9th International Conference of the Immersive Learning Research Network
Academic Proceedings iLRN2023

DOI: https://doi.org/10.56198/ZH3JOKGN7

Work-in-Progress—Exploring the Effectiveness of Multi-User Basic Workflow Assessment Training in Web XR

César Daniel Rojas Ferrer [1], Masayoshi Ishibashi [1] and Takayuki Fujiwara [1]

[1] Research & Development Group, Hitachi, Ltd., Tokyo, Japan
cesar.rojasferrer.bv@hitachi.com

Abstract. This paper introduces a Web XR system for immersive multi-user basic workflow assessment training in the industrial sector. As society ages and remote work becomes more prevalent, the need for virtual training solutions in education and the workforce has become increasingly important. The proposed system uses off-the-shelf headsets and advanced Human-Computer Interaction (HCI) design principles to provide an immersive remote online training solution. To evaluate the effectiveness of the system, experiments were conducted with voluntary engineers (n=9) using a custom questionnaire. The results of the experiments showed that the system was well-received, with good acceptance scores on a 4-point Likert Scale (M=3.3, scale range=1-4). The open comments also provided useful feedback on important HCI elements that could enhance the effectiveness of multi-user virtual training with Web XR in future iterations of the system. This work-in-progress paper contributes to the development of effective virtual training solutions that can address the challenges faced by the industrial sector in the age of remote work and social distancing.

Keywords: Web XR, Immersive, Training, Industry, HCI, Education, UX.

1 Introduction

In recent years, the world has been faced with new challenges such as an aging society and the new normal brought about by the COVID-19 pandemic, leading to an increasing demand for remote and flexible training solutions. Virtual Reality (VR) has been used for over a decade with promising results in multi-user safety training and basic procedural workflows, such as crane dismantling and identifying working target locations or car assembly lines [1, 2].

From a practical perspective, making these training solutions accessible online while ensuring good performance, smooth interactions, and accessible hardware is crucial to ensure their long-term adoption by business clients and overcome the Proof of Concept (PoC) phase, which is a common fate for industrial solutions. In this context, Web-based Virtual and Augmented Reality (Web XR) has emerged as a promising technology, offering numerous applications in industrial training [3].

One of the key advantages of Web XR is that it enables interactive, immersive, and accessible virtual training experiences from anywhere with an internet connection. However, to ensure the long-term success of these Web XR implementations, it is necessary to identify the elements required for effectively training multiple users performing their task-relevant basic workflow actions, while also allowing their trainers to assess the trainee's performance during the virtual session. In this work-in-progress, we aim to examine the effectiveness of our system in facilitating Basic Workflow Assessment Training (hereinafter referred to as BWAT) in Web XR, focusing on identifying the crucial HCI elements, such as shared gaze among users, and the effect of avatar interaction with the UI and the Virtual Environment (VE).

Given the challenges posed by an aging society and the new normal, our research has the potential to provide valuable insights into the effectiveness of Web XR in enhancing the training and development of multi-user basic workflow skills and to inform the design of more effective and accessible training programs in the future.

2 Method

2.1 Experiment Setup

We designed and implemented an experimental system to evaluate multi-user BWAT in a real-world business application through collaboration with relevant industry players as potential end-users in their work environment. The system was developed using Three.js, a cross-browser JavaScript-based library for developing 3D content with WebGL on web browsers, and increasingly popular as a framework for Web XR solutions for multiple users [4].

The experiment was conducted in our collaborator's office spaces with one experimenter as the facilitator (i.e., role-playing Engineer) of the BWAT and three subjects participating remotely as observers. The subjects and facilitator used standard off-the-shelf Head Mounted Displays (i.e., Meta Quest 2 headsets) to experience the simulation via Quest's native web browser.

The users were represented by adapted 3D half-body avatar models from Hubs by Mozilla [5]. Previous work on avatar communication in videoconferencing scenarios [6] and guidance in industrial settings [7] showed positive results for avatar mediated XR interaction. However, we needed to test the effect of the chosen avatar format in a multi-user BWAT scenario. Therefore, besides standard hand and head tracking features, the system included voice communication via Microsoft Teams voice call, a new dynamic hand-locked menu for task-relevant instructions, and shared gaze pointers to help users understand the Engineer visual attention during each task.

2.2 Task

The participants first receive an explanation of the experimental content. Then, they put on the HMD and receive instructions on the headset and controller operation, followed by a demo to reduce the novelty effect exerted by new technology. After completing the demo, they take off the HMDs and answer a preparatory questionnaire.

