2023 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct) | 979-8-3503-2891-2/23/$31.00 ©2023 IEEE | DOI: 10.1109/ISMAR-Adjunct60411.2023.00041

Comparative Evaluation of Virtual Reality and In-Person Onboarding for Assembly Trainings in Manufacturing

Tanja Zigart[a]*, Felix Stürzl[a], Daniel Niedermayr[b], Josef Wolfartsberger[b], Sabrina R. Sorko[c]

[a]Technische Universität Wien, Taubstummengasse 27, 1040 Wien, Austria
[b]Fachhochschule Oberösterreich Campus Steyr, Steyr, Austria
[c]Fachhochschule Joanneum Kapfenberg, Kapfenberg, Austria

ABSTRACT

In recent years, employee turnover has increased to the point where the manufacturing workforce is constantly changing. The onboarding and training processes are time-consuming and costly. Virtual reality (VR) allows skills to be trained before working on the actual production line and learning in a safe environment. This approach promises cost and productivity benefits for companies and a personalized learning experience for users. In this paper, we present the results of a user evaluation of an industrial VR use case for training the assembly of a compressor. The goal of the evaluation is to compare conventional onboarding training in-person with VR training in terms of learning success, workload, net promotion score, and qualitative feedback. Additionally, user acceptance and usability of the VR training are questioned. The results show that the users still prefer conventional training, as the net promoter score is higher for the in-person training. Also, the results of the learning success show better results for in-person training than the VR training.

Keywords: Virtual Reality, Evaluation, User Study, Manufacturing, Assembly Training.

Index Terms: Empirical studies in HCI; User-centered design; Virtual Reality

1 INTRODUCTION

Training for industrial assembly processes is important in manufacturing to meet quality standards and productivity goals. Due to the high turnover in the workforce, efficient and effective training is essential. One way to make training more accessible, safer, and more cost-effective is to use Virtual reality [1]. Manufacturing companies are eager to take advantage of the benefits of VR, such as reducing training time, increasing learning success, and increasing flexibility in the training process [2].

This paper evaluates an industrial VR assembly training of a compressor part compared to a conventional in-person training. We compared the as-is learning scenario in the manufacturing company (an in-person trainer) and a VR scenario, which the company considered replacing the in-person onboarding training for new employees with. The aim was to compare the learning outcome of the two scenarios and evaluate human factors while using the system. For the evaluation, 57 students were divided into two groups: one group tested the in-person training, while the other group tested the VR training. In addition to learning success, human factors such as perceived user acceptance, workload, and usability were evaluated and analyzed in the results section. Qualitative feedback from the participants is summarized, and potential improvements are discussed.

2 THEORY

VR has become an increasingly popular tool for training and education. It is common in industrial training to learn manual tasks in a controlled environment. In this domain, it is used to memorize the sequence of assembly or maintenance steps.

2.1 Virtual reality in training processes

VR as a learning medium can be used in different training scenarios along the employee life cycle: from training new employees as part of onboarding to individual on-demand training. The basis for the use of VR as a learning medium is learning theory. Learning theory research does not follow a unified approach, but differentiates between five views: Behaviorism, Cognitivism, Constructivism, Experientalism, and Connectivism. The respective paradigms provide insight into the course of learning processes and thus offer an essential basis for the use of new technologies as a learning medium. A fundamental understanding of the interdependencies enables a proper determination of adequate VR learning scenarios [3]. In this respect, a key finding is the relevance of the widespread known 'enactment effect'. According to this, the learning success of VR training depends on the learner's level of activity. The more activity based training, the more likely it is that what is learned is transferred to long-term memory. Furthermore, the closeness to reality plays an important role. Successful VR training ensures that what is learned can be easily transferred to the learner's real life world [4]. Most studies in this domain have reported the positive effects of VR-based training for both training effectiveness and user acceptance [5].

Other studies have explored the use of VR for training complex surgical procedures. For example, Zhou et al. [6] investigated the benefits of haptic feedback in laparoscopic surgery training simulators. The authors reported that learning with haptic feedback was significantly better during initial training. However, the cost and complexity of implementing haptic feedback should be considered. Strandholt et al. [7] integrate physical tools into VR to enable the sense of touch and increase immersion. Although participants in this study perceived the interactions as more realistic, specific tools were implemented. Therefore, applying all types of interactions in an arbitrary training process is difficult. While these examples show that VR benefits from more realistic interactions, they are difficult to generalize to any application.

The way in which a training is implemented can have a significant impact on the outcome. Huang et al. [8] have shown that greater immersion in VR can lead to positive learning effects and higher motivation and engagement. Wolfartsberger et al. [9] compared variations of assembly training with different levels of virtual instructions. The results showed significant differences in learning outcomes even with minor modifications in the virtual training. In particular, the training benefits from a lower level of guidance because the user is more cognitively challenged, which supports learning transfer [10].

