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

DOI: https://doi.org/10.56198/ZH3JOA4QX

# Work-in-Progress—Effects of Attention Guidance on Virtual Reality Training for an Industrial Assembly Task

Joy Gisler[1], Dominik Dedic[1], Valentin Holzwarth[2], and Andreas Kunz[1]

[1] ETH Zurich, Zurich, Switzerland
[2] RhySearch, Buchs, Switzerland
gj@ethz.ch

**Abstract.** One of the main objectives of Virtual Training Environments (VTEs) for industrial training is to train workers for a real world task. Prior work identified a multitude of factors influencing a VTE's effectiveness. In this work-in-progress paper, we add to this body of research by evaluating the effect of attention guidance (AG) on a VTE's effectiveness. In a controlled between-subject design pilot study with 42 participants, participants were trained in a VTE either with or without AG. Subsequently, learning transfer was assessed in a Real-World Evaluation (RWE). Our findings indicate that, while not necessary for a VTE's efficacy, AG appears to be a substantial factor in a VTE's effectiveness.

**Keywords:** Virtual Reality, Attention Guidance, Training Simulator.

## 1 Introduction

Virtual Reality (VR) based immersive learning environments offer a multitude of benefits for industrial training. Industrial training, e.g. for assembly tasks or machine operation in VR reduces the risk of injury and avoids wasting raw material [1]. Furthermore, learning any industrial task in a Virtual Training Environment (VTE) reduces the reliance on instructors and training facilities [2]. One of the main objectives of any VTE is to maximize transfer of learning from the VTE to a corresponding real-world task. However, studies in the 1990s did not provide any evidence for transfer of learning to real-world tasks [3]. Since then, a multitude of factors influencing the efficacy of VTEs have been identified, e.g. trainees' motivation and prior knowledge [4] as well as more technical aspects such as the degree of immersion [5]. However, the quest for factors assuring efficacy (i.e. whether transfer of learning actually occurs) and increasing effectiveness of VTEs (i.e. how well a training supports learning transfer) is still ongoing [6]. Aiming to contribute to this ongoing research, we hypothesize attention guidance (AG) to be another addition to the long list of factors influencing VTEs' effectiveness.

AG can be defined as deliberately shifting the attention of a trainee towards a target object [7]. In any educational setting, capturing learners' attention is essential for any learning to occur [8], especially in non-mediated setting (i.e. in absence of instructors or teachers). While there is considerable work on guiding the attention of users in Mixed Reality (MR) and VR, the effect of AG on a VTE's effectiveness regarding transfer of learning to a real-world task is virtually unexplored. While previous work suggests that training effectiveness could be increased by adding some forms of guidance [9], we hypothesize that AG could also allow trainees to absent-mindedly follow the training procedure without dealing with the learning content, thus reducing learning transfer.

In this work, we address these concerns by developing two VTEs for an industrial assembly task either with or without AG. We conduct a pilot study with controlled between-subject study design, wherein each participant is randomly assigned to be either (A) trained in a VTE with AG, (B) trained in a VTE without AG or (C) not trained at all. Subsequently, the corresponding real-world assembly task is performed to assess training effectiveness. The chosen use case is the assembly of Modular Support Systems (MSS). MSSs are installed in commercial buildings to accommodate electrical, mechanical and plumbing components. The assembly of MSS involves procedural knowledge (e.g. executing the required steps in the correct sequence) as well as psychomotor skills (e.g. dexterity and muscle memory). Thus, our use case represents a broad variety of assembly tasks currently relevant in industry settings, e.g. construction sites or factories.

We organize the remainder of this paper as follows: In section 2, we outline the related work on AG in the context of industrial VTEs and beyond. Subsequently, in section 3 we introduce the industrial assembly task to be trained in the scope of this work. In section 4, we describe the methodology consisting of the VTE, the implemented AG, a description of our pilot study including the collected metrics, and the employed apparatus. In section 5, we present and discuss the results of our pilot study before concluding the paper and addressing potential directions for future research in section 0.

