Virtual Learning Environment for an Industrial Assembly Task

Christian Hirt
ETH Zurich
Innovation Center Virtual Reality
Zurich, Switzerland
hirtc@ethz.ch

Valentin Holzwarth
University of Liechtenstein
Institute of Information Systems
Vaduz, Liechtenstein
valentin.holzwarth@uni.li

Joy Gisler
ETH Zurich
Innovation Center Virtual Reality
Zurich, Switzerland
gj@ethz.ch

Johannes Schneider
University of Liechtenstein
Institute of Information Systems
Vaduz, Liechtenstein
johannes.schneider@uni.li

Andreas Kunz
ETH Zurich
Innovation Center Virtual Reality
Zurich, Switzerland
kunz@iwf.mavt.ethz.ch

Abstract—In this paper, we present a virtual learning environment for an industrial assembly task, which combines an easy-to-use interaction with an intuitive user experience. It is shown that such a virtual environment can be used for initial training to introduce tasks to new employees, but also experts may benefit from advanced training in case of new products, or new assembly routines. Consequently, this application was validated twice. First, in a lab pilot study showing that the simple interactions and helpful instructions were appreciated by the participants. Second, professionals from industry were asked to perform the task and to evaluate the usefulness of the virtual learning environment considering its industrial applicability. Based on the achieved scores in common evaluative questionnaires and the post-study interviews, both, the lab pilot and the industrial study, have performed well and will be further developed in close collaboration with the industrial partner.

Index Terms—Virtual Reality, Industrial Application, Assembly Training

I. INTRODUCTION

Industrial trends like mass customization and customer oriented design lead to increasingly knowledge-intensive and highly specific assembly tasks. In order to fulfill these tasks, assembly workers have to undergo repetitive, time-consuming, and expensive training processes [5]. Such training processes traditionally involve print-out manuals, instructional videos, and hands-on exercises with experts [7]. Besides being cost-intensive and time-consuming, traditional training processes comprise high effort in preparation, as changes in the assembly routine require new instructional material to be produced and distributed.

A promising solution to the issues related to traditional training processes is training in virtual environments, which allows for low-cost setups and fast adjustments to changing tasks or environments. A virtual environment can either be realized employing virtual reality (VR) or mixed reality (MR) technologies. VR describes a synthetic environment that fully immerses a user by wearing a head-mounted display (HMD). MR describes an environment, in which a user experiences both, virtual objects and the real world, in a combined form, e.g. by wearing mixed reality glasses or using a tablet computer. While VR and MR technologies are appropriate for industrial assembly training [5], VR technologies offer the decisive advantage of a comparably inexpensive consumer-grade hardware. In contrast, currently available MR hardware comes with multiple issues such as tracking inaccuracy, low display resolution, or narrow field of view. Additionally, fully immersive virtual environments are more favorable for training the manual assembly of larger scale and complex products, where workers have to access intricate or dangerous spaces [6]. This favor towards an immersive virtual learning environment (VLE) is also confirmed by applications in other fields. An example for such an application field is the medical sector, where fully immersive VLEs are acceptable solutions for learning new tasks e.g. myocardial infarction treatment training [10], and for repetitive training of previously learned tasks to increase task fulfillment performance or regulation compliance e.g. hand-hygiene training [11]. Besides the described advantages of VLEs, their independence of physical workplaces and their mobility make them further interesting for industry.

However, most VLEs have not proven their efficiency in industrial use cases yet [3], e.g. whether they help detecting possible design errors by a hands-on experience, and whether employees could learn from these mistakes. Furthermore, most studies on VLEs do not include product designers as a target group, which is crucial, since they have a major influence on how the product will be assembled in subsequent steps.

This paper thus describes a real industrial use case and how such a VLE could improve workers’ efficiency and also raise the designers’ awareness on product assembly. For this, we organize the paper as follows: In Section II, we review existing literature on VLEs for assembly training, followed by a description of the study design in Section III. Section
IV presents the results of a lab study and an industry study, conducted to evaluate the VLE developed within this work. Section VI concludes this paper by reflecting its key findings, while Section VII shows future research to be conducted.