---

* tanja.zigart@tuwien.ac.at

Work licensed under Creative Commons Attribution 4.0 License.
https://creativecommons.org/licenses/by/4.0/

DOI 10.1109/ISMAR-Adjunct60411.2023.00041

Mayrhofer et al. [11] describe the training to be interactive, visual, and activate learners with tests, quizzes, and games to stimulate action and reflection. Especially for on-the-job training, this is important and prepares learners for follow-up hands-on training that can be conducted more efficiently with having more time for questions and practical training [11].

2.2 Measuring learning success

When implementing training arrangements in practice, companies, in particular, focus on the quality of the learning sequences due to cost-benefit considerations. To consider the advantage of the training, the learning success must be made visible and measurable. Learning success consists of an objective and a subjective part, requiring different measurement systems [12, 13]. The measurement framework developed by Kirkpatrick takes both perspectives into account [14]. The model provides a guideline using different evaluation instruments on four levels: Reaction, Learning, Behavior, and Results [15].

Level one (reaction) focuses on the learners and their satisfaction within the learning scenario. This result indicates the engagement of the learners towards the training and, thus, the motivation of attending or fulfilling the training. Those aspects mainly address subjective learning success, so methods such as NASA-TLX [16] or usability tests are suitable. Level two (learning) reflects on the training itself and the learning outcome. Within this level, the gained competencies and a possible change in the learners' attitudes are measured. In terms of VR-supported learning, the understanding and training of competences is evaluated, as well as the handling of the technology. On this level, short- and long-term measurements should be conducted [17, 18]. Taking the individual learning curve into account, Kirkpatrick focuses on the behavior of the learners at level three (behavior). The aim is to regularly monitor the learning outcome over time in daily business. To gain a realistic measurement result, it is essential to include heterogeneous perspectives at this level [19] to adjust the training arrangement and thus increases its quality. The last step of Kirkpatrick's approach (results) is focusing on the organizational benefits resulted from the training. This is done by calculating the Return on Expectations (ROE) and comparing the expectations of all stakeholders involved (learners, management, leaders) with the results achieved [20]. In practices, this can be realized by using a Balances Score Card (BSC) [21], as this instrument combines hard and soft facts from different perspectives [22]. The approach of Kirkpatrick builds the basis for the methodological framework for measuring the learning success of the research results presented in this paper.

3 METHODOLOGY

A use case was developed for a manufacturing company to onboard new employees to the assembly process of a compressor part. The setup, the experimental procedure, participants, and evaluation method are explained in this section.

3.1 Experimental setup of the use case

To conduct the study two experimental training scenarios were designed. One scenario provided the assembly training via a VR training. The second scenario was a one-on-one in-person training and is the current state of the training in the manufacturing company. The aim was to compare the learning outcome of the two scenarios and evaluate human factors while using the system.

The virtual training of the assembly was implemented in Unity and a standard HTC Vive Focus 3 headset with its included tracked controllers was used tethered to a PC. At the beginning of the training, all required assembly parts and tools are spread out on a virtual table. Users can pick up parts and tools with the controllers and they can be assembled by moving them to a per-object defined assembly zone on other parts of the assembly. Objects do not snap to their target position, but are guided towards the target when they are close. For example, when a screw is moved close to a corresponding screw hole with its tip, it will only slide along the hole axis when it is pushed in further until it reaches its end stop. This approach was chosen because it allows precise assembly and is closer to the real assembly. Therefore, it could be more intuitive compared to snapping into position when it is released within reach of its target. Screws additionally had to be tightened with a tool, either a wrench or a screwdriver, after they were placed on top of the screw hole. The tool has to be put on the screw head and rotated realistically to tighten the screw into place (see Fig. 1 top). This approach was chosen to investigate the impact of using virtual tools in a realistic fashion on the learning process and is being evaluated in another study. Another step that required the use of a tool was the wiring. Participants had to thread cables into the housing and connect the wires according to a sketch (see Fig. 1 bottom). Users had to use a screwdriver to pry open the connector holes of a terminal block and wire the cables. This approach was chosen to investigate the impact of using virtual tools in a realistic fashion on the learning process and is being evaluated in another study. Because the wiring takes place in a small housing and therefore is a very delicate process, the housing was enlarged during wiring to make it more comfortable to the users. Due to technical limitations with high numbers of complex model parts, there are no collisions between objects and objects stay in mid air when released instead of falling to the ground.

The instructions for the virtual training are displayed on a screen next to the assembly (see Fig. 1 top). They are automatically advanced once a step has been successfully completed. The instructions always consist of a short informational text and optionally an image (a sketch or a photo). This dedicated monitor for the instructions was chosen instead of showing the necessary information directly on top of the assembly, because it better challenges users to think about the given information on the monitor and interpret it instead of blindly following animations on top of the virtual assembly objects. If users get stuck, they can request additional assistance at the push of a button. Here, the object to be used is highlighted and animated in a semi-transparent manner moving onto its target position to show more precisely how a component should be assembled. Completed steps are indicated with a success sound, so progress is always perceived, even if the screen is not in the user's view. The participants could go back and forth to repeat each step in their own pace and as often as they needed to. Figure 1 shows the VR use case training setting.