After the introductory phase is complete, the participants then put the HMDs back on and proceed in groups of three to experience the virtual content, with an experimenter acting as an Engineer performing the work and providing relevant explanations for the BWAT task. The participants act as evaluators, providing their opinions based on their previous knowledge of the work presented in the VE. After the experience, they remove the HMDs and answer a custom questionnaire evaluating key aspects of the system. Finally, they participate in a group interview with an experimenter and instructor to analyze their impressions and provide subjective opinions on the prototype.

The group format allows for dynamic interviews and more time for each participant to express their opinions while the experimenter can easily follow each participant's level of engagement. Fig. 1 shows a representative illustration of the experiment workflow.

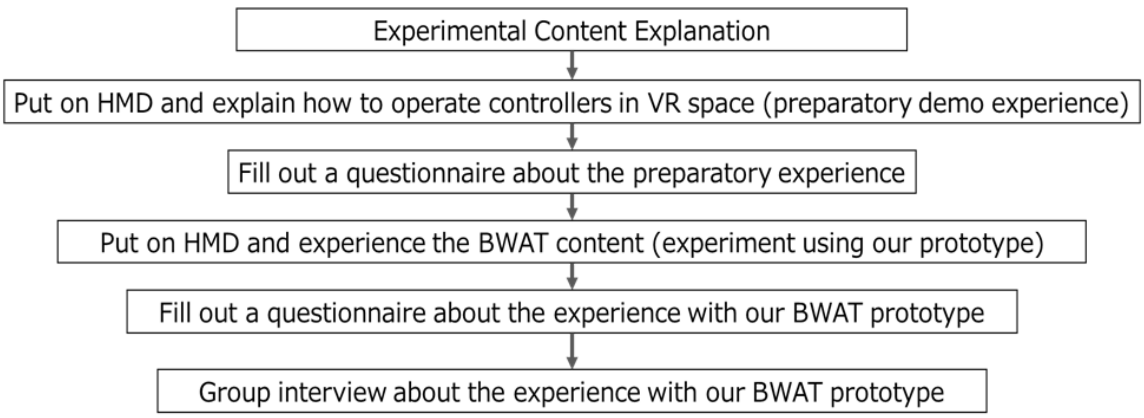

Fig. 1. Experimental workflow.

2.3 Participants

For the experiment, we gathered 9 volunteer experienced engineers who were familiar with the type of task presented (i.e., they had experience performing BWAT content in real life as part of their jobs) from our industry collaborator's staff (9 males), with ages ranging from 40 to 50 (M= 42.22, SD= 9.71).

Most of the participants were not familiar with Web XR or VR, with only one participant reporting limited experience with VR entertainment applications. A picture depicting the users and the VE during the experiment can be seen in Fig. 2.

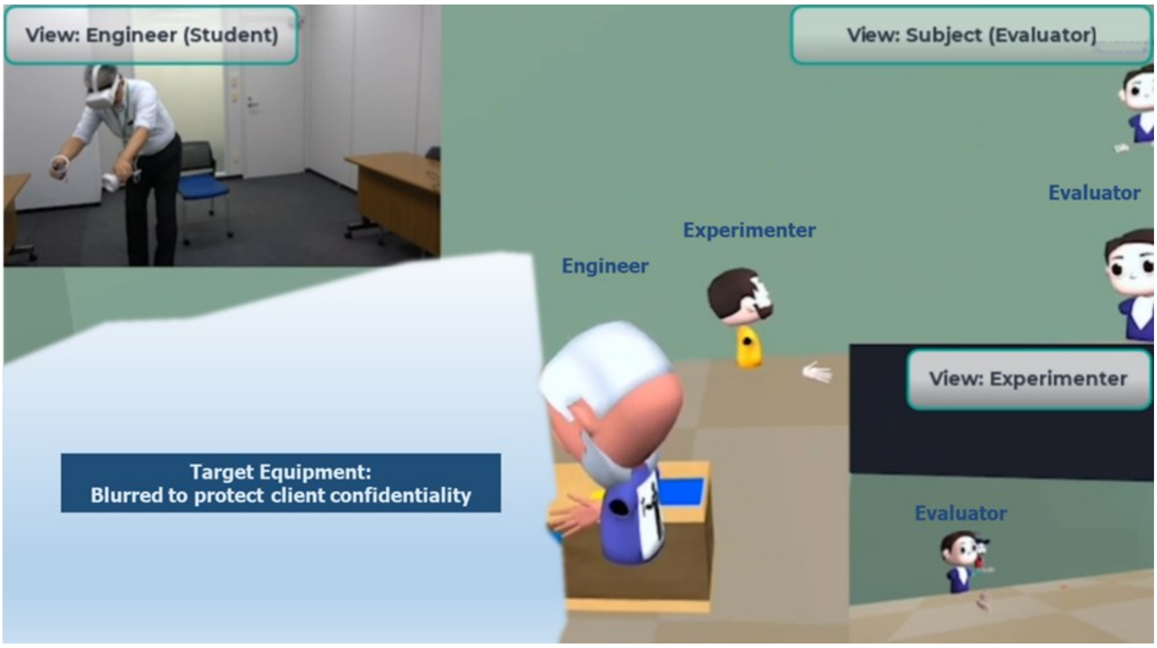

Fig. 2. Experimental Content View. Top Left: Engineer (Role-Playing Experimenter). Center: Virtual Environment View. Bottom Right: First-Person View of a Participant.

