms" application and were able to collect more balls.

Based on the data obtained, it can additionally be observed that people feel better and find themselves more in virtual reality using HMD, but the sense of moving in virtual reality is more convincing in CAVE (where, unlike HMD, a person sees their own body).

In summary, the thesis that CAVE is more immersive has been refuted. However, the thesis that it is possible to measure the level of immersion has been proven.

6 Threats to validity

The way the study was conducted, the difference between the devices used, and the encountered technical and organizational issues suggest the following threats to validity.

As mentioned in Discussion, in CAVE, there were problems with participants interacting with the surrounding world. This was due to the fact that the CAVE used for the study was established in 2014. This means that the controllers and tracking used were about 10 years old (at the time of the study). At that time, technology was at a completely different level than it is today, and the solutions available then did not work as well as contemporary ones. During the study, there were regular problems, such as losing track of the controllers in the corners of the CAVE. On the other hand, the HMD used in the study was from 2019. Five years is a significant difference in the world of technology. If contemporary controllers were used in CAVE, the results might have been different.

An organizational validity concern arises from the HMD study being conducted in a lab environment where multiple individuals typically worked on computers, occasionally engaged in conversations, and moved about. The ambient noises of footsteps and dialogue, coupled with the feeling of being watched, might have created distractions for participants. In contrast, the CAVE study was carried out in a controlled, sterile setting. This discrepancy in study environments may have contributed to a reduced immersion experience for participants using the HMD.

A technical challenge involved the differing methods used to signal the boundaries of the virtual environment. In the CAVE system, approaching the screen's edge triggered a local color change to red and displayed "STOP" warnings. Conversely, in the HMD setup, nearing the boundary caused the visual field to dim, leading to a loss of vision for the participant. This inconsistency could potentially detract from immersion levels within the CAVE experience.

Moreover, the complexity of entering the CAVE posed another technical validity concern. For HMD users, setup involved simply donning the headset and controllers—a process taking mere seconds. CAVE participants, however, were required to undertake a more elaborate preparation, including shoe cleaning, donning a microphone for staff communication, glasses, slippers for navigating the screen-floor, and handling controllers. The door-closing process for the CAVE also added about 30 seconds to the procedure. Combined with strict rules against touching the CAVE walls or floor without slippers, and the necessity of door closure, this preparatory regimen likely increased stress levels among CAVE participants, as indicated in Fig. 5, question number one - “Was participating in the study stressful for you, did you feel observed/stressed by it?”. Such factors could negatively influence CAVE's immersion levels.

Additionally, the group of subjects was unbalanced. There were significantly more younger people (students) than those from other age groups, and more men than women. This caused the research group to not be very representative

To conclude, fully mitigating validity threats was unattainable, suggesting potential influences on the research findings. Nonetheless, quantifying these effects with precision poses a challenge.

7 Conclusions

This research aimed to investigate the simulation capabilities of Cave Automatic Virtual Environments (CAVE) and Head Mounted Displays (HMD), specifically by measuring user immersion to compare their simulation effectiveness. Conducted at the Gdańsk University of Technology's Immersive 3D Visualization Lab, our study employed sophisticated VR applications to evaluate and document how participants responded to virtual reality in both CAVE and HMD.

Contrary to our initial hypothesis that Cave Automatic Virtual Environments (CAVE) would offer superior immersion, our findings revealed that Head-Mounted Displays (HMDs) provided a slightly higher immersion experience for users. This difference in immersion levels can be attributed to the fact that we utilized a more technologically advanced HMD and an older CAVE system for our study. The newer technology of the HMD likely contributed to its enhanced immersive experience. However, it's important to note that when users were solely observing the virtual environment without interacting with it, the level of immersion was higher in CAVE than in HMD, suggesting that the immersive potential of CAVE systems might be more effectively realized in passive viewing scenarios.

Future research directions should focus on addressing the identified validity threats, exploring the impact of updated CAVE systems on user immersion, and extending the demographic diversity of participants to improve the external validity of findings.

In closing, while our hypothesis that CAVE systems are more immersive than HMDs was refuted, the research successfully demonstrated the feasibility of measuring immersion levels in VR environments. These insights not only advance our understanding of virtual reality's simulation potential but also pave the way for future studies in this evolving field

Acknowledgments. We would like to thank Dr. Eng. Jacek Lebiedź for overseeing this research project, providing invaluable guidance. Our thanks also go to the staff of the I3DVL for their support and to all participants for their contribution to our study.