Step 2: Mount actuator housing to sealing housing

Step 2.1:

Tighten with four screws (M12)

Tighten to a torque of 35Nm ± 5Nm

Step 3: Wiring

Step 3.3:

Connect data cables.

Step 4/6

Figure 1: VR training setting. All assembly parts and tools are lying on the table. The required actions are described on the blue monitor. On the screenshot a screw is being tightened with a wrench (top). Users had to connect wires according to a picture in an enlarged copy of the housing (bottom).

The second scenario was a one-on-one in-person training. The trainer followed a standardized script to present the step-by-step instructions in the same order and with the same information value each time. Figure 2 shows the use case setting for the in-person training scenario which consisted of various assembly parts and tools. The compressor part and the needed tools were placed on the workbench. In the first run-through, the trainer demonstrated the assembly by conducting all the steps from start to finish and explaining necessary information. The provided information included details such as screw dimensions, required torques for screw tightening, and the proper sequence and fitting of wiring cables. The participants could then assemble the part themselves and ask for instructions and information they did not remember from the first run-through. There was no additional paper or digital instructions for the participants. The assembly process consisted of 16 assembly steps. To conclude the steps, the participants had to use a variety of tools and had to apply the learned information to do so correctly. The trainers in the in-person training had an expert script to ensure replicable trainings for all participants. The same information like in the VR training was given verbally. No additional information was given in the in-person training compared to the VR training. The company expert, who does the in-person training, created the content. The VR use case was developed by the researchers and tested with the company expert. As soon as the expert decided the scenarios to be comparable, the evaluation took place.

Figure 2: In-person training setting. All assembly parts and tools are placed on the table. The step-by-step instructions were given by a trainer.

3.2 Experimental procedure

At the beginning of the study, participants received information about the study, read and signed the ethics and privacy statement, and completed the online demographic questionnaire. Then they completed the training either with an in-person trainer or in VR. The participants had a maximum training time of 30 minutes for both scenarios and could stop as soon as they felt ready for the test. Participants were randomly allocated to the two groups. Participants in the VR group received a brief introduction to the VR controls. After completing the training, participants were tested on their acquired knowledge by undergoing an examination that required them to assemble the real compressor part and by an evaluation of the learning success. The test after the training included all 16 assembly steps and required information learned in training such as screw dimensions, required torques for screw tightening, and the proper sequence and fitting of wiring cables. Furthermore, basic assembly skills like the use of a screw driver and torque wrench were required. In addition, participants were tested on their ability to wire cables by prying open terminal block connector holes with a screwdriver. At the end of the study, participants completed an online questionnaire on user acceptance, workload, usability, net promoter score, and answered some open-ended questions.

3.3 Participants

The participants were students from the university where the study took place. They did not have any previous knowledge in the field of assembly neither with VR trainings. 28 students did the in-person training, 29 students did the VR training. 77.97 percent of the participants considered themselves male, 22.03 percent female, and 0 percent did not say. Most students were 18-23 years old (81.36%), 15.25 percent were between 24 and 29, and 3.39 percent were over 30. 84.85 percent had a high school diploma as their highest qualification, 10.17 percent a bachelor's degree, 3.39 percent a compulsory education degree, and 1.69 percent had a master's degree.

3.4 Evaluation method

The evaluation was divided into learning success and human factors. The learning success was evaluated using the approach of Kirkpatrick [15], taking objective and subjective learning success into account. As a limitation, due to the limited time within the project, only the first two steps, reaction and learning, have been carried out so far. It is recommended to continue the research over time to get a long-time measurement result. To test the gained knowledge and skills, a practical assembly was carried out after completing the training. The participants had to assemble the real compressor part themselves without help. During some of the steps, participants had to answer questions about the information they had learned, such as the dimensions of the screws used or the torque required to tighten a screw. These questions were asked to test whether the participants had learned all the detailed information, rather than just joining the parts together.

For the evaluation of human factors, user acceptance with the 'technology acceptance model' (TAM) [23-25], workload with the 'NASA raw task load index' (NASA-RTLX) [26], usability with the 'system usability scale' (SUS) [27], and the net promoter score with 'NPS' [28] are queried with a questionnaire after 30 minutes of training.

4 RESULTS

The objective learning success, the evaluation of the human factors, and the qualitative feedback were analyzed in this section.