## 2 State of the Art

VTEs have been developed for various industrial contexts, such as safety [10], maintenance [11], or assembly [5]. However, research efforts to evaluate the effectiveness of VTEs still yield mixed results, concluding that training in VR is either better [12], worse [13], or comparable [14] to traditional training methods. These findings further highlight the currently incomplete understanding of factors influencing VTEs' effectiveness as described by [6].

Based on the well-documented relationship between attention and learning for real-world educational settings [8], we hypothesize that AG could be another factor influencing a VTE's effectiveness. Prior work has evaluated a decent amount of AG methods in many different fields of application such as cinematic VR, VR tours, and instructional videos [15]. The performance of AG methods is typically judged based on search time, i.e. how long it takes a user to find a target object. [16] compared the target-finding times of three AG methods (butterfly guide, fixed radar, and 3D arrow) and found that the 3D arrow performed best. In a comparison of different AG methods for an MR application, [17] report that highlighting the target object performs best when the target object is already in a user's Field of View (FOV), whereas an in-situ guiding line drawn from the middle of the FOV to the target object performs best for target objects outside a user's FOV. Despite the abundance of related work on AG methods, their effectiveness in terms of learning transfer are virtually unexplored [18]. In this work, we aim to address this gap and add to the body of knowledge concerning factors influencing the effectiveness of VTEs.

## 3 Use Case

The industrial task to be trained in VR is the assembly of a MSS. The MSS consists of two different types of extruded aluminum profiles, angular conjunctions, and so-called push buttons that connect and fix the other components. Additionally, a socket wrench is required to fix components in place. A push-buttons is a load-bearing connector, which facilitates the assembly. The working principle of a push-button however is supposedly not self-evident. The full assembly consist of 18 push-buttons, nine angular conjunctions (called *angles*), and seven unique extruded profiles (called *rails* in the remainder of this paper) of different lengths and/or type. The VTE representations of all the components necessary to assemble the MSS are shown in Fig. 1: Rails, the socket wrench, push-buttons, and angular conjunctions. The assembled MSS is shown in Fig. 2.

**Fig. 1** Components required to assemble the MSS: Rails (leftmost table), socket wrench (middle table), angles and push-buttons with instruction board above (rightmost table).

To successfully assemble the MSS, participants need to understand four distinct core concepts: (a) distinguishing the different extruded profiles, (b) identifying the extruded profiles' correct positions and orientations on a 2D layout, (c) understanding the working principle of push-buttons, and (d) the correct assembly sequence. For safety reasons, the rail that connects the MSS to the ceiling is already pre-installed in the correct position and orientation.

**Fig. 2** Fully assembled MSS in the VTE (left) and in reality (right).

# 4 Methodology

To address the identified research gap, we implemented a VTE for MSS assembly (see subsection 4.1). Next, we enhanced the VTE with AG methods (see subsection 4.2) that can be switched on or off for each training session. Finally, we conducted a pilot study (see subsection 4.3) to evaluate the effect of AG on learning transfer from a VTE to the corresponding real-world task, using the apparatus described in subsection 4.4.

## 4.1 Virtual Training Environment

The VTE was implemented in Unity (Version 2021.3.13). The virtual environment was adapted from the PBR Workshop asset from the Unity Asset Store [19]. Fig. 4 shows all the relevant areas of the VTE: A table containing the rails on the, the pre-installed rail to which the MSS has to be mounted in the center, and a table containing the angular conjunctions, push-buttons as well as a instruction board to the right. The instruction board contained assembly instructions and a 2D layout of the MSS. Furthermore, the VTE included pre-recorded sounds that were played accordingly e.g. when components were dropped or push-buttons mounted. The SteamVR Unity Plugin was included to allow for natural interaction with all the MSSs' components and to fully assemble it in the VTE. As in the real-world evaluation task of the pilot study, the uppermost rail of the MSS was pre-installed in the correct position and orientation. Furthermore, position indicators were added to the pre-installed rail in order to allow participants to correctly position angular connections without measuring distances. The same position indicators were also present in the real-world evaluation task of the pilot study. For every rail that needs to be added, the procedure consist of three steps: i.) loosely attaching the new rail to the already mounted components using an angular conjunction and a push-button, ii.) correctly positioning the newly added angular conjunction and rail, and iii.) tightening the newly added components at the correct position using the provided socket wrench. Once a newly added component was fixed with the socket wrench, the corresponding push-button turned green and could not be removed anymore. The fully assembled MSS in the VTE is shown in Fig. 2. Furthermore, a state machine (see Fig. 3) was implemented to keep track of the assembly sequence.