As the experiment was conducted in a real-world business scenario at our client's premises, the target equipment has been blurred to ensure the confidentiality of the client's products.

3 Metrics

We measured two types of metrics using mixed methods, one focusing on the user experience in relation to the presented virtual environment (VE) via a custom-designed psychometric questionnaire and a guided interview to further explore and corroborate each user's opinions.

Based on the real-world task presented and the cooperating client's business needs and concerns, we created a customized questionnaire consisting of 22 items measured on a 4-point Likert scale aimed at evaluating the system acceptance level by the users for the BWAT task. Typically, 4-point Likert scales are designed to force the respondent to provide opinions either in favor or against a hypothesis, resulting in an ipsative opinion (i.e., there is no neutral answer) [8]. In this study, a score of 3 or higher indicates acceptance of the presented elements. The questionnaire items were grouped into 6 main factors, with some questions aiming to evaluate more than one factor: 10 items for Shared Space (i.e., VE quality), 13 items for Avatar Configuration (i.e., appearance and overall usability), 4 items for Avatar Movement, 5 items for Avatar Voice Communication, 10 items for Avatar Cooperation, and 4 items for 3D Training Object Manipulation (e.g., placing target equipment in the VE).

The decision to create a custom questionnaire instead of using standard psychometric tools in the field, such as the System Usability Scale (SUS), was based on the needs of the task at hand and the trainer's well-defined requirements.

4 Results and Discussion

4.1 Web XR

We calculated the scores from the questionnaire and determined the level of acceptance for each item evaluated. However, given the small sample size and evidence of a non-normal distribution in the data after visual inspection, we cannot state that the scores are statistically significant. Nevertheless, they provided initial insights into the direction for future iterations of the system, which were later confirmed through group discussions of the score results. The average scores were Shared Space (M=3.2), Avatar Configuration (M=3.3), Avatar Movement (M=3.5), Avatar Voice Communication (M=3.8), Avatar Cooperation (M=3.2), and 3D Training Object Manipulation (M=3.0).

The overall score indicates acceptance of the proposed system for BWAT (M=3.3). However, among the six factors, 3D Training Object Manipulation received the lowest score. Upon analysis of user comments, a common concern was that the system seemed effective for the task presented, but its efficacy for more complex manual procedures encountered in daily work could not be evaluated. Since evaluating the system's adaptability to various training scenarios was a main goal of the study, further exploration of more complex 3D object manipulations is necessary in future prototypes. The scores can be seen in Fig. 3.

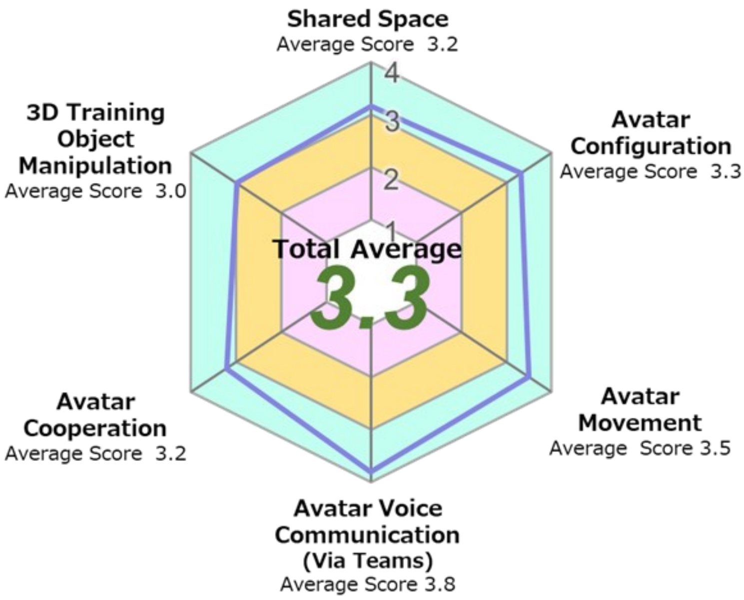

Fig. 3. Average scores of the 6 core factors and the total average obtained during the system evaluation.