4.1 Evaluation of the learning success

Table 1 shows the objective learning success for fulfilling the given task of assembling a compressor after a maximum of 30 minutes training. In the in-person scenario, all participants did two run-throughs of the given tasks with the trainer before taking the test. In the VR scenario, the participants did in average 2.04 run-throughs in a training time of 22:26 minutes. 16 assembly steps had to be conducted. The average time for conducting the test after the in-person training was 10:08 and 12:43 minutes after the VR training. Correct and incorrect execution of each step has been recorded, and the percentage of correct tasks is shown. The number of completed questionnaires were 26, as some did not answer all questions.

Table 1: Learning success for both training scenarios.

	Trainer	VR
Average time for conducting the test	10:08 min	12:43 min
Average correct tasks	14.58	11.85
Average incorrect tasks	1.42	4.15
Percentage of correct tasks	91.11 %	74.04 %
n	26	26

The results show higher learning success and shorter test-taking times for the in-person scenario than the VR training scenario. This means the number of correctly executed assembly steps was 17.07 percent higher and 20.31 percent faster after the in-person training than in the VR training. This can be explained by several limitations in the research design: (1) The measurement of learning success took place immediately after the in-person training, so what the participants had heard could easily be recalled from their short-term memory. Whether long-term higher learning success can be achieved through in-person training than through VR-based training would have to be examined in a long-term survey; (2) The measured learning success includes not only the development of technical competence but also the user safety of the participants. In order to make an unbiased statement about the learning success of VR training, a high level of operational safety of the VR device would have to be ensured in advance.

Out of 16 assembly steps, there were five specific steps the participants trained in the VR had much more problems. These problems were mainly connected with transferring the digitally learned elements into the real-world setting. In the VR the whole compressor was shown, which is a few meters in size. In the real-world setting, there was only the part from the assembly, which led to confusions of some participants. The joining steps were remembered best by all participants, additional information like the torque or tolerances was poorly memorized.

4.2 Evaluation of human factors

To evaluate human factors, a questionnaire about workload, net promoter score, and open questions was filled out after each test scenario. Additionally, questions about user acceptance and usability were asked in the VR scenario.

The workload was rated with the NASA-raw task load index (NASA-RTLX). Participants rated the workload slightly higher in the trainer scenario than in the VR training, see Table 2. The NASA-RTLX scores show a medium workload for both scenarios [29].

Table 2. Workload (NASA-RTLX) for the VR application.

	Trainer	VR
NASA-RTLX score (0-100)	28.57	28.30

The net promoter score was calculated by subtracting the number of detractors from the number of promoters [17]. The NPS shows better results for the trainer scenario (28.57 %) than the VR scenario (6.90 %), see Table 3. The higher the NPS, the more likely it is that the training will be recommended to others. 28.57 percent is already considered as a "good" net promoter score, while 6.90 percent needs improvement. Figure 3 shows the distribution of the promoters, passives, and detractors in the trainer and VR scenario.

Table 3. Net promoter score (NPS) for both training scenarios.

	Trainer	VR
Promoters (Pr)	13 / 46.43 %	11 / 37.93 %
Passives (Pa)	10 / 35.71 %	9 / 31.03 %
Detractors (De)	5 / 17.86 %	9 / 31.03 %
n	28	29
NPS score (Pr-De [%])	28.57 %	6.90 %
Standard deviation	2.46	2.55

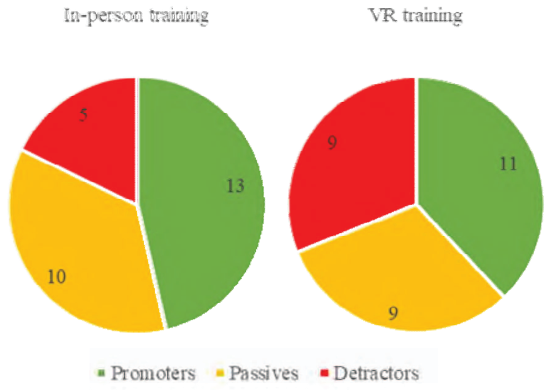

Figure 3. Net promoter score for both training scenarios.

The results of the user acceptance of the VR training were evaluated with the technology acceptance model (TAM) (Figure 4).

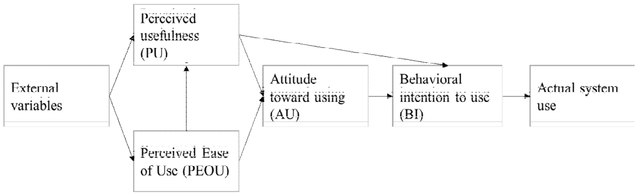

Figure 4. Technology acceptance model (TAM).

The following five hypotheses were tested:
H1: PEOU has a positive influence on PU
H2: PEOU has a positive influence on AU
H3: PU has a positive influence on BI
H4: PU has a positive influence on AU
H5: AU has a positive influence on BI

Table 4 shows the descriptive analysis of the constructs with response formats of 1 (do not agree at all) to 5 (fully agree).

Table 4. Descriptive analysis of the constructs.

Construct	n	Mean	SD
PU	29	3.52	0.81
PEOU	29	3.57	0.63
AU	29	3.76	1.09
BI	29	3.68	1.01

Cronbach's alpha measures the scale's internal consistency and was calculated for each construct. Hair et al. [30] recommend Cronbach's alpha values of 0.6 to 0.7 as the limit of acceptability. A maximum alpha value of 0.9 is recommended [31]. A very high value for Cronbach's alpha indicates that some items are redundant and may be testing the same question. The Cronbach's alpha values in the present study (see Table 5) suggest the use for further analysis. AU consists of one item only, therefore no Cronbach's alpha value is available.

Table 5. Reliability with Cronbach's alphas.

Construct	Cronbach's alpha	Evaluation	Number of items
PU	0.87	Very good	6
PEOU	0.84	Very good	6
AU	-	-	1
BI	0.75	Good	3

Table 6 shows the construct and dependable variables of the hypotheses, the regression coefficient, standard error, Beta, T, and significances.

Table 6. Coefficients of the constructs and dependable variables.

		Not standardized coefficient		Standardized coefficient		
H	Construct/ dependable variable	Regression coefficient	Standard error	Beta	T	Sig.
1	PEOU/PU	0.42	0.13	0.54	3.32	0.003
2	PEOU/AU	0.24	0.10	0.42	2.37	0.025
3	PU/BI	0.67	0.08	0.84	8.08	< 0.001
4	PU/AU	0.32	0.13	0.44	2.58	0.016
5	BI/AU	0.53	0.15	0.57	3.64	0.001

The correlation analysis with one-sided significance between the items shows significant relationship (null hypothesis rejected) between:

H1: PEOU → PU (PU can be explained by 53.90% from PEOU)
H3: PU → BI (BI can be explained by 84.10% from PU)
H4: PU → AU (44.40% of AU can be explained by PU)
H5: AU → BI (57.30% of BI can be explained by AU)

There is no significant association (null hypothesis accepted and random association cannot be excluded) between:
H2: PEOU → AU (AU can be explained by 41.5% from PEOU, but this relationship can be considered as random with a 12.5% significance level).

Table 7. Correlation analysis (one-sided significance) of the constructs.

		PEOU	PU	AU
PU	Pearson-correlation	0.54**		
	Sig. (1-sided)	0.0013		
AU	Pearson-correlation	0.42*	0.44**	
	Sig. (1-sided)	0.0125	0.0079	
BI	Pearson-correlation	0.56**	0.84**	0.57**
	Sig. (1-sided)	<0.001	<0.001	<0.001

* Correlation is significant at the 0.01 level (one-sided)

** Correlation is significant at the 0.05 level (one-sided)

The participants rated the usability of the VR training with an average SUS score of 78.71 (Table 8). According to Brooke [27], the usability of the VR training is rated "acceptable". The standard deviation for the SUS score is 10.62 percent. Figure 5 shows the SUS score in a boxplot diagram. VR SUS scores in literature show 80.00 [32], 75.50 [33], 76.40 [34], and 86.56 and 85.94 [35]. Common SUS scores for everyday products are e.g. for Google search 93.40, for Amazon 81.80, for Word 76.20, for Excel 56.50 [36].

Table 8. Usability (SUS) results of the VR application.

	VR
SUS score (0-100)	78.71
Standard deviation	10.62
95% confidence interval	3.87
Upper limit confidence interval	80.60
Lower limit confidence interval	76.80

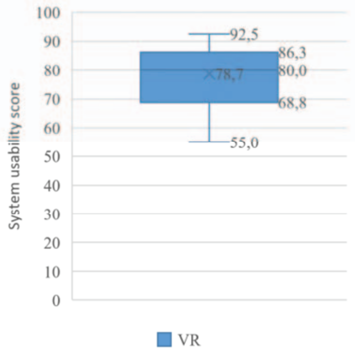

Figure 5. Usability (SUS) diagram of the VR application.

4.3 Qualitative feedback

In addition to the questions about usability, user acceptance, workload, and learning success, the users were asked open questions about their opinion on the advantages and disadvantages of VR training for manufacturing.

The main advantages were seen in training the processes and procedures individually and that no real material, like a workpiece or tools, was needed for the training. The parallel training process and the relief for trainers were mentioned as well. It was described as fast learning, flexible, adaptable, and cost-effective. Table 9 shows the mentioned advantages and the number of participants mentioning them.

Table 9. Mentioned advantages of a VR training.

Advantages	Number of mentions
Train/exercise/get to know processes and procedures	9
No real workpiece/workshop/material necessary to become familiar with workpiece/workshop	6
Safe learning for users	3
Parallel training of several employees/relief of the trainers	3
Fun, playful learning/ no strenuous learning/ exciting	3
Better understanding/ practical learning effect	3
Fast learning	3
In the final assembly of a production/for assembly instructions	2
Objects are quick to assemble	1
Flexible/ adaptable to tasks and areas	1
More productive after enrollment	1
Repeatability of practice	1
Cost-effective	1

On the other hand, participants mentioned disadvantages with some being contrary to the advantages (Table 10). The main disadvantage is the missing haptic feedback and the lack of practical experience during the training. The danger of learning something wrong and repeating it several times is also greater in VR than next to a trainer. One explanation for this effect could be that a trainee does not receive immediate corrective feedback from a trainer. Instead, it is possible that an inefficient or incorrect action is not detected by the VR application and is therefore incorrectly learned. The knowledge transfer differs too much from the virtual part to the real workpiece that the learnings are hard to use. Since a realistic avatar including displayed hands was used, a mishandling of tools due to missing visual features is unlikely. However, part of this problem can be explained by the lack of haptic feedback, which occurred when participants tried to pry open the terminal block connector holes while wiring the cables. Furthermore, differences in room layout and not having the entire compressor in the room, as well as a lack of positive feedback from the program upon completion of a step, could lead to participants feeling insecure and second-guessing themselves. Additionally, not everyone is comfortable using VR, some people reported headaches, eye strains, cognitive overload, and motion sickness.

Table 10. Mentioned disadvantages of VR training.

Disadvantages	Number of mentions
No haptic feedback/ hands-on-feeling/ lack of practical experience	4
Teach-in error/ processes not displayed correctly	3
Costly/ Extra equipment	3
Knowledge transfer when VR and reality differ too much	3
Physical exertion (headaches, eye strain, cognitive overload/motion sickness)	3
Problems with operation/finding one's way around (especially for people with less technical affinity)	2
Customization/creation of the VR environment cumbersome/expensive	2
Takes a long time	1
Occurrence of technical difficulties	1
Will be replaced by AR/MR	1
Lack of social interaction	1

5 Discussion

The study shows a tendency for participants to prefer the in-person training scenario. The VR training received a good usability score of 78.71 out of 100. Even though some participants reported physical issues like headaches, cognitive overloads, and eye strain, the workload is only slightly different in the two scenarios. It shows that, in average, the VR scenario is mentally and physically not more exhausting than an in-person training. The net promoter score shows a clear tendency toward the recommendation for in-person training compared to VR training. Also, the learning success is 17.07 percent higher in the in-person training compared to the VR scenario. The participants had problems transferring their digital gained knowledge into the real-world setting. Regarding user acceptance, four out of five hypotheses were confirmed.

As described in [11], the VR training could be an addition to the in-person training. Getting first insights into the steps and the basic information about the assembly in VR and then having more time in the in-person training for questions and practical advice might be an efficient way to go.

VR novices were not excluded, every participant received a VR intro training in the beginning. This is representing the current situation of our target group, as most of the manufacturing workers are VR novices.

5.1 Limitations

The results from this study are related to this particular scenario. It is not clear to what extent the results are transferable to other scenarios. Although the VR environment replicates the real scenario as closely as possible, there are still differences (virtual working environment, lighting, etc.). These factors can have an influence on the results.

In addition, the VR use case was developed in TRL stage 7; this means that there are still some usability improvements that can be made for the serial operation. The use case setting of the VR training was purely focused on VR without any hands-on experience in real life. The participants assembled the compressor for the first time during the test. This might be an unrealistic setting for industry, but it compares digital-only to in-person training.

The in-person training was highly standardized in the process for all trainers to always act very similarly. Nevertheless, two different people conducted the training, resulting in small differences between the trainers.

Due to the limited number of participants, the analysis of the technology acceptance model (TAM) is very limited, e.g., a multiple regression analysis, typically preferred, could not be employed in this study.

6 CONCLUSION

The presented research aimed to test a VR training process and evaluate the objective learning success, user acceptance, workload, and usability compared to conventional in-person training.

We presented the results of a VR user study with 57 participants. The feedback varied within the groups, with some finding the VR system good, but due to problems of transferring the digital learnings into the real assembly setting, it is not seen useful for a standalone training, similar to the reported research in [1]. Reasons might be the missing haptic feedback, the lacking of practical experience, and that some steps are challenging to train virtually and therefore certain simplifications in VR were necessary.

Others complain about using VR in the onboarding process because there is a lack of social interaction, and some people cannot use VR due to physical issues. Additionally, the result shows that the participants finishing the VR training had a lower learning success than the ones trained in person. Like in [1], the users appreciated the intuitive use of visual feedback for interaction. Compared to [34], the NASA-RTLX was higher in this study (28.30/100), while in their study, it was 21.40 for VR, showing a higher workload in this VR setting.

This study shows that the VR system used in the case study needs further improvement to be entirely accepted by the workers, such as integrating advanced haptic feedback for a better transfer into the real world.

ACKNOWLEDGMENTS

The research presented in this paper is supported by the Austrian Research Promotion Agency (FFG) through the project IMPACT-sXR (FFG- 888744). We thank the company Hoerbiger Wien GmbH for providing the industrial use case and the actual components.

REFERENCES

[1] D. Niedermayr und J. Wolfartsberger, „Design and Evaluation of a Virtual Training Environment for Industrial Assembly Tasks“, *SSRN Journal*, 2021, doi: 10.2139/ssrn.3862367.

[2] F. Holly *u. a.*, „Gaining Impact with Mixed Reality in Industry – A Sustainable Approach“, in *2022 8th International Conference on Computer Technology Applications*, Kapfenberg Austria: ACM, Mai 2022, S. 128–134. doi: 10.1145/3543712.3543729.

[3] J. Radianti, T. A. Majchrzak, J. Fromm, und I. Wohlgenannt, „A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda“, *Computers & Education*, Bd. 147, S. 103778, Apr. 2020, doi: 10.1016/j.compedu.2019.103778.

[4] S. A. Smith, „Virtual reality in episodic memory research: A review“, *Psychon Bull Rev*, Bd. 26, Nr. 4, S. 1213–1237, Aug. 2019, doi: 10.3758/s13423-019-01605-w.

[5] S. K. Renganayagalu, S. C. Mallam, und S. Nazir, „Effectiveness of VR head mounted displays in professional training: A systematic review“, *Technology, Knowledge and Learning*, S. 1–43, 2021.

[6] M. Zhou, S. Tse, A. Derevianko, D. Jones, S. Schwaitzberg, und C. Cao, „Effect of haptic feedback in laparoscopic surgery skill acquisition“, *Surgical endoscopy*, Bd. 26, S. 1128–1134, 2012.

[7] P. L. Strandholt, O. A. Dogaru, N. C. Nilsson, R. Nordahl, und S. Serafin, „Knock on wood: Combining redirected touching and physical props for tool-based interaction in virtual reality“, gehalten auf der Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 2020, S. 1–13.

[8] W. Huang, R. D. Roscoe, M. C. Johnson-Glenberg, und S. D. Craig, „Motivation, engagement, and performance across multiple virtual reality sessions and levels of immersion“, *Journal of Computer Assisted Learning*, Bd. 37, Nr. 3, S. 745–758, 2021.

[9] J. Wolfartsberger, R. Zimmermann, G. Obermeier, und D. Niedermayr, „Analyzing the potential of virtual reality-supported training for industrial assembly tasks“, *Computers in Industry*, Bd. 147, S. 103838, Mai 2023, doi: 10.1016/j.compind.2022.103838.

[10] F. Schuster, B. Engelmann, U. Sponholz, und J. Schmitt, „Human acceptance evaluation of AR-assisted assembly scenarios“, *Journal of Manufacturing Systems*, Bd. 61, S. 660–672, Okt. 2021, doi: 10.1016/j.jmsy.2020.12.012.

[11] W. Mayrhofer, S. Nixdorf, C. Fischer, T. Zigart, C. Schmidbauer, und S. Schlund, „Learning Nuggets for Cobot Education: A Conceptual Framework, Implementation, and Evaluation of Adaptive Learning Content“, *SSRN Journal*, 2021, doi: 10.2139/ssrn.3868713.

[12] W. H. Bommer, J. L. Johnson, G. A. Rich, P. M. Podsakoff, und S. B. Mackenzie, „On the interchangeability of objective and subjective measures of employee performance: a meta-analysis“, *Personnel*

Psychology, Bd. 48, Nr. 3, S. 587–605, Sep. 1995, doi: 10.1111/j.1744-6570.1995.tb01772.x.

[13] H. Clasen, „Die Messung von Lernerfolg: eine grundsätzliche Aufgabe der Evaluation von Lehr-bzw. Trainingsinterventionen“, *Technischen Universität Dresden*, 2010.

[14] Kirkpatrick D. L., „Techniques for Evaluation Training Programs“. J. Am. Soc. Train. Dir, 1959.

[15] J. Kirkpatrick, „An Introduction to The New World Kirkpatrick Model“, *Kirkpatrick Partners*, 2021.

[16] N. A. A. S. Administration, „NASA TLX Task Load Index“, *https://humansystems.arc.nasa.gov/groups/tlx/*, 2020.

[17] A. Frick-Salzmann, *Gedächtnis: Erinnern und Vergessen*. in essentials. Wiesbaden: Springer Fachmedien Wiesbaden, 2017. doi: 10.1007/978-3-658-16720-2.

[18] H. Ebbinghaus, „Memory: A Contribution to Experimental Psychology“, *ANS*, Bd. 20, Nr. 4, Okt. 2013, doi: 10.5214/ans.0972.7531.200408.

[19] D. L. Kirkpatrick und J. D. Kirkpatrick, *Implementing the four levels: a practical guide for effective evaluation of training programs*. in Read How You Want. San Francisco: Berrett-Koehler Publishers, 2007.

[20] Preparedness and Emergency Response Learning Centers, „Kirkpatrick Level 3 (Behavior) Evaluation Strategies“, *http://www.phf.org/programs/preparednessresponse/evaluationrepository/Documents/Kirkpatrick_Level_3_Behavior_Evaluation_Strategies.pdf*, missing year.

[21] R. S. Kaplan und D. P. Norton, „Balanced Scorecard“, in *Das Summa Summarum des Managements*, C. Boersch und R. Elschen, Hrsg., Wiesbaden: Gabler, 2007, S. 137–148. [Online]. Verfügbar unter: 10.1007/978-3-8349-9320-5_12

[22] D. L. Kirkpatrick und J. D. Kirkpatrick, *Transferring learning to behavior: using the four levels to improve performance*, 1st ed. San Francisco, Calif.: Berrett-Koehler Publishers, 2005.

[23] F. D. Davis, „Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology“, *MIS Quarterly*, Bd. 13, Nr. 3, S. 319–340, 1989.

[24] F. Weng, R.-J. Yang, H.-J. Ho, und H.-M. Su, „A TAM-Based Study of the Attitude towards Use Intention of Multimedia among School Teachers“, *ASI*, Bd. 1, Nr. 3, S. 36, Sep. 2018, doi: 10.3390/asi1030036.

[25] P. A. Ratna und S. Mehra, „Exploring the acceptance for e-learning using technology acceptance model among university students in India“, *IJPMB*, Bd. 5, Nr. 2, S. 194, 2015, doi: 10.1504/IJPMB.2015.068667.

[26] M. Georgsson, „NASA RTLX as a Novel Assessment Tool for Determining Cognitive Load and User Acceptance of Expert and User-based Usability Evaluation Methods“, *European Journal of Biomedical Informatics*, 2020.

[27] J. Brooke, „SUS - a Retrospective“, *Journal of Usability Studies*, Bd. 8, Nr. 2, S. 29–40, 2013.

[28] F. F. Reichheld, „The One Number You Need to Grow“, *Havard Business Review On Point*, Bosten, 2003.

[29] A. D. Prabaswari, C. Basumerda, und B. W. Utomo, „The Mental Workload Analysis of Staff in Study Program of Private Educational Organization“, *IOP Conf. Ser.: Mater. Sci. Eng.*, Bd. 528, Nr. 1, S. 012018, Mai 2019, doi: 10.1088/1757-899X/528/1/012018.

[30] Jr. J. Hair, R. E. Anderson, R. L. Tatham, und W. C. Black, *Multivariate Data Analysis*, 5th edition. India: Pearson Education, 2003.

[31] D. L. Streiner, „Starting at the Beginning: An Introduction to Coefficient Alpha and Internal Consistency“, *Journal of Personality Assessment*, Bd. 80, Nr. 1, S. 99–103, Feb. 2003, doi: 10.1207/S15327752JPA8001_18.

[32] N. Wenk, J. Penalver-Andres, K. A. Buetler, T. Nef, R. M. Müri, und L. Marchal-Crespo, „Effect of immersive visualization technologies on cognitive load, motivation, usability, and embodiment“, *Virtual Reality*, Bd. 27, Nr. 1, S. 307–331, März 2023, doi: 10.1007/s10055-021-00565-8.

[33] M. K. Othman, A. Nogoibaeva, L. S. Leong, und M. H. Barawi, „Usability evaluation of a virtual reality smartphone app for a living museum“, *Univ Access Inf Soc*, Bd. 21, Nr. 4, S. 995–1012, Nov. 2022, doi: 10.1007/s10209-021-00820-4.

[34] H. Lovasz-Bukvova, M. Hölzl, G. Kormann-Hainzl, T. Moser, T. Zigart, und S. Schlund, „Usability and Task Load of Applications in Augmented and Virtual Reality: How Applicable are the Technologies in Corporate Settings?“, in *Systems, Software and Services Process Improvement*, M. Yilmaz, P. Clarke, R. Messnarz, und M. Reiner, Hrsg., in Communications in Computer and Information Science, vol. 1442. Cham: Springer International Publishing, 2021, S. 708–718. doi: 10.1007/978-3-030-85521-5_48.

[35] E. Marino, L. Barbieri, B. Colacino, A. K. Fleri, und F. Bruno, „An Augmented Reality inspection tool to support workers in Industry 4.0 environments“, *Computers in Industry*, Bd. 127, S. 103412, Mai 2021, doi: 10.1016/j.compind.2021.103412.

[36] P. T. Kortum und A. Bangor, „Usability Ratings for Everyday Products Measured With the System Usability Scale“, *International Journal of Human-Computer Interaction*, Bd. 29, Nr. 2, S. 67–76, Jan. 2013, doi: 10.1080/10447318.2012.681221.