Our VTE was fully functional and realistic i.e. participants could connect any two components in any arrangement that was also feasible in reality. Consequently, this allowed components to be mounted incorrectly. We chose a fully functional approach in order to support the participants in building a understanding of the supposedly non-obvious working principle of the push-button.

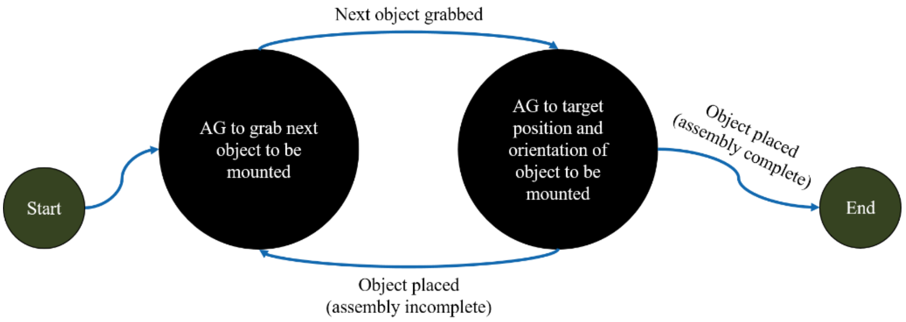

**Fig. 3** State machine.

**Fig. 4** Overview of the relevant areas of the VTE: Table with rails (to the left), assembly area with pre-installed rail (in the center), and table with angular conjunctions, push-buttons as well as an instruction board (to the right).

### 4.2 Attention Guidance

Based the findings of [17], three AG methods were implemented: i.) highlighting the target object (i.e. the next component to be mounted), ii.) an in-situ line originating from the center of the FOV going to either the next target object or a grabbed object's target position, and iii.) an indication of a grabbed component's target position and orientation. For every component (i.e. target object) that needs to be installed next, a trainee is first guided to the corresponding component by an in-situ line and a highlight on the target object (see Fig. 5, left). Once the trainee grabs the component, the in-situ line guides them to the component's target position where the correct orientation is indicated (see Fig. 5, right). As soon as the component is installed, the whole process starts again by guiding the trainee to the next component to be mounted. During the whole process, the implemented state machine keeps track and guides trainees through the assembly sequence.

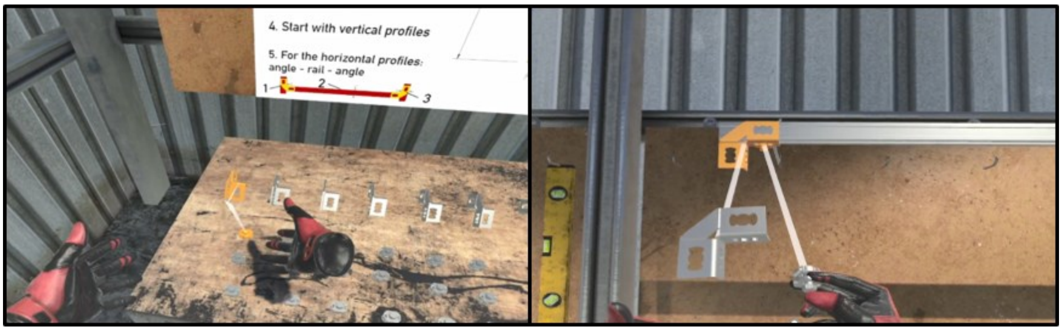

**Fig. 5** Implemented AG methods - Highlight and in-situ line from center of FOV to target object (left). Target orientation indication and in-situ line from center of FOV to target position (right).