4.2 Open Comments (Group Discussion)

We collected several comments during the post-test group discussions, but for the purposes of this brief paper, we present a summarized conclusion of the feedback received for each of the six factors:

a) *Shared Space:* The elements that led to general acceptance were the ability to see the Engineer's object of attention (i.e., gaze pointer), move around the scene to observe the procedures from a bird's-eye view, and getting close to the engineer without worrying about physical obstruction. On the downside, some comments questioned the necessity of using a VR headset, suggesting that they could observe the action using a 2D screen. In other words, VR mode may be fitting for certain roles more than others.

b) *Avatar Configuration:* Several users reported that half-body avatars were suitable for the task at hand. However, for more intricate work processes, a full-body avatar may be necessary (e.g., monitoring lower body movements may be crucial in avoiding certain dangers).

c) *Avatar Movement:* Despite being able to learn how to move in the VE in a short time, some users expressed concern about certain mechanics such as teleporting. This was because the subjects became disoriented for a moment, not knowing where the Engineer (i.e., target) was located. Better mechanics for navigating the VE, such as a pointing arrow gizmo towards the object of interest after teleporting may help in this regard.

d) Avatar Voice Communication: We used Teams as a temporary solution for testing voice communication. Nevertheless, several subjects expressed satisfaction with using a tool they were familiar with for long-distance communication.

e) Avatar Cooperation: Several users expressed satisfaction with the shared gaze mechanism, suggesting that adding extra markers to distinguish the Engineer's pointing action could help them understand the overall working procedures better. On the negative side, some subjects expressed concern about the difficulty in distinguishing the gaze pointer in the presence of complex and intricate equipment.

f) 3D Training Object Manipulation: Users were satisfied with the presented content. However, they also raised concerns once again about the effectiveness of manipulating 3D objects in the presence of more complex machinery scenarios. The need to test more complex scenarios became evident.

5 Conclusions and Future Work

In this initial iteration of our system, we obtained positive indicators that it is feasible to implement a multi-user BWAT solution with Web XR while achieving a high level of user acceptance in a real-world business environment.

In the design of the virtual environment (VE), we incorporated core human-computer interaction (HCI) concepts, along with shared gaze interaction, which enabled users to focus on the most relevant task. As a result, we expect that our initial findings will facilitate the development of the system in future iterations. These iterations may include more realistic work scenarios with even more complex tasks, a better graphical user interface (GUI), and inter-user visual cues for behavioral understanding, such as gaze and finger pointer gizmos, among other potential enhancements to promote the long-term adoption of our system. Furthermore, we recognize the importance of evaluating our system with existing standard tools to compare and validate its advantages relative to other options in the market. To that end, we plan to conduct follow-up experiments and incorporate standard usability evaluations, such as the System Usability Scale (SUS) questionnaire.

References

1. Li, H., Chan, G., Skitmore, M.: Multiuser Virtual Safety Training System for Tower Crane Dismantlement. 26, (2012). https://doi.org/10.1061/(ASCE)CP.1943-5487.0000170.
2. Andaluz, V.H., Sánchez, J.S., Sánchez, C.R., Quevedo, W.X., Varela, J., Morales, J.L., Cuzco, G.: Multi-user industrial training and education environment. 5, 533–546 (2018). https://doi.org/https://doi.org/10.1007/978-3-319-95282-6_38.
3. Doolani, S., Wessels, C., Kanal, V., Sevastopoulos, C., Jaiswal, A., Nambiappan, H.R., Makedon, F.: A Review of Extended Reality (XR) Technologies for Manufacturing Training. 8, (2020). https://doi.org/10.3390/TECHNOLOGIES8040077.
4. Billinghurst, M., Nebeling, M.: Rapid Prototyping of XR Experiences. In: CHI Conference on Human Factors in Computing Systems Extended Abstracts (CHI '21 Extended Abstracts). ACM, Yokohama Japan (2021).
5. Creating Custom Avatars · Hubs by Mozilla, https://hubs.mozilla.com/docs/index.html.
6. Rojas Ferrer, C.D., Fujiwara, T.: Work-in-progress—Avatar vs Teams: Co-Presence and Visual Behavior in a Videoconference Cooperative Task. In: 8th International Conference of the Immersive Learning Research Network (iLRN), pp. 1-3. IEEE, Vienna, Austria (2022). doi: 10.23919/iLRN55037.2022.9815938.
7. S. K. Dwivedi, C. D. Rojas Ferrer and T. Fujiwara.: Effectiveness of Avatar Guided Instructions and Feedback in the Server Repair and Maintenance Task. In: 8th International Conference of the Immersive Learning Research Network (iLRN), pp. 1-3. IEEE, Vienna, Austria (2022). doi: 10.23919/iLRN55037.2022.9815927.
8. Bertram, D.: Likert Scales …are the meaning of life. (2007).