References

1. Kuhlen T. W., Hentschel B., "Quo Vadis CAVE: Does Immersive Visualization Still Matter?", IEEE Computer Graphics and Applications, vol 34, issue 5, 14-21
2. Lee H., Byun W., Lee H., Kang Y. "Integration and Evaluation of an Immersive Virtual Platform", IEEE Access, vol 11, 1335-1347
3. Mystakidis S., Besharat J., Papantzikos G., Christopoulos A., Stylios C., Agorgianitis S., Tselentis D. "Design, Development, and Evaluation of a Virtual Reality Serious Game for School Fire Preparedness Training", Education Sciences, vol 12, issue 4
4. Maruhn P. "VR Pedestrian Simulator Studies at Home: Comparing Google Cardboards to Simulators in the Lab and Reality", Frontiers in Virtual Reality, vol 2, 746971
5. Leder R., Laudan M. "Comparing a VR Ship Simulator Using an HMD With a Commercial Ship Handling Simulator in a CAVE Setup", International Conference on Harbour, Maritime and Multimodal Logistics Modelling and Simulation 2021, 1-8

6. Elor A., Powell M., Mahmoodi E., Hawthorne N., Teodorescu M., Kurniawan S. "On shooting stars: Comparing CAVE and HMD immersive virtual reality exergaming for adults with mixed ability", ACM Transactions on Computing for Healthcare, vol 1, issue 4
7. Ronchi E., Mayorga D., Lovreglio R., Wahlqvist J., Nilsson D. "Mobile-powered head-mounted displays versus cave automatic virtual environment experiments for evacuation research", Computer Animation and Virtual Worlds, vol 30, issue 6
8. Siess A., Hepperle D., Wölfel M., Johansson M. "Worldmaking: Designing for Audience Participation, Immersion and Interaction in Virtual and Real Spaces", Social-Informatics and Telecommunications Engineering, vol 265, 58-68
9. Tcha-Tokey K., Loup-Escande E., Christmann O., Richir S. "Effects on user experience in an edutainment virtual environment: Comparison between CAVE and HMD", ACM International Conference Proceeding Series, vol F131193, 1-8
10. Brade J., Lorenz M., Busch M., Hammer N., Tscheligi M., Klimant P. "Being there again – Presence in real and virtual environments and its relation to usability and user experience using a mobile navigation task". International Journal of Human-Computer Studies, vol 101, 76-87
11. Chessa M., Caroggio L., Huang H., Solari F. "Insert your own body in the oculus rift to improve proprioception", VISIGRAPP 2016, vol 4, 755-762
12. Ng, A.K.T., Chan, L.K.Y., Lau, H.Y.K. "Depth Perception in Virtual Environment: The Effects of Immersive System and Freedom of Movement" In: Lackey, S., Shumaker, R. (eds) Virtual, Augmented and Mixed Reality 2016, Lecture Notes in Computer Science(), vol 9740
13. Wu, S., Chen, Z., Liao, N., Chen, X. "Study on User Experience of Panoramic Images on Different Immersive Devices" In: Hong, R., Cheng, WH., Yamasaki, T., Wang, M., Ngo, CW. (eds) “Advances in Multimedia Information Processing” – PCM. Lecture Notes in Computer Science(), vol 11165
14. Abade, T., Campos, J.C., Moreira, R., Silva, C.C.L., Silva, J.L. "Immersiveness of Ubiquitous Computing Environments Prototypes": A Case Study. In: Streitz, N., Markopoulos, P. (eds) Distributed, Ambient, and Pervasive Interactions. DAPI 2015. Lecture Notes in Computer Science(), vol 9189
15. Martirosov, S., Bureš, M. & Zítka, T. "Cyber sickness in low-immersive, semi-immersive, and fully immersive virtual reality", Virtual Reality 26, 15–32
16. Kelly, N.J., Hallam, J. & Bignell, S. "Using interpretative phenomenological analysis to gain a qualitative understanding of presence in virtual reality", Virtual Reality 27, 1173–1185