### 4.3 Pilot Study

The study design of our pilot study consists of five distinct phases, shown in Fig. 6. In phases one, three, and five, we ask participants to answer questionnaires on their perception of the VTE, their subjective data, and their well-being. These questionnaires include questions on demographics, Manual Dexterity (MD, adapted from [20]), Simulator Sickness (SSQ) [21], Presence [22], Learner Satisfaction (LS, adapted from [23]), and NASA Task Load Index (TLX) [24]. The MD questionnaire yields a score $\in [0;7]$, where zero indicates very poor MD and seven indicates excellent MD. The SSQ is administered twice for each participant, once before and once after their VR session (phase two). The within-subject difference (ΔSSQ) between post- and pre-SSQ indicates how the VR session affected a participants well-being. The LS score indicates how satisfied participants are with their training and ranges from zero (not at all satisfied) to seen (very much satisfied).

In phases two and four, the research gap of this work is directly addressed. Phase two is conducted in VR and further divided into three separate conditions: (A) VR training with AG enabled, (B) VR training with AG disabled. Participants in the control group (C) receive no training and are instead playing an unrelated VR game for 15 minutes in phase two. The control group is introduced to detect a possibly occurring ceiling effect: Should the chosen assembly task be too trivial, training would be obsolete and task performance of participants belonging to condition (C) would perform similarly as those participants trained in either version of the implemented VTE. Participants are given a maximum of 15 minutes to complete phase two. Furthermore, we assess their Task Completion Time (TCT) and whether they are able to successfully assemble the MSS in the VTE.

In phase four, participants are asked to assemble the MSS in a real-world setting. In this Real-World Evaluation (RWE), participants are again given a maximum of 15 minutes to complete this task. Furthermore, their TCT and success are assessed. In phases three and five, successful task completion is a binary metric, i.e. either they are able to correctly assemble the MSS in the given time or not. In both phases, TCT is set to 15 Minutes if participants are not able to complete the assembly in the time given.

For the pilot study, we recruited 42 participants. Their Mean (M) age was 24.36 years with a Standard Deviation (SD) of 2.56 years. Participants were either students (39 participants) or people with desk jobs (three participants). People familiar with the use case and/or field workers were excluded from the study in order avoid biased results. The participants were randomly assigned to one of the three conditions.

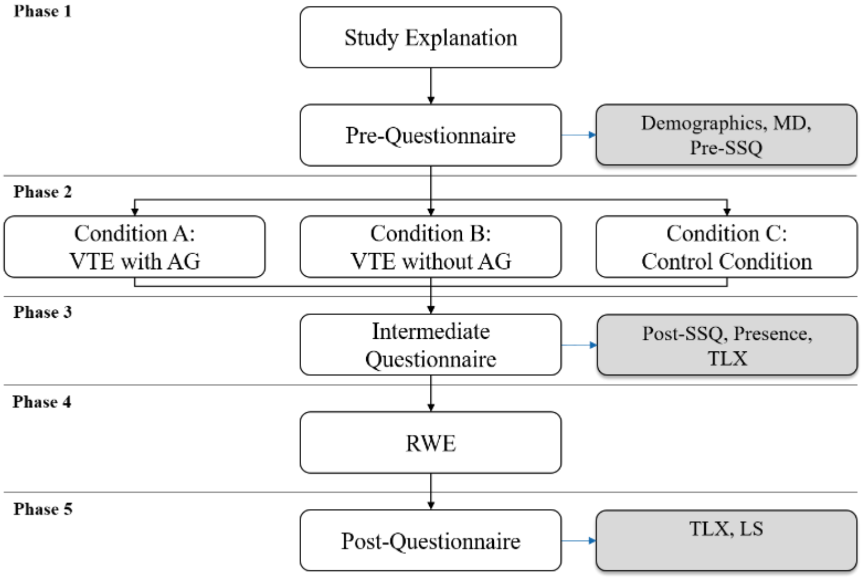

**Fig. 6** Study procedure.

### 4.4 Apparatus

The VTE was rendered at 90Hertz on a computer with an Nvidia RTX 2080, 32 gigabyte RAM and an Intel i7-9700K core. We used an HTC VIVE Pro head-mounted display with an HTC Wireless Adapter and four SteamVR base stations 2.0 to allow uninhibited real walking in a 5m x 6m room-scale tracking space. Two HTC VIVE Controllers 2.0 allowed for natural interaction (i.e. grab, rotate, carry, drop, manipulate, and place) with virtual objects. Furthermore, we used an ordinary laptop and Google Forms to collect participants' answers to the questionnaires.

## 5 Results and Discussion

The descriptive statistics of the ΔSSQ, MD questionnaire, Presence questionnaire and LS questionnaire are given in Table 1. The MD questionnaire did not yield any statistically significant differences between the three conditions. On average, the MD score of participants in condition (A) is merely 0.2 point higher than for participants in conditions (B) and (C). We therefore assume that our findings were not biased by an unequal distribution of MD across the three conditions.

The SSQ score of participants on average decreased by 2.5±15.8 points from phase 1 to phase 3. This indicates that the implemented VTE did not cause peculiar symptoms for simulator sickness. No participant had to be excluded due to simulator sickness. The mean Presence score of participants trained without AG with 5.4/7 is insignificantly higher than for participants trained with AG. This is expected, as the implemented AG methods represent an unnatural extension of the VTE.

**Table 1** Descriptive statistics (M±SD) of questionnaire scores: ΔSSQ, MD, Presence, and LS.

| Training Condition | n | ΔSSQ | MD | Presence | LS |
|---|---|---|---|---|---|
| Overall | 42 | -2.5±15.8 | 3.2±1.5 | 5.0±0.9 | 5.6±0.9 |
| AG Enabled (A) | 14 | -4.5±11.6 | 3.3±1.3 | 5.2±0.7 | 5.5±0.9 |
| AG Disabled (B) | 14 | 1.1±17.4 | 3.1±1.5 | 5.4±0.6 | 5.8±0.9 |
| No Training (C) | 14 | -4.3±18.2 | 3.1±1.7 | 4.5±1.2 | n.a. |

The TCTs and success rates for the VTE as well as for the RWE are given in Table 2. Five out of 14 participants in condition (B) – i.e. participants who trained in the VTE without AG – were able to successfully complete the assembly in the VTE, resulting in a success rate of 35.7%. For participants in condition (A), the success rate is twice as high at 71.4%, i.e. 10 out of 14 participants in this condition successfully assembled the MSS in the VTE. This difference between conditions (A) and (B) indicates that AG increases training success in a VTE. Furthermore, participants trained with AG completed the VR training significantly faster than participants trained without AG: In a one-tailed t-test with significance level $\alpha = 0.05$, the differences in TCT in the VTE between participants trained with and without AG exhibited statistical significance ($p = 0.029$).

In the RWE, all participants trained in the VTE with AG successfully assembled the VTE, whereas only 71.4% of participants trained without AG were successful. This indicates that AG is not a necessary factor for a VTE's efficacy regarding learning transfer. However, the success rate of 100% is a strong indicator that AG positively affects the effectiveness of a VTE. Based on these findings, the hypothesis that AG reduces learning transfer due to trainees blindly following instructions does not seem plausible. Further, participants trained with AG assembled the MSS significantly faster ($p = 0.010$) in the RWE. A Whisker plot of the TCTs in the RWE for all three conditions is shown in Fig. 7. Contrasting these objective findings, participants trained without AG on average reported insignificantly higher LS than participants trained with AG (see Table 1). LS, TCT in the VTE and success rate in the VTE were not assessed (n.a.) for participants in condition (C), since they did not receive any training.

Limitations of our findings lie in the chosen measure for success: Attempts which are classified as failed include both participants who just missed the last step as well as participants who did not manage a single step.

**Table 2** TCTs (M±SD) and success rates for the VTE as well as the RWE.

| Training Condition | n | TCT VTE | Success Rate VTE | TCT RWE | Success Rate RWE |
|---|---|---|---|---|---|
| AG Enabled (A) | 14 | 12.7±2.6min | 71.4% | 9.5±2.0min | 100.0% |
| AG Disabled (B) | 14 | 14.2±1.4min | 35.7% | 11.7±2.5min | 71.4% |
| No Training (C) | 14 | n.a. | n.a. | 15.0±0.0min | 00.0% |

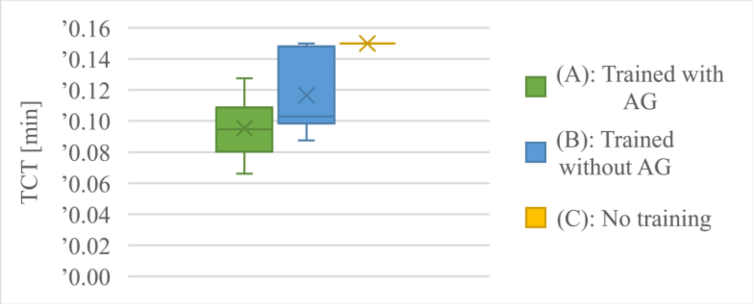

**Fig. 7** Whisker plot of TCTs in the RWE for the three conditions (A), (B), and (C).

## 6 Conclusion and Future Work

In this work, we investigated the effects of AG on the efficacy and effectiveness of a VTE for an industrial assembly task. Our results indicate that, while AG does not seem to be a necessary factor for a VTE’s efficacy, it does nevertheless influence a VTE’s effectiveness. I.e. participants trained in a VTE with AG were significantly faster and had a higher probability of successfully completing the training. The same holds true for the RWE: Participants trained in the VTE with AG were significantly faster in completing the trained task in a real-world setting. Furthermore, all of the participants receiving training in a VTE with AG were subsequently able to complete the corresponding real-world assembly task. For participants trained in a VTE without AG, the success rate was only 71.4%. Based on these findings, we conclude that AG not only reduces training time while increasing success rate in a VTE but also positively affects TCT and task performance in a corresponding real-world task. Further, our findings clearly suggest that AG does not reduce learning transfer by allowing trainees to blindly follow the instructions.

Future work should investigate the external validity of our findings, and compare the AG methods implemented in this work to other AG methods. To overcome the limitations imposed by our binary metric of transfer learning, future work could develop a more nuanced method for measuring success.

## References

1. Gisler, J., Holzwarth, V., Hirt, C., Kunz, A.: Work-in-Progress–Enhancing Training in Virtual Reality with Hand Tracking and a Real Tool. In: 2021 7th International Conference of the Immersive Learning Research Network (iLRN), pp. 1–3. IEEE, Eureka (2021).
2. Kalkan, Ö.K., Karabulut, Ş, Höke, G.: Effect of Virtual Reality-Based Training on Complex Industrial Assembly Task Performance. Arabian Journal for Science and Engineering 46(12), 12697–12708 (2021)
3. Kozak, J.J., Hancock, P.A., Arthur, E.J., Chrysler, S.T.: Transfer of training from virtual reality. Ergonomics 36(7), 777–784 (1993)
4. Holton, E.F.: Holton’s Evaluation Model: New Evidence and Construct Elaborations. Advances in Developing Human Resources 7(1), 37–54 (2005)
5. Gisler, J., Hift, C., Kunz, A., Holzwarth, V.: Designing Virtual Training Environments: Does Immersion increase Task Performance? In: 2020 International Conference on Cyberworlds (CW), pp. 125–128. IEEE, Caen (2020).
6. Dengel, A.: What Is Immersive Learning? In: 2022 8th International Conference of the Immersive Learning Research Network (iLRN), pp. 1–5. IEEE, Vienna (2022).
7. Grogorick, S., Magnor, M.: Subtle visual attention guidance in vr. Real VR–Immersive Digital Reality: How to Import the Real World into Head-Mounted Immersive Displays, 272–284 (2020)
8. Lodge, J.M., Harrison, W.J.: Focus: Attention science: The role of attention in learning in the digital age. The Yale journal of biology and medicine 92(1), 21(2019)
9. Hoareau, C., Querrec, R., Buche, C., Ganier, F.: Evaluation of internal and external validity of a virtual environment for learning a long procedure. International Journal of Human–Computer Interaction 33(10), 786–798 (2017)
10. Buttussi, F., Chittaro, L.: Effects of Different Types of Virtual Reality Display on Presence and Learning in a Safety Training Scenario. IEEE Transactions on Visualization and Computer Graphics 24(2), 1063–1076 (2018)
11. Ganier, F., Hoareau, C., Tisseau, J.: Evaluation of procedural learning transfer from a virtual environment to a real situation: a case study on tank maintenance training. Ergonomics 57(6), 828–843 (2014)
12. Ho, N., Wong, P.-M., Chua, M., Chui, C.-K.: Virtual reality training for assembly of hybrid medical devices. Multimedia Tools and Applications 77(23), 30651–30682 (2018)
13. Makransky, G., Terkildsen, T.S., Mayer, R.E.: Adding immersive virtual reality to a science lab simulation causes more presence but less learning. Learning and Instruction 60(December 2017), 225–236 (2019)

14. Rose, F.D., Attree, E.A., Brooks, B.M., Parslow, D.M., Penn, P.R.: Training in virtual environments: transfer to real world tasks and equivalence to real task training. Ergonomics 43(4), 494–511 (2000)
15. Nonino, E., Gisler, J., Holzwarth, V., Hirt, C., Kunz, A.: Subtle attention guidance for real walking in virtual environments. In: 2021 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), pp. 310–315 (2021). IEEE
16. Wallgrün, J.O., Bagher, M.M., Sajjadi, P., Klippel, A.: A comparison of visual attention guiding approaches for 360 image-based vr tours. In: 2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), pp. 83–91 (2020). IEEE
17. Renner, P., Pfeiffer, T.: Evaluation of Attention Guiding Techniques for Augmented Reality-based Assistance in Picking and Assembly Tasks. In: Proceedings of the 22nd International Conference on Intelligent User Interfaces Companion, pp. 89–92. ACM, New York, NY, USA (2017).
18. Prattico, F.G., De Lorenzis, F., Lamberti, F.: Look at It ThisWay: A Comparison of Metaphors for Directing the User's Gaze in eXtended Reality Training Systems. In: 2021 7th International Conference of the Immersive Learning Research Network (iLRN), pp. 1–8. IEEE, Eureka (2021).
19. Unity Asset Store: "PRB Workshop." Accessed Dec. 02, 2022.
20. Francoise, P.: Dexterity as a valid measure of hand function. Occupational Therapy In Health Care 4(3), 69–83 (1988)
21. Kennedy, R.S., Lane, N.E., Berbaum, K.S., Lilienthal, M.G.: Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. The international journal of aviation psychology 3(3), 203–220 (1993)
22. Slater, M., Usoh, M., Steed, A.: Depth of presence in virtual environments. Presence: Teleoperators & Virtual Environments 3(2), 130–144 (1994)
23. Wan, Z., Compeau, D., Haggerty, N.: The effects of self-regulated learning process on e-learning outcomes in organizational settings. Journal of Management Information Systems 29(1), 307–340 (2012)
24. Hart, S.G., Staveland, L.E.: Development of nasa-tlx (task load index): Results of empirical and theoretical research. In: Advances in Psychology vol. 52, pp. 139–183. Elsevier, North-Holland (1988)