### II. RELATED WORK

Research on VLEs for assembly tasks spans across multiple domains such as engineering [6], [8], computer science [2], [7], [9] and education [5], [12].

We compare the most relevant and recent works in VLE research among five categories, which are presented in Table I. First, we subdivide the category technology into MR and VR, since both technologies are appropriate for a VLE [5]. Second, we identify the target group, which could be in the case of an industrial assembly task either worker or designer. Third, we distinguish the training mode of the VLE by initial, dedicated to employees being unfamiliar (c.f. [10]) to the task, and advanced, dedicated to employees being familiar with industrial assembly tasks. Fourth, it is important for the evaluation of a VLE to consider both, subjective parameters, i.e. self-reported data of study participants on usability, and objective parameters, i.e. measured data such as task completion time and errors [3]. Fifth, the setting covers an industrial task, since virtual learning environments should be evaluated within an authentic setting, i.e. an actual assembly task conducted by real workers [12] instead of a lab setting. Additionally, we state the number of participants for each of the reviewed studies, indicated as Sample Size in Table I.

Within the reviewed literature, VR is the dominant technology with nine out of ten studies employing it within their VLE. One study solely relies on a MR-based VLE, wherein participants were required to assemble a gully trap on an actual workplace, either with paper-based or MR-based instructions [4]. However, the set-up of such a study requires a physical workplace, which would not be needed in an immersive VLE. The sole target group of all VLEs in the sample except one is assembly workers. However, designers also need to be included as an additional target group due to their major influence on the industrial assembly process. This hypothesis is supported by one study within the sample, where participants first had to model and then assemble a bearing onto a spindle in a VLE [8]. Such a VLE is dedicated to product designers as an environment for immersive product and assembly reviews. The dominant training mode of most of the reviewed studies with seven out of ten is initial training, i.e. participants had no prior experience in industrial assembly. Three works dedicate the purpose of their VLE towards advanced training, i.e. participants with prior experience in industrial assembly were trained to improve task fulfillment performance. A VLE dedicated to advanced training should be evaluated in an authentic setting, with employees that have at least medium experience in assembly operations [5]. Further, the evaluation of a VLE should be both objective, i.e. measured values such as task completion time, and subjective, i.e. self-reported data of participants on usability [3]. However, more than half of the studies in the sample (6/10) either evaluated their VLE objectively or subjectively. Since the context in which a VLE is evaluated has a major influence, a VLE should be not only evaluated in a lab setting, but also in an authentic setting, i.e. with a realistic task and real employees. Most of the retrieved studies (6/10) were conducted in labs, whereas three studies were conducted in an authentic setting. The study on gully trap assembly [4] is indicated in parentheses in Table I, since it partially fulfills the criteria of an authentic setting, by investigating a realistic task, but without integrating real employees. So far, there is no subjective and objective evaluation of a VLE that provides an authentic setting for advanced training of designers.

### III. STUDY DESIGN

The study was designed to respond to the need for a VLE, wherein designers can be trained in the assembly of the products they develop. The VLE was evaluated subjectively and objectively in our lab as well as in an authentic setting on our industrial partner’s site. Therefore, a multi-phase study design was chosen, in which the participants had to switch between filling out questionnaires and performing virtual tasks. Each participant had to undergo all phases, which took about 45 minutes to complete.

The VLE was realized using an HTC Vive Pro with both controllers, calibrated to a room-scale setting, allowing for free movement in a 4x4m space. The VLE extends over this room-

### TABLE I: Overview on recent work in VLE research

<table>
<thead>
<tr>
<th>Study</th>
<th>Sample Size</th>
<th>Technology</th>
<th>Target Group</th>
<th>Training Mode</th>
<th>Evaluation</th>
<th>Settling</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smart factory assembly [1]</td>
<td>10</td>
<td>MR</td>
<td>Worker</td>
<td>Initial</td>
<td>Advanced</td>
<td>Subjective, Objective, Authentic</td>
</tr>
<tr>
<td>Medical device assembly [2]</td>
<td>30</td>
<td>VR, MR</td>
<td>Designer</td>
<td>Initial</td>
<td>Advanced</td>
<td>Subjective, Objective, Authentic</td>
</tr>
<tr>
<td>Car door assembly (1) [3]</td>
<td>30</td>
<td>VR</td>
<td>Worker</td>
<td>Initial</td>
<td>Advanced</td>
<td>Subjective, Objective, Authentic</td>
</tr>
<tr>
<td>Car door assembly (2) [3]</td>
<td>30</td>
<td>VR</td>
<td>Worker</td>
<td>Initial</td>
<td>Advanced</td>
<td>Subjective, Objective, Authentic</td>
</tr>
<tr>
<td>Gully trap assembly [4]</td>
<td>30</td>
<td>VR</td>
<td>Worker</td>
<td>Initial</td>
<td>Advanced</td>
<td>Subjective, Objective, Authentic</td>
</tr>
<tr>
<td>Electronics assembly [5]</td>
<td>40</td>
<td>VR</td>
<td>Designer</td>
<td>Initial</td>
<td>Advanced</td>
<td>Subjective, Objective, Authentic</td>
</tr>
<tr>
<td>Haptics-based assembly [6]</td>
<td>39</td>
<td>VR</td>
<td>Worker</td>
<td>Initial</td>
<td>Advanced</td>
<td>Subjective, Objective, Authentic</td>
</tr>
<tr>
<td>Model assembly [7]</td>
<td>39</td>
<td>VR</td>
<td>Worker</td>
<td>Initial</td>
<td>Advanced</td>
<td>Subjective, Objective, Authentic</td>
</tr>
<tr>
<td>Part design and assembly [8]</td>
<td>30</td>
<td>VR</td>
<td>Worker</td>
<td>Initial</td>
<td>Advanced</td>
<td>Subjective, Objective, Authentic</td>
</tr>
<tr>
<td>Transmission assembly [9]</td>
<td>30</td>
<td>VR</td>
<td>Worker</td>
<td>Initial</td>
<td>Advanced</td>
<td>Subjective, Objective, Authentic</td>
</tr>
<tr>
<td><strong>This study</strong></td>
<td><strong>21</strong></td>
<td><strong>VR</strong></td>
<td><strong>Worker</strong></td>
<td><strong>Initial</strong></td>
<td><strong>Advanced</strong></td>
<td><strong>Subjective, Objective, Authentic</strong></td>
</tr>
</tbody>
</table>
scale setting and thus, an out-of-the-box SteamVR teleporting is utilized. In order to grab virtual components, SteamVR’s grab function is adapted to the task.

The virtual environment was created in Unity and consists of three scenes (i.e. the Tutorial, the Impossible Task, and the Completion Task) which had to be experienced sequentially by all participants. The VLE’s overall task required the participants to assemble a ceiling-mounted installation system from three different types of extruded profiles, two types of rails, and one angle type. The assembly information was given by a 2D layout plan, which had to be followed by the participants (see Fig. 1). During the Tutorial scene, participants were instructed on how to use the controllers, how to move (i.e. walking and teleporting), how to interact with the virtual environment, and how to assemble the various components. In the second scenario, the Impossible Task, the assembly instruction provided to the participants was intentionally faulty such that the assembly could not be finalized with the available extruded profiles. Accordingly, only in the third scenario, the Completion Task, a successful assembly was possible and the ceiling-mounted installation system could be completed (see Fig. 2a).

1) Phase 1 – Introduction and Pre Questionnaires: In Phase 1, participants filled out an initial questionnaire consisting of personal information regarding their background and expertise both in VR and industrial assembly, followed by an initial simulator sickness questionnaire (SSQ) to retrieve a participant’s personal entry condition. Next, the assembly task was explained using 3D cardboard models of the extruded profiles and multiple images describing how (based on position and orientation) the single pieces could be assembled.

2) Phase 2 – VLE Tutorial: In Phase 2, the participants were guided to the VLE and a tutorial scene was started. In this scene, the instructor explained the interaction possibilities to the participants. In addition to the preexisting functionalities like walking, teleporting, and grabbing, the VLE incorporates the additional custom-made interaction of assembling the components. Single assembly components can be put together by holding them close to adjacent components. Whenever a loose component is allowed to be assembled (i.e. the components fit together and could be assembled in reality), they are highlighted in orange. Additionally, if the component was brought to the right position and orientation according to the layout plan, it was highlighted green to indicate its correct placement (see Fig. 2b). Once a piece is highlighted in either colour, it is attached to its counterpart as soon as the participant releases the grab button. Each participant interacted with each of the available components and connected them to understand how the components interact with each other.

3) Phase 3 – VLE – The Impossible Task: In Phase 3, participants were provided with the actual assembly task in the VLE. In the Impossible Task, the assembly is not entirely possible, since the provided layout plan was consciously designed faulty. However, even though the plan has a flaw (i.e. the central rail does not allow left- and right-handed connections simultaneously), the virtual assembly pieces are according to the plan. The main goal of the Impossible Task for the participant is to find the previously mentioned flaw in the layout plan and notify the study instructor about this issue. After pointing out the error, participants are instructed to take off the HMD and the situation is explained to them.

4) Phase 4 – VLE – The Completion Task: In Phase 4, the participants were provided with the same task as in Phase 3, but this time, the components fit together and the assembly can be completed (see Fig. 2a).

5) Phase 5 – Post Questionnaires and Interview: In Phase 5, participants filled out a second questionnaire, which consisted of the post SSQ, Task Load Index (TLX), System Usability Scale (SUS), presence, followed by an open interview. Accordingly, participants were able to further describe their experience and thoughts beyond the questionnaires.

Measurements
Beside background information about the participants and the mentioned subjective questionnaires, an objective
measurement was performed. First, the task completion time (TCT) was measured for each participant in Phases 3 and 4. The start time for both scenes was taken at the moment the participant entered the VLE. The end time for Phase 3 was defined as the moment when a participant correctly announced the error, while for Phase 4, the final timestamp was taken when the participant completed the instructed assembly. Additionally, measures (i.e. objective handling errors) consisted of the number of components that participants dropped and the number of components that were incorrectly placed during the assembly task.

Setting
The study was conducted twice, once with students and scientific staff in the lab, and in an industrial environment addressing designers and other professionals closely related to this particular assembly task. The first run, the lab study, was conducted mainly to test the feasibility of the interaction with the single components and the assembly itself. Since the lab study only involved participants who were unfamiliar with the assembly, it was expected that they would focus more on the interaction’s natural and intuitive feeling. In the second run, the industry study, the study was conducted on the site of the industrial partner. Only professionals familiar with the assembly of installation systems were chosen as participants. In both runs, the setups were prepared identically.

IV. RESULTS AND COMPARISON
In this section, we present and discuss the results found in the two runs: the lab study and the industry study. The most relevant values are shown in Table II.

A. Results: Lab Study
The lab study was conducted with 13 participants (scientific staff, students and interns) that were 28.85±2.32 years old (M±SD). Many participants had prior experience with VR (8/13). The SSQ score on average slightly decreased from pre to post questionnaire. Presence was assessed to be 5.47/7 on average. Objective measurements resulted in a TCT of 447.58s on average for the Impossible Task, and 483.67s for the Completion Task.

B. Results: Industry Study
The industry study consisted of eight participants, from a variety of assembly related professional backgrounds (designer, design project coordinator, design technician, logistician), and 33.37±4.82 years of age (M±SD). Most participants had prior experience with VR (6/8). The SSQ score also slightly decreased from pre to post questionnaire on average. Presence was assessed to be 5.39/7 on average. Objective measurements
resulted in a TCT of 286.63s on average for the Impossible Task, and 403.25s for the Completion Task.

C. Comparison

Notable differences are shown in Table II. According to the TLX evaluation, the mental demand and effort are perceived slightly higher in the lab study than in the industry study, however, for both not statistically significant. On the other hand, the difference in the values between the lab and the industry study for the perceived performance and frustration are both statistically significant with a one-tailed t-test ($\alpha = 0.05$) with $t_{perf} = 0.0118 < \alpha$ resp. $t_{frust} = 0.0028 < \alpha$. Further notable differences were observed in the objective measurements. TCTs for both, the Impossible Task and the Completion Task, were lower in the industry study as shown in Fig. 3. In the Impossible Task, professional designers were 36% faster in noticing that the layout plan is faulty. Notably, all participants out of both target groups identified the faulty plan in the Impossible Task. For the Completion Task, professional designers assembled the complete system 17% faster on average. For the object handling errors, none of the groups performed notably better.

V. DISCUSSION

The results obtained in the evaluation of the VLE were generally very positive. The overall sample size of 21 participants is comparable to other studies (see Table I).

A. SSQ Evaluation

An evaluation of the SSQ for all participants shows that the SSQ score decreased on average. In cases where the SSQ score for a participant did increase at all, it was mainly attributed to an increase in self reported sweating. Given that both studies were conducted on hot summer days with outside temperatures well above 25°C, combined with the inherently given heat source of the HTC Vive Pro, this seems reasonable. Overall, the observed decrease in the SSQ scores indicates that the participants felt comfortable using our VLE and that it did not induce simulator sickness.

B. Feeling of Presence

The SUS presence score with 5.44/7 is rather high which confirms that the system is easy and intuitive to use. However, during the interview in Phase 5, all participants mentioned aspects of the VLE that could be enhanced and extended, when asked if they could imagine any further improvement of the VLE. Specifically, two participants mentioned auditory and five participants mentioned haptic feedback as favorable features to be added to future iterations of the VLE. Further, in terms of hardware, two participants complained about cable management and five participants about resolution. However, complaints about the resolution were unexpected, as the HTC Vive Pro offers full HD resolution for each eye. On the software side, a participant mentioned that even though the assembly components looked realistic, they sometimes behaved in a strange way, which is however part of Unity’s physics engine. Furthermore, the virtual environment was perceived to be too empty and artificial for another participant, since a construction site, where the assembly task is carried out in reality, would be more lively, open and more “under construction” than our virtual training room.

C. Familiarity

There is a noticeable difference in four TLX measurements between the lab and the industry study. Mental demand, required effort and frustration level are perceived much less in the industrial environment. Further, this is also visible in the TCT, in which the industrial employees performed 36% respectively 17% faster than the scientific staff. Most likely, this is closely related to the professionals’ familiarity with the assembly components used in the VLE. Additionally from the questionnaires, we know that the VR expertise shows a slight advantage towards the industrial employees.

D. Usability

The overall SUS score of 85.24/100 is very high, which can be also referred to as excellent usability [13]. Participants in the industry study assigned higher scores (88.43) than participants in the lab study (83.27). This indicates that the VLE’s usability is even higher for professionals familiar with assembly tasks.

VI. CONCLUSION

In this paper, we presented a VLE to train industry professionals in assembly tasks. Designers were specifically chosen as the main target group, since they have critical influence on a product’s assembly, but rarely get the opportunity to carry out such an assembly task by themselves. The evaluation, which was carried out using subjective and objective criteria, showed that the VLE is easy and intuitive to use and does not induce simulator sickness. Further, all participants were able to complete all three scenarios without help from the instructor. Additionally, the VLE has proven to be suitable for initial and advanced training, as the lab and industry study both yielded similarly positive results. The feedback obtained through the interviews was positive and constructive, which suggests that a more extensive and elaborate VLE would be appreciated.
Future work is based on participants’ answers in the interviews of the user studies, and on insights from the development process. Immediate improvements mentioned by participants include that the VLE should contain more features and components. Besides previously mentioned aspects (see Section Feeling of Presence), users wished for a more complex task, more variation in the assembly components, and more realistic interactions with the components. For example, participants stated that it feels wrong to not use tools like screwdrivers to connect the components. Similarly, most professionals noticed that such an angle piece would never be fixed to a rail without properly measuring the distance between a reference and the point of fixation. They asked to further extend the catalog of available components such that an inventory can be accessed while assembling, from which the participants could choose what component they want to use next. Additionally, two participants mentioned that real rails would be labeled (i.e. yoke number, component length, etc.) before being delivered to the construction site. Considering the experience as a whole, one participant envisioned to extend this existing VLE towