An Evaluation of Virtual Reality Maintenance Training for Industrial Hydraulic Machines

Thuong Hoang[1], Stefan Greuter[2] and Simeon Taylor[2]
[1]School of Information Technology, [2]School of Communication & Creative Arts / Deakin Motion Lab
Deakin University, Geelong, Australia

2022 IEEE on Conference Virtual Reality and 3D User Interfaces (VR) | 978-1-6654-9617-9/22/$31.00 ©2022 IEEE | DOI: 10.1109/VR51125.2022.00077

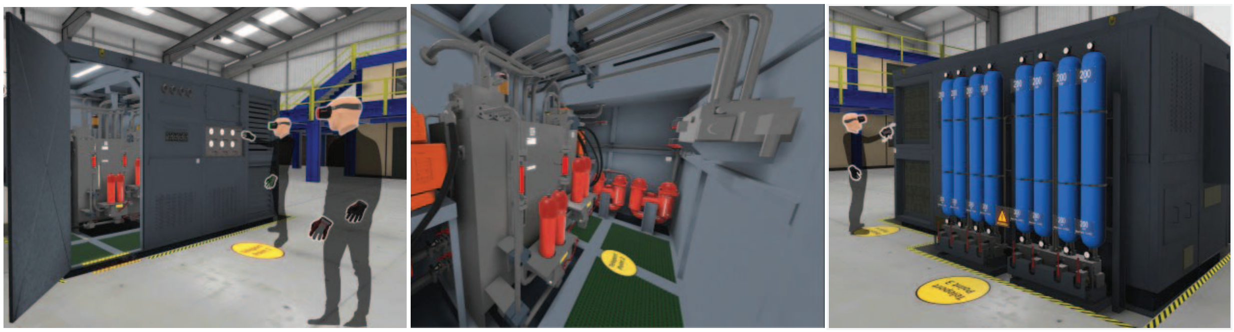

Fig 1. The virtual hydraulic power unit inside a virtual warehouse: Left: front view with doorway entrance; Middle: inside view with red hydraulic pressure filters units; Right: accumulator bladders at the back of the unit

Abstract

Virtual reality applications for industrial training have widespread benefits for simulating various scenarios and conditions. We present an empirical evaluation of VR training approach using kinesthetics learning strategy in industrial maintenance training, specifically the hydraulic manufacturing industry. Through our collaboration with a leading industry partner, a remote multi-user training platform using head-mounted display was designed and implemented. We present the evaluation of the platform with two diverse cohorts of novice users and industry contractors, in comparison to traditional training using slides, photos, and videos. The results show that VR training is engaging and effective in boosting trainee's confidence, especially for novice users. Our studies highlight the impact of virtual reality training on trainee experience, performance, and skills transfer, with reflections on the differences between novice and industry trainees.

Keywords: Virtual Reality, Industrial Training, Multimedia Learning in Virtual Reality, Kinesthetics Learning

Index Terms: [Human-centred Computing]: Human Computer Interaction – Empirical studies in HCI

1 Introduction

Virtual reality has widely been recognised as a beneficial platform for education and training, especially in an industrial setting. A scoping review by Naranjo et al. [1] outlines the benefits of VR training as reducing training time, preventing errors leading to improved product quality in a diverse range of industrial domains, including automobile, steel, and oil and gas. Traditional training in an industrial setting often uses slideshows or demonstration to convey procedural or task process training. Multimedia training theory [2] explains the benefits of VR training through the usage of interactive multimedia. Prior works apply different learning strategies in VR, including summarising [3], practice testing [4], and enactment [5], and compare VR to traditional methods of training, such as slideshows or videos. The findings lean towards higher enjoyment and engagement, but not better learning outcomes. Moreover, VR training could introduce significant cognitive load and distraction [4], [6] in multiple training domains and scenarios.

In this paper, we are motivated to employ a metacognitive strategy of kinesthetics learning [7], or 'learning by doing', for VR training in an industrial training setting, specifically hydraulic manufacturing industry. We have established a successful collaboration with an industrial partner, whose name is coded as HI, as the national branch of an international corporation specialised in manufacturing fluid power products, typically hydraulic systems used in multiple industries, such as defence, mining, or constructions. HI offers nationally accredited training courses covering basic hydraulic principles and components, as well as practical maintenance courses on assessment, testing, repair and practical maintenance courses on assessment, testing, repair, and replacement of faulty components. Our industry partner in hydraulic machines, HI, aims to implement a training solution using virtual reality for diverse cohorts of trainees, including maintenance technicians, trades and sales, and junior engineers. We are motivated to conduct a cost-benefit analysis of VR training solutions regarding existing training methods offered at HI. Currently, HI conducts presentations and hands-on training with components. In this paper, we conducted a comparative evaluation of training methods for maintenance tasks of a large industrial hydraulic systems. We create a virtual digital twin of a room-scale wheel lock hydraulic power unit (HPU) designed by a leading international manufacturer for the iron ore industry and build an HMD based VR training platform to perform three maintenance tasks. Our evaluation compares the VR training system with traditional training using a paper checklist and classroom-style presentations.

We present the parallel empirical media comparison of traditional and virtual reality training for industrial maintenance tasks in two separate cohorts with different backgrounds: novice users and industrial contractors with professional experience. We recruited two cohorts of participants representing the typical

{thuong.hoang,stefan.greuter,simeon.taylor}@deakin.edu.au

978-1-6654-9617-9/22/$31.00 ©2022 IEEE
DOI 10.1109/VR51125.2022.00077

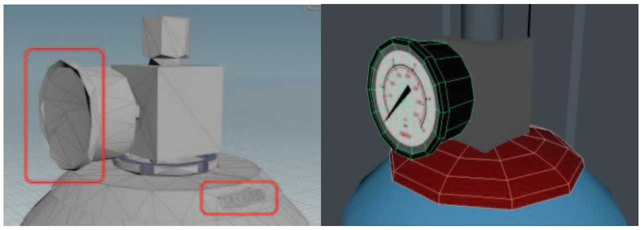

Fig 2. Optimised 3D models used in VR. Left: Houdini result, note the unnecessary geometry (in red rectangles) on the label on the side and markings on the pressure gauge. Right: Remodelled in Maya with efficient geometry

groups of trainees at HI: university students with varying levels of technical skills and professional contractors with prior experience working on similar hydraulic components. We applied two study designs to best suit both cohorts' skill, experience, and time commitment. We also reflected on the differences in training experiences, skill transfer, and assessment feedback. Based on our findings, we present our reflections on the impact of virtual reality training platform on novice users and professional contractors, as guidelines for future designers, researchers, and stakeholders.

Our work contributes to the body of work of VR learning in understanding the impact of kinesthetics learning strategy in VR industrial training versus traditional methods. Our study offers unique insights into the differences in training experiences and performances by both cohorts, novice and expert trainees. Our comparison differs from prior work by deploying genuine training scenarios through a partnership with an international hydraulic manufacturing corporation with validated hands-on assessment in an authentic industrial factory.

2 Related Work

Virtual reality training for the industrial domain has been a proliferating direction of research. In this section, we explore current theoretical standpoint on VR training, as well as relevant prior work in VR training in industrial settings. We also discuss the impact of VR learning on novice and expert learners.

2.1 Multimedia Learning in Immersive Systems

Prior studies [8], [9] have shown successful procedural knowledge transfer comparable to traditional training. Parong and Mayer [3] explain the benefits of VR training as supported by the theory of multimedia learning, in which the learners are presented with interactive animations in the virtual world. They conducted a media comparison study between VR and slideshows for a biology lesson and found that VR provides higher enjoyment, engagement, and motivation, but suffers lower post-test score for knowledge retention. The authors further suggested the inclusion of generative learning theory in VR by introducing meaningful cognitive processing activities during learning, such as selecting, organising, and integrating knowledge. They conducted a similar media comparison study in VR with and without summarising task, and the results showed that the inclusion of cognitive processing strategy of summarising increased learning outcome.

Parong and Mayer [4] later applied a different generative learning strategy of practice testing (frequent testing for recall) in VR and did not find any significant result for knowledge retention, despite higher enjoyment with VR learning. The authors further concluded that VR learning contributed to cognitive distraction with higher workload due to the immersive nature of the instructional medium, leading to decreased performance on learning outcome. Makransky et al. [5] employed another learning strategy of enactment, in which the learners replicated procedural knowledge learnt with physical objects, and performed two experiments to understand the impact of VR learning versus video instructions, with and without enactment strategy. Their findings showed that enactment contributed to higher procedural knowledge and transfer for VR training, but not in video instructions. VR learning induced higher enjoyment but not more effective than video training.

Our work contributes to multimedia learning in VR theory by employing the metacognitive strategy of 'learning by doing' [7], or kinesthetics learning in an industrial training setting. Kinesthetics learning in VR has been applied in other domains, such as language learning [10], remote higher education [11], and lab-based science subjects [12]. Prior studies show that kinesthetics approach in VR learning is more efficient than traditional methods [11], even though there may not be immediate gain but often leading to better retention rate [10]. *Our study contributes empirical evidence for kinesthetics VR training for both novice and expert trainees in industrial training domain.*

2.2 Virtual Reality for Industrial Training

A scoping review by Naranjo et al. [1] surveyed papers published from 2015 to 2020 and found a total of 44 articles that addresses the multiple aspects of VR usage for industrial training. The review expanded to VR for non-industrial training and industrial training methods that are not VR for comparison. Naranjo et al. found that most studies for industrial training are limited in task complexity, comparisons, and measured variables. The range of tasks includes manual and simulated assembly [9], [13], [14], remote robotic operations [15], power distribution [16].

Pérez et al. [15] combined a virtual simulation environment with a robotic arm to build a training solution for robotic operations. The authors performed an evaluation with 12 robotic domain experts to rate their preference through a Likert scale. They concluded that the experts enjoyed the realism with usable and friendly virtual interfaces. Lacko [17] built a system to promote safety practice for stairs at a worksite through a first-person view. The author performed a comparison study with 37 users in traditional video training and 32 users using VR, to show that on average, participants who received VR training retain higher scores on health and safety tests one month after the training than participants trained only via videos. The data was based on post-training test scores. These studies represent a body of work that evaluate training effectiveness with only expert users, typically via a self-rated questionnaire. Another industrial safety training application was built by Hoang et al. [18] as a virtual construction site containing workplace scenarios that triggered accidents, such as falling, crushing or working with faulty tools. The authors combined fear-arousal with experiential VR to deliver impactful safety messages to contractors and office workers of a construction company. Their evaluation confirmed the effectiveness of improving safety training attitudes among the workers. *Our study complements prior work with the addition of VR use cases in maintenance tasks for industrial hydraulic machines. Unlike prior work discussed in the scoping review [1], our study employed genuine tasks with validation on actual hydraulic machines, as well as applying well-validate measures for training experience, satisfaction and self-confidence* [19]*, as well as educational practice questionnaires* [17].

2.3 Novice and Expert Learners in VR

Other work compares the impact of VR industrial training on both expert users and novices. Mirauda et al. [20] developed a virtual laboratory for simulating fieldwork in hydraulic engineering education. The authors evaluated the system through Likert questionnaires on system quality, interactivity, and training performance, against the traditional approach with 24 participants from a mixed technical and non-technical background. The outcome showed that virtual tools could effectively assist all types of users. Roldán et al. [14] introduced a VR-based process mining system that translates expert knowledge to train novice users through an example assembly domain. Process mining applies data analysis techniques to model, evaluate and optimise processes, which is, in this case, a sequence of assembly tasks. Expert users built a virtual assembly made up of coloured virtual boxes based on their expertise, which is then captured through processing mining and presented as virtual guidance for a novice user to reconstruct. Time and error performance data, NASA-TLX workload [21], subjective evaluation, perception and learning facility questionnaire was collected from twenty university students. The analysis found that immersive training has a significant advantage over traditional training on most scores. Winther et al. [22] conducted a lab-based comparison study between VR and traditional training methods (video and pairwise) for 36 novices with the task of replacing a rubber ring on an industrial pump. The results showed that traditional training method still provided better outcomes in terms of completion time and number of errors. *Prior work explored novice and expert trainees separately, while our study combines the two cohorts under similar evaluation protocols to offer insights into the differences in VR training for novice and expert trainees.*

3 System Design

During our collaboration, HI was in the process of delivering a room-scale wheel lock hydraulic power unit for an iron ore client. The HPU was built to power an array of hydraulic cylinders that lock onto carriage trains so their iron ore contents can be dumped safely onto a stacker reclaimer for storage. The unit was measured at 4.8 x 2.1 x 3m, with a large door for maintenance workers to access the various components inside. Maintenance procedures for the HPU included tasks conducted both inside and outside the unit, such as replacing an oil filter in a pressurised hydraulic unit, testing accumulator pressureand cleaning air vents of cooling units. Our VR training system aimed to enable maintenance tasks on a virtual digital twin of the HPU.

3.1 Digital Twin of Hydraulic Power Unit

HI provided a detailed virtual model of the HPU. We performed model simplification to enable real-time rendering performance on our target mobile VR headset, the Oculus Quest 1 and 2. First, we optimised the model in SolidWorks and export polygon models at a lower level of detail. Second, we ran the polygon through Houdini[1], as a 3D animation software commonly used for visual effects creation in film and games, to further optimise the model and reduce the polygon count. Third, we manually repaired model simplification faults to simplify the polygon topology further using Maya. Fourth, we organised the model into static and movable objects and map UV coordinates for surfaces that require textured surfaces. Lastly, we developed textures and bump maps to maximise the visual appearance. Through this process, we reduced the polygon complexity from 15-million polygons to 1.5-million polygons. Fig. 2 shows the results of automatic and manual polygon reduction. The second stage of creating the digital twin was to isolate components in the model to enable interactions. This process was manually completed for approximately 30 interactive machine elements, including valves, doors, and gauges.

3.2 Virtual Maintenance Scenarios

Based on the digital twin of the HPU, we implemented training scenarios for three maintenance tasks, including changing the pressure filter, offline filter, and checking the accumulator pressure. Two filter tasks were conducted inside the HPU container while the accumulator task was performed on the outside (see Fig 1 middle and right).

Our training scenarios were based on three maintenance tasks for the wheel lock HPU: changing pressure and offline filters and pressurising the accumulators. Hydraulic filters remove contaminants in the hydraulic oil and require changing periodically. For the purpose of the VR training solution, there were two types of filter elements inside the HPU: pressure filter and offline filter, housed within steel cylinders. The filter changing process involved 1) Identify the correct filter, 2) De-pressurise the filter, 3) Drain the oil, 4) Open the filter cap, 5) Replace the filter element, 6) Seal the filter, and 7) Clean up. Offline and pressure filters had different ways to complete step 4: A drill was needed to remove 4 bolts on the cap; for pressure filters, a spanner is needed to unscrew the cap. The locations of the oil drain in step 3 were also different.

The standard HI training program for these three tasks provided a printed checklist with a textual description of the actions and typically a photo of the target element, as shown in Fig. 3. The picture focused directly on the element and occasionally depicted the required action to be performed; however, it did not offer instruction on where to locate the element within the HPU.

3.3 Virtual Checklist Interface and Interactions

We created a VR training platform for these three maintenance tasks using Unity game engine with multi-user support through Photon Unity Networking package[2]. The VR platform supported up to 20 participants remotely connected, each represented with a floating head and hands avatar, without a torso. A custom wrist interface enabled participants to input their name, displayed as a floating label atop their head avatar. The wrist interface also enabled access to a laser pointer for teleportation and virtual tools such as a spanner, a hand drill, a dust spray, and an oil drain container. Voice communication was enabled with spatial sound so that users can only hear one another if they are nearby on the virtual site. The virtual environment was set in a warehouse with the digital twin HPU set at the centre.

We built training scenarios with a virtual checklist interface for each of the three tasks. Virtual QR codes are placed at 3 locations inside and outside the virtual HPU, which can be activated when the trainee looks at or touch the QR code. The usage of QR codes enabled future implementation in an augmented reality solution. Once activated, a floating UI panel appeared showing a textual checklist on semi-transparent background outlining the steps (approximately 20 steps for each task) with a checkbox for completion and an accompanying video with voice narration of how to complete the described task on the digital twin, as seen in Fig. 3. The floating panel could be grabbed and repositioned anywhere within the virtual environment. Tilt viewing angle could also be adjusted via direct manipulation. The trainee accessed the wrist interface to obtain the correct tool, if required, and completes the task in VR as per textual and video instructions. The videos captured actions in the virtual environment using digital twin versions of the machine. Upon completion, the trainee could touch the checkbox next to each step and the UI advances to the next instruction.

We implemented kinesthetics learning by providing realistic interactions with virtual tools through the wrist interface. When activated, each virtual tool (drill, spanner, oil container, and laser pointer) was rendered as a floating 3D model in front of the trainee. Using the Oculus controller, trainee would reach out to grab the tool by pressing down the trigger. Interaction with the

[1] https://www.sidefx.com/products/houdini/

[2] https://www.photonengine.com/pun

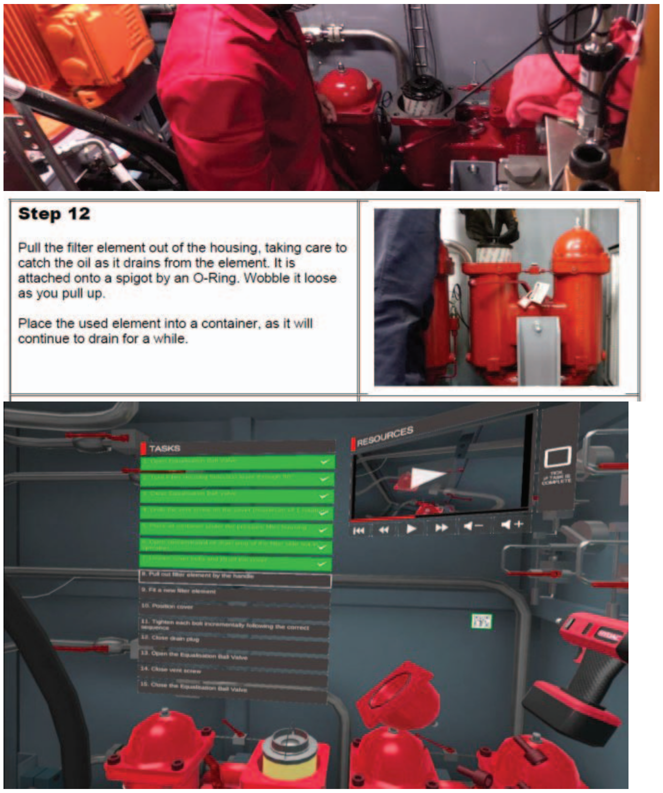

Fig 3. Top: Participant completing assessment inside the HPU Middle: the corresponding instruction in traditional training; Bottom: view of the corresponding step within VR with floating checklist, tools, and open filter cap

tool mimicked similar real-life actions when using physical tools. For example, when using the spanner, the trainee would grab the spanner, place the tip of the spanner to the correct virtual bolt or nut, and motion their hand back and forth while holding down the trigger on the controller and maintaining contact between the spanner and the component. Sound effects for the interactions were also added.

4 EVALUATION

We conducted user studies with two cohorts: university students (*Student)* and professional contractors (*Industry*), to understand the impact of VR training on the diverse range of target demographics, compared to traditional training methods of written instructions. Both studies received ethics approval from our university ethics committee. Our research question was "*How does kinesthetics VR training impact on the training experience and performance of novice and expert users?'*

Our studies compared two conditions: *VR* and *Traditional.* The former uses our VR training platform as described, and the latter follows current training methods at HI through classroom style presentation and written instructions, as shown in Fig. 2. Prior to the studies, we organised consultations with the technical training team at HI to discuss traditional training arrangements. The training team prepared a classroom-style presentation that introduced hydraulic machines, their operating environments, risks and safety issues during maintenance, and detailed instructions for the three maintenance tasks. The team used slides, photos, close-up video of maintenance operations and physical machine components for demonstration.

For both cohorts, each participant received training for either or both conditions, and completed an assessment by performing the maintenance tasks on the physical HPU on the factory floor of HI, under the supervision of a technical training technician.

We conducted the study on-site at HI factory over 4 non-consecutive days, which was scheduled after HI has completed building the wheel lock HPU and before it was shipped to the iron ore client. The two cohorts of participants were divided into 8 sessions, including 2 sessions for Traditional Students, 2 sessions for VR Students, and 4 sessions for Industry participants. The participants arrived as a group for each session at the HI factory. Upon arrival, the participants were signed in as visitors to the factory and provided plain language statement and consent forms. Each participant completed the onboarding questionnaires, including demographics, and learning styles (for student participants only).

4.1 Student Evaluation

We applied a between-subject design for Student cohorts.

4.1.1 Student Participants

We recruited 23 participants (7 females, 16 males) from our university. The participants self-rated their technical experience with a total of 8 novices, 7 somewhat experienced to high experienced and 1 expert. The majority of the participants were in the 18-24 years of age bracket (14 participants), with 7 participants in the 3 age brackets from 25 to 45, and 2 participants aged 50 years or above. The participants were divided into two groups: *Traditional* (12 participants) and *VR* (11 participants). Each group received the corresponding types of training (Traditional or VR) for 3 maintenance tasks on the machine.

4.1.2 Student Measures

As our *Student* cohort received a full version of the training, we aimed to capture their learning experiences. In addition to demographics information (age group, gender, VR experience, and industry), we asked our student participants to complete a VARK learning style questionnaire with 16 statements [23] before the training of either traditional or VR condition. The VARK questionnaire measures perceptual preference for the information delivery mode, including visual (charts, graphs or symbolic devices), aural (spoken lessons and talking), read/write (printed text), and kinaesthetic (learning through direct practice). Each statement presented a learning scenario with 4 answers indicating different learning approaches. A scoring sheet was used to allocate score for each of the 4 learning styles to find the dominant learning style (highest score) for the participant.

We asked participants to complete a survey on their training experience and environment (*TX*), satisfaction and self-confidence in learning (*SSCL*), as well as educational practices questionnaire (*EPQ*) [19]. The participants were presented with statements and used a 5-point Likert scale to indicate how much they agree with each statement (Strongly [dis]agree, somewhat [dis]agree, and neutral). The TX component contained statements related to training objectives, interactions, content, trainer, time and facilities, adapted from the workplace learning evaluation framework [24]. The student satisfaction component contained 5 statements of satisfaction, teaching methods, diversity of learning materials, facilitation, motivation and suitability. The SSCL subscale [19] contained 6 items measuring self-confidence in contents mastery, necessity, skill development, available resources and problem-solving knowledge. The EPQ subscale [19] contained 10 statements measuring opportunities for active learning and participation in training. Scoring for both SSCL and EPQ subscales [19] included summing the responses of the Likert statements ranging from 1 point for Strongly disagree to 5 points for Strongly agree. Higher SSCL score corresponded to more satisfaction and more self-confidence. Higher EPQ score indicated increased recognition of educational best practices. There was no summative scoring for the TX questionnaire.

We also asked our Student participants to complete a non-weighted version of NASA-TLX [21] in the VR condition, to understand impact of cognitive workload which could have been cause by the VR medium itself. NASA TLX measures the level of demand or workload in six subscales: mental, physical, temporal, performance, effort, and frustration. This was a standard practice for immersive VR training [4].

After survey collection, we asked Student participants to perform the maintenance task inside the physical wheel lock HPU under the supervision of a HI trainer based on printed checklists for the task (different from the printed training material as it does not contain photo instructions or guidance). The trainer provided a numeric score to measure task performance from 0 to 10 on the participant's performance in terms of overall performance, confidence, time, cleanliness, and safe work practice. The student participants performed one maintenance task as they were provided only one instructional condition (VR or Traditional, between subject). The participants were also asked to self-rate their performance on completing the task in the real-life HPU. The rating factors included Task Completion, Difficulty, Efficiency and Satisfaction. The participants were also asked if they would recommend the training provided to others. Each of these rating factors were provided with 5-point Likert scales.

We asked the participants in the VR condition (half of the student cohort) to rate their sense of presence [25] and sense of embodiment [26] in the virtual environment through a visual analogue scale (0-100). We also conducted a group interview at the end of the session, where we asked participants to reflect on: 1) how the training has prepared them for the tasks, in terms of expectations, confidence and know-how; 2) their confidence level during and after completing the assessment on the actual machine; 3) suggestions on how the training for either or both conditions could be improved. The meeting was voice recorded for transcription and analysis.

4.1.3 Student Procedure

For student participants assigned to traditional training, a group training session with the HI trainer was conducted for one hour using presentation slides, photos, and sample components. The training covered all 3 maintenance tasks (pressure filter, offline filter, and accumulator). Three VR stations were set up for student participants assigned to VR, each with one research assistant to guide up to a pair of participants through the 3 maintenance task scenarios in VR. For each station, 2 participants attended the training together, where each took turn performing the task in VR as instructed by the research assistant. The training was also conducted within one hour.

Student participants were randomly assigned one of the 3 tasks to perform on the physical wheel lock machine on the factory floor, under the supervision of a HI training personnel. Each participant had 20 minutes to complete the task and was provided with a written checklist (without photos nor further instructions). Figure 3 shows a comparison view during assessment (top), traditional training checklist (middle), and VR view (bottom) of the same step in replacing the pressure filter. The figure in VR noted floating tools and filter cap as we enabled such features for ease of tool placement in VR.

After training and assessment were completed, the participants were seated inside a training room to complete the questionnaires, including TX, SSCL and EPQ, self-rated performance, NASA-TLX and presence and embodiment. The questionnaires were built using Qualtrics, and a QR code was provided for participants to complete on their mobile phones or a supplied mobile or tablet device. A group interview was completed with all participants from each session, Overall, student participants spent around 3 hours with the study.

4.2 Industry Evaluation

We applied a within-subject design for industry, due to the availability of the participants in this cohort. The technical training manager at HI suggested that the standard training material would be redundant for the industry cohort. While they have not seen nor worked with the wheel lock HPU before, the industry participants are professional contractors whose daily job involves working directly with hydraulic components. Therefore, we adopted a different traditional training condition for the *Industry* participants where they were provided with printed sheets of task instructions, similar to Fig. 2, without the classroom presentation as in the *Student* cohort.

4.2.1 Industry Participants

Our industry partner HI assisted in recruiting 12 *Industry* participants (all males) with extensive professional experience with hydraulic and engineering systems. Each industry participants received training for both conditions, Traditional and VR, with randomisation of the 2 maintenance tasks (Pressure and Offline filters). After receiving the training, the participants were asked to perform the task on real-life machine under the supervision of training personnel from HI, who observed and provided the numeric score for the participants' performance.

4.2.2 Industry Measures

For the *Industry* participants, we collected demographics information, including age group, gender, VR experience, and years of industrial experience. As the industry participants only received a written task list for the Traditional condition, we only collected the training experience survey (TX) after they have finished training with each condition.

After survey collection, we asked the Industry participants to complete the maintenance task inside the physical wheel lock HPU under the supervision and scoring of a HI trainer. The procedure was similar to the Student cohort as described in 4.1.2. The main difference was that the industry participants performed two maintenance tasks as they experienced both training conditions (within-subject), as compared to one maintenance task performed by the Student participants. The Industry participants were also asked to self-rate their performance similarly to Student participants. We collected sense of presence and embodiment questionnaire from all the Industry participants similarly to the Student participants.

4.2.3 Industry Procedure

Due to time limitation and higher existing experience with similar systems, the training component was significantly reduced for industry participants. Each industry participant received both VR and Traditional training. Each condition only covered one of two tasks (pressure filter or offline filter). We applied Latin square on 2 conditions (VR, Traditional) and 2 tasks (pressure or offline filter) to balance and assign to each participant. For the traditional condition, the industry participants were provided with a written task checklist, with photos instructions to study for 15 minutes, without a trainer. For the VR condition, the industry participants spent 15 minutes with a research assistant to train the assigned task in VR. Immediately after each training, the industry participants completed the assessment component for that task before switching to the other training condition.

The remaining procedure for Industry participants was similar to Student participants as outlined in 4.1.3. Industry participants were asked to perform two maintenance tasks for 10 minutes each under the supervision and assessment of a HI trainer. After training and assessment were completed, the Industry participants were seated in a training room to complete the TX post training questionnaires, self-rated score, and presence and embodiment questionnaires. An audio-recorded group discussion was conducted afterwards, for each of the 4 sessions for Industry participants. Overall, industry participants spent nearly 1 hour throughout the process.

5 RESULTS

We compared the results for each cohort to understand the differences between VR and traditional training. We performed statistical comparison tests for quantitative results and thematic analysis on the interview transcription and written comments. An independent research assistant was employed to conduct thematic analysis using Nvivo and the results were verified by two authors. In this section, we use numeric coding for participants in the *Student* cohort (P1-P23), and letter coding for industry participants (PA-PL).

5.1 Training Performance and Outcomes

We perform comparative tests for the quantitative data on TX, SSCL, EPQ scores, assessment scores, and self-rated performances for two cohorts between both conditions.

5.1.1 Student Learning Experience and Performance

The VARK learning style questionnaire for *Student* participants showed that except for 2 participants, the dominant learning styles for all our student participants were kinesthetics. The survey indicated that almost all our participants preferred learning through hands-on experience and actively engaging with a multi-sensory learning environment.

We calculated the SSCL and EPQ score in both traditional and VR conditions for each participant by summing the answer for each statement (1 point for Strongly disagree and 5 points for Strongly agree), as suggested by Franklin et al. [19]. We performed a Shapiro-Wilk test for normality and confirmed that the scores were normally distributed. We conducted independent t-Tests and found a significant difference ($t(20) = 1.74$, $p < 0.05$) in the SSCL score between traditional condition (mean 48.27, SD 4.5) and VR condition (mean 51.09, SD 2.7), which showed that *participants experienced more satisfaction and self-confidence when training with VR compared to traditional method.* There was no significant difference between VR and traditional condition for the EPQ summative score.

To understand the differences between the two conditions further, we performed a Mann-Whitney test for the TX, SSCL, and EPQ scores between the two conditions, VR and Traditional, and found significant differences in the following statements with advantage for the VR condition.

TX3: The content was organised and easy to follow (Median value VR: Strongly Agree, Traditional: Somewhat Agree, z-score 2.19, p=0.01)

TX8: The time allotted for training was sufficient (Median value VR: Strongly Agree, Traditional: Somewhat Agree, z-score 2.63, p=0.004)

SSSCL3: I enjoyed how my instructor delivered the training. (Median value VR: Strongly Agree, Traditional: Strongly Agree, z-score 1.76, p=0.03)

SSSCL4: The teaching materials used motivating and helped me to learn. (Median value VR: Strongly Agree, Traditional: Somewhat Agree, z-score 2.65, p=0.003)

SSSCL6: I am mastering the content of the training. (Median value VR: Strongly Agree, Traditional: Somewhat Agree, z-score 1.68, p=0.04)

EPQ2: I actively participated in the training session (Median value VR: Strongly Agree, Traditional: Somewhat Agree, z-score 2.24, p=0.01)

EPQ9: Training activities made my learning time more productive (Median value VR: Strongly Agree, Traditional: Somewhat Agree, z-score 2.21, p=0.01).

There was no significant difference for the remaining statements in TX, Satisfaction, SSCL and EPQ subscales. Figure 4 summarises the scores for each statement. Full description of the statements can be found in supplementary material.

Regarding assessment scores by the trainer, we performed an independent t-test on these 5 scores and found no significant difference between the two conditions.

Regarding self-rated performance score on task completion, difficulty, efficiency, recommendation, and satisfaction, we performed a Mann-Whitney test for two independent samples on the rating factors and found a significant difference that 100% of participants would strongly recommend VR training experience while only 54% would strongly recommend traditional training (z-score 2.44, p=0.007).

For the sense of presence scores, the results show that the participants were highly immersed (mean 87.64, SD 10.48) in the virtual world and felt genuinely as if they were on HI factory floor. The participants had a high sense of agency (mean 72, SD 19.95). Within the virtual factory floor, each participant was represented by a head model and two hand models wearing gloves. Overall, there was a high sense of embodiment with the virtual entity (mean 65.18, SD 26.24) with full agency and control of the virtual avatars (mean 83.09, SD 11.95). It should be noted that the participants were only represented with a floating head and hands virtual model without a torso; therefore, we did not expect a 'full' embodiment effect.

For NASA-TLX score, we perform an independent t-test to compare the scores between the traditional and VR groups. We found a significant different in workload for the temporal demand dimension: How hurried or rushed was the pace of the task? in which for the VR conditions, the participant felt less demanding in time to complete the training (mean score 22.54, SD 10.5) compared to traditional condition (mean score 48.8, SD 10.5).

5.1.2 Industry Participants Training Experience and Performance

The industry participants completed a workplace training experience questionnaire with statements related to training objectives, interactions, content, trainer, time and facilities. We completed a Wilcoxon Signed-Ranked test for paired samples between VR and Traditional conditions and did not find significant results between conditions.

Regarding the assessment score, we tested the scores for both VR and traditional conditions for normality using Shapiro-Wilk test and confirmed that the data did not follow the normal distribution. We conducted Wilcoxon Signed-Rank test for paired samples on the score between traditional and VR conditions. There were 5 factors for scoring: overall performance, confidence, time, cleanliness, and safe work practice. There was no significant difference for all the five scores between the conditions.

Regarding the self-rated performance score on task completion, difficulty, efficiency, recommendation, and satisfaction, we performed a Mann-Whitney test for two independent samples on the rating factors, and found no significant difference between VR and traditional conditions. For presence and embodiment, industry participants felt immersed in the training environment in VR with a strong sense of 'being there' (mean 85.08, SD 19.48) and had a strong sense of agency within the virtual world (mean 69.67, SD 25.86).

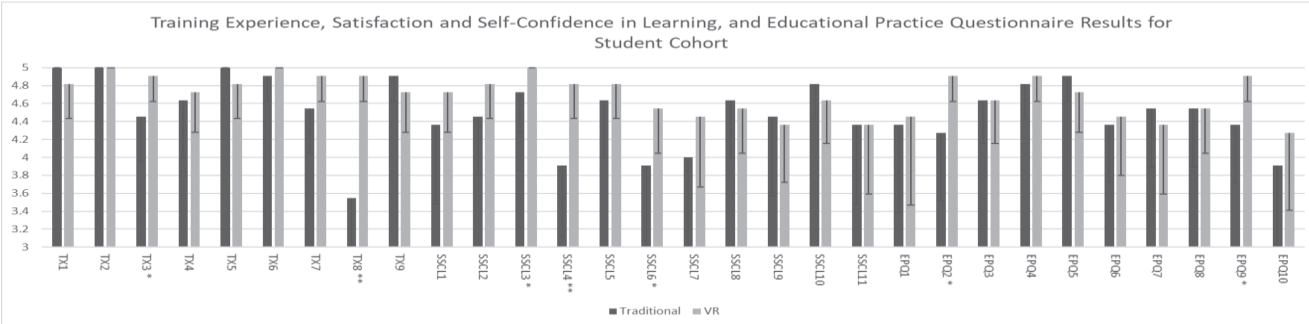

Figure 4: Questionnaires results for TX, SSCL, and EPQ with * and ** indicating significant results (p<0.01 and p<0.001 respectively)

5.2 Training Experiences

We compared the training experiences of student and industry participants from two conditions (VR and Traditional) based on interview data. We reported the results of both studies together in this section. Thematic analysis [27] on interview transcript was conducted by an external independent researcher who was not involved in data collection. The themes were verified independently by the first and second author.

5.2.1 Training Modality

Student participants found both training conditions helpful in preparing them for the assessment on the real machine. For VR, they appreciated the scale, size, and positioning of machine elements (P7,11,23), leading to realistic expectations on the real machine (P12,10); while for Traditional training, the student participants found being able to touch and pick up components helpful (P6,14,18) to contextualise the operation, especially with additional classroom-style presentations from the trainers (P13). However, the participants also noted the struggle in understanding the instruction details and pictures in Traditional training, due to lack of reference for the location of components in the context of the larger machine and quality of the printed photos on papers (P5,15). The current traditional training did not use diagrams, which could have been useful and more engaging (P4,5). Participants commented that performing the actions in VR is more effective than video or verbal instructions (P7,23), as they were encouraged to follow the video demonstration or the researcher's movements within VR. The majority of our student participants have a kinesthetic learning style, as indicated through our questionnaire. Task observation was also facilitated by the ability to move around freely in the virtual environment compared to watching the instructor on the physical machine (as reflected after assessment experience by P11).

For our industry participants, the VR experience provided valuable orientation and assisted in building a mental model of the HPU, which was more useful than the checklist instruction with photos in traditional condition (PA,PF,PH,PG,PI). The mental model was useful for understanding the machine's scale, components locality, and spatial awareness while inside the HPU. One participant (PA) extrapolated that the contextualised mental model supported by VR was more valuable than a checklist with photos, as it is expected that colours of machine parts degrade over time and become less recognisable: "with the VR (training), even though the thing (component) is faded, you still roughly know that is where it is going to be". The industry participants also commented on the benefits for kinesthetic learners while using VR (PC,PH,PG,PI,PJ,PK) or learners with less experience (PC,PF). Industry participants particularly highlighted the value of VR training as an occupational health and safety tool allowing a user to experience mistakes in a safe environment (PA,PC,PE).

5.2.2 Confidence through Repeated Progression

Student participants reported varying confidence levels between VR and Traditional training conditions. For VR, visual realism helped increase participants' confidence before performing the task on the real machine. In VR, participants felt that they could have multiple practice runs (P11,19), in which mistakes made them more aware of their actions. However, participants in the traditional condition felt that there was a disconnection between classroom activities and the actual task (P2,4,16). Despite having access to actual components as props during traditional training, participants wanted more practical training before performing on the machine (P2,5,16,18). Student participants in the VR condition appreciated a progression of task difficulties. The participants started with learning how to use a virtual air-gun to clean a vent at the back of the machine as a way for the researcher to get them familiarised with VR controllers (P7). The ability to repeat tasks also helped them remember what they needed to do in the actual machine, thus boosting their confidence (P10,23).

While the student cohort appreciated the repeated exposure to VR to gain confidence, industry participants commented on the user experience of VR training that assisted them in learning the task. Industry participants appreciated that the VR training was easy to use with an authentic experience (PC), especially the floating checklist interface (PA,PE). Physics simulation was another feature that improved realism for industry participants (PE), including such examples as the way a steel cover slightly wobbled when padlocked or when the valves only slowly rotated to indicate the effort required to loosen them. Industry participants also pointed out the possibility of using the VR system for fault-finding training by incorporating visually damaged components.

5.3 Skill Transfer

We asked specific questions about skill transfer during the interview to understand how the two different cohorts applied their training to assessment tasks. Student participants with traditional training mentally referred to the photo instructions during the assessment to assist with recognising machine components (P15,17). Approximately 63% of student participants self-rated their technical experience as being Novice. Therefore many participants felt that they were overwhelmed by the complexity and concerned about the safety of completing the task (P1,6,17). Our student participants in the Traditional condition resorted to on-site learning during the assessment by watching the trainer's assistance and while doing the task (P5,6,15).

For VR student participants, the skill transfer happened through muscle memory (P7,22). They commented that doing the task in VR assisted with remembering the required hand movements and coordination (P10,23), especially during assessment when they needed it to figure out what to do next. The VR training made the operation easier by providing augmented elements, such as sound, floating checklist, making skills transfer from VR to reality easy (P8). While the similarity of the virtual and physical tools are useful in increasing confidence (P11,12), the VR implementation oversimplified the usage of tools making the steps less memorable to the participants (P9,12). For example, when needing to remove a valve on the filter using a spanner, the VR system only required participants to touch the tip of the valve with the virtual spanner rather than implement the twisting motion. Participants commented that they 'completely forgot that (using spanner) even happened', making the step much less memorable during assessment with the actual machine (P9,12). Such procedural differences made it harder to perform the tasks, especially highlighting the lack of physical efforts required in VR (P7,10,19,22), leading to some concern regarding safety on the real machine (P12). Furthermore, the augmented checklist in VR was helpful during training but causing decreased participants' confidence during assessment (P19).

5.4 Training Improvements

Participants from both cohorts suggested a range of improvements for both training conditions. For Traditional training, student participants expected to have some form of evaluation during training (P6), while industry participants preferred a factory tour before training to provide background orientation (PB,PC). Inconsistent language and terms between training material and maintenance checklist were identified as areas for improvements and additional checklist information was requested.

For VR training, performance feedback is suggested to ensure the trainees are performing the correct action (P11). Safety is of utmost concern for both cohorts, as industry participants requested verification tools such as pressure gauges within the VR environment (PA). Increased realism, such as grime, dirt, and noise in the virtual environment, is also suggested (PD,PF). Additional annotations such as labels and exploded view diagrams, with different levels of difficulties as a form of scaffolding is also recommended (PE).

6 Discussion

Our results show significant results as well as thematic reflections within each cohort between the two conditions of VR and Traditional training. We did not find any significant difference between the conditions. Despite combining between- and within-subject approaches in our cohorts, we recruited an equal number of participants across cohorts and conditions: 11 Student VR, 12 Student Conditional and 12 Industry VR and Traditional participants. We performed Mann-Whitney tests for two independent samples on the 5 categories of assessment scores between Student and Industry cohorts, for each condition VR and Traditional. There was no significant result between the assessment scores. This could be due to the ceiling effects [28] of the assessment scores awarded by our trainers. In addition, the complexity of the tasks was relatively low, only involving identifying components, performing basic hand tools skills (spanner, drill), and following task sequences.

Our student participants did not have prior exposure to industrial settings. Therefore, despite being simplistic, the usage of virtual tools provided a point of reference for our student participants to familiarise themselves with the task to perform on physical machines. On the other hand, our industry participants focused more on the benefits of the user interface, virtual checklist and video panel. They commented that the ability to move the panel in virtual space is useful (PA,PC,PE).

Regarding the training experiences of the students, participants agreed that they did not feel time pressured as compared to traditional condition (TX8 statement). In VR, the learning started immediately as the participants put on the VR headset. As they immersed in the virtual environment, they were exposed to the setting, location, scale of the factory floor. On the other hand, for traditional training, the learning was only completed when the participants arrived at the factory floor, on-site and completed the practical components as part of the assessment. Therefore, this could lead to the differences in students experience as reported that in VR it appeared that more training can be completed in the same period of the allotted time.

Our results agree with prior work on multimedia learning in VR [3]–[5] showing no significant performance gain or skill transfer against traditional multimedia training of slideshows and videos. Our Traditional training condition used a combination of images, videos, and physical mechanical components as study prop. Our empirical contribution unpacks the impact of VR training on novice and expert cohorts and confirms that both the self-rated and trainer's score for assessment task on real life machine did not significantly differ. Our studies however did not detect the impact of cognitive load caused by VR for the student participants, as Parong and Mayer [4] or Frederiksen et al. [6].

Our study contains some limitations. We could not apply similar testing protocols for both cohorts for a fair comparison, due to the limitation of recruitment and time constraint for our industry participants. We also could not complete a retention task as our industry partner HI was required to ship the Wheel Lock HPU to their client. Another variance between VR and traditional training that might have confounded the results was that traditional training was delivered by the Technical Training Manager at HI who had significant professional experience with conducting industrial training. The VR training for both Student and Industry participants were delivered by research assistants who were VR experts and developers but did not have professional experience as an industrial trainer. The research assistants were familiar with the procedural tasks and could answer any questions they had about the VR system.

7 Conclusion

Virtual reality training has provided multiple benefits in the industrial domain. We have presented a virtual reality training platform in collaboration with an international hydraulic manufacturer for training maintenance tasks on a room-scale hydraulic power unit. Our training approach embeds kinaesthetic strategy to VR training. Our training solution enables a checklist UI that presents step by step textual and video instructions, which can be positioned as floating panels anywhere within the virtual environment. We conducted an evaluation of the training system with two cohorts of participants, professional contractors and novice students. We compared our VR training platform with the existing traditional training method of slides, photos, and videos. We found that VR training is engaging, effective in boosting trainees' confidence, especially for novice users. Professional contractors were mostly concerned with the safety features of the training system. Our work contributes to the current literature in VR training with kinesthetics approach. Our studies present parallel empirical comparison of traditional and virtual reality training on novice and professional trainees, as a reflection for future designers, researchers, and stakeholders in this domain.

Acknowledgements

Dr Thuong Hoang is the recipient of an Australian Research Council Discovery Early Career Researcher Award (DE200100898) funded by the Australian Government. This project is supported by HYDAC Australia and Innovation Connection Grant.

REFERENCES

[1] J. E. Naranjo, D. G. Sanchez, A. Robalino-Lopez, P. Robalino-Lopez, A. Alarcon-Ortiz, and M. V. Garcia, “A Scoping Review on Virtual Reality-Based Industrial Training,” *Appl. Sci.*, vol. 10, no. 22, p. 8224, Nov. 2020, doi: 10.3390/app10228224.

[2] R. E. Mayer, “Multimedia learning,” *Psychol. Learn. Motiv.*, vol. 41, pp. 85–139, 2002, doi: 10.1016/S0079-7421(02)80005-6.

[3] J. Parong and R. E. Mayer, “Learning science in immersive virtual reality.,” *J. Educ. Psychol.*, vol. 110, no. 6, pp. 785–797, Aug. 2018, doi: 10.1037/edu0000241.

[4] J. Parong and R. E. Mayer, “Cognitive and affective processes for learning science in immersive virtual reality,” *J. Comput. Assist. Learn.*, vol. 37, no. 1, pp. 226–241, Feb. 2021, doi: 10.1111/jcal.12482.

[5] G. Makransky, N. K. Andreasen, S. Baceviciute, and R. E. Mayer, “Immersive virtual reality increases liking but not learning with a science simulation and generative learning strategies promote learning in immersive virtual reality.,” *J. Educ. Psychol.*, vol. 113, no. 4, pp. 719–735, May 2021, doi: 10.1037/edu0000473.

[6] J. G. Frederiksen *et al.*, “Cognitive load and performance in immersive virtual reality versus conventional virtual reality simulation training of laparoscopic surgery: a randomized trial,” *Surg. Endosc.*, vol. 34, no. 3, pp. 1244–1252, Mar. 2020, doi: 10.1007/s00464-019-06887-8.

[7] V. A. W. M. M. Aleven and K. R. Koedinger, “An effective metacognitive strategy: learning by doing and explaining with a computer-based Cognitive Tutor,” *Cogn. Sci.*, vol. 26, no. 2, pp. 147–179, Mar. 2002, doi: 10.1207/s15516709cog2602_1.

[8] F. Ganier, C. Hoareau, and J. Tisseau, “Evaluation of procedural learning transfer from a virtual environment to a real situation: a case study on tank maintenance training,” *Ergonomics*, vol. 57, no. 6, pp. 828–843, Jun. 2014, doi: 10.1080/00140139.2014.899628.

[9] N. Gavish *et al.*, “Evaluating virtual reality and augmented reality training for industrial maintenance and assembly tasks,” *Interact. Learn. Environ.*, vol. 23, no. 6, pp. 778–798, Nov. 2015, doi: 10.1080/10494820.2013.815221.

[10] C. Vazquez, L. Xia, T. Aikawa, and P. Maes, “Words in Motion: Kinesthetic Language Learning in Virtual Reality,” in *2018 IEEE 18th International Conference on Advanced Learning Technologies (ICALT)*, Jul. 2018, pp. 272–276, doi: 10.1109/ICALT.2018.00069.

[11] K. Nesenbergs, V. Abolins, J. Ormanis, and A. Mednis, “Use of Augmented and Virtual Reality in Remote Higher Education: A Systematic Umbrella Review,” *Educ. Sci.*, vol. 11, no. 1, p. 8, Dec. 2020, doi: 10.3390/educsci11010008.

[12] A. Janonis, E. Kiudys, M. Girdžiūna, T. Blažauskas, L. Paulauskas, and A. Andrejevas, “Escape the Lab: Chemical Experiments in Virtual Reality,” *Inf. Softw. Technol. ICIST 2020. Commun. Comput. Inf. Sci.*, vol. 1283, pp. 273–282, 2020, doi: 10.1007/978-3-030-59506-7_22.

[13] J. Wolfartsberger and D. Niedermayr, “Authoring-by-Doing: Animating Work Instructions for Industrial Virtual Reality Learning Environments,” *Proc. - 2020 IEEE Conf. Virtual Real. 3D User Interfaces, VRW 2020*, pp. 173–176, 2020, doi: 10.1109/VRW50115.2020.00038.

[14] J. J. Roldán, E. Crespo, A. Martín-Barrio, E. Peña-Tapia, and A. Barrientos, “A training system for Industry 4.0 operators in complex assemblies based on virtual reality and process mining,” *Robot. Comput. Integr. Manuf.*, vol. 59, pp. 305–316, Oct. 2019, doi: 10.1016/j.rcim.2019.05.004.

[15] L. Pérez, E. Diez, R. Usamentiaga, and D. F. García, “Industrial robot control and operator training using virtual reality interfaces,” *Comput. Ind.*, vol. 109, pp. 114–120, Aug. 2019, doi: 10.1016/j.compind.2019.05.001.

[16] M. Perez-Ramirez, G. Arroyo-Figueroa, and A. Ayala, “The use of a virtual reality training system to improve technical skill in the maintenance of live-line power distribution networks,” *Interact. Learn. Environ.*, vol. 29, no. 4, pp. 527–544, 2021, doi: 10.1080/10494820.2019.1587636.

[17] J. Lacko, “Health Safety Training for Industry in Virtual Reality,” in *2020 Cybernetics & Informatics (K&I)*, Jan. 2020, pp. 1–5, doi: 10.1109/KI48306.2020.9039854.

[18] T. Hoang, S. Greuter, S. Taylor, G. Aranda, and G. T. Mulvany, “An Evaluation of Virtual Reality for Fear Arousal Safety Training in the Construction Industry,” in *2021 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, Oct. 2021, pp. 177–182, doi: 10.1109/ISMAR-Adjunct54149.2021.00044.

[19] A. E. Franklin, P. Burns, and C. S. Lee, “Psychometric testing on the NLN Student Satisfaction and Self-Confidence in Learning, Simulation Design Scale, and Educational Practices Questionnaire using a sample of pre-licensure novice nurses,” *Nurse Educ. Today*, vol. 34, no. 10, pp. 1298–1304, Oct. 2014, doi: 10.1016/j.nedt.2014.06.011.

[20] D. Mirauda, N. Capece, and U. Erra, “Sustainable water management: Virtual reality training for open-channel flow monitoring,” *Sustain.*, vol. 12, no. 3, 2020, doi: 10.3390/su12030757.

[21] S. G. Hart and L. E. Staveland, “Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research,” 1988, pp. 139–183.

[22] F. Winther, L. Ravindran, K. P. Svendsen, and T. Feuchtner, “Design and Evaluation of a VR Training Simulation for Pump Maintenance Based on a Use Case at Grundfos,” in *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, Mar. 2020, pp. 738–746, doi: 10.1109/VR46266.2020.00097.

[23] W. L. Leite, M. Svinicki, and Y. Shi, “Attempted Validation of the Scores of the VARK: Learning Styles Inventory With Multitrait–Multimethod Confirmatory Factor Analysis Models,” *Educ. Psychol. Meas.*, vol. 70, no. 2, pp. 323–339, Apr. 2010, doi: 10.1177/0013164409344507.

[24] R. Paul Griffin, “Workplace learning evaluation: a conceptual model and framework,” *Ind. Commer. Train.*, vol. 43, no. 3, pp. 172–178, Apr. 2011, doi: 10.1108/00197851111123631.

[25] B. G. Witmer and M. J. Singer, “Measuring Presence in Virtual Environments: A Presence Questionnaire,” *Presence Teleoperators Virtual Environ.*, vol. 7, no. 3, pp. 225–240, Jun. 1998, doi: 10.1162/105474698565686.

[26] K. Kilteni, R. Groten, and M. Slater, “The Sense of Embodiment in Virtual Reality,” *Presence Teleoperators Virtual Environ.*, vol. 21, no. 4, pp. 373–387, Nov. 2012, doi: 10.1162/PRES_a_00124.

[27] V. Braun and V. Clarke, “Thematic analysis.,” in *APA handbook of research methods in psychology, Vol 2: Research designs: Quantitative, qualitative, neuropsychological, and biological.*, Washington: American Psychological Association, 2012, pp. 57–71.

[28] M. McBee, “Modeling Outcomes With Floor or Ceiling Effects: An Introduction to the Tobit Model,” *Gift. Child Q.*, vol. 54, no. 4, pp. 314–320, 2010, doi: 10.1177/0016986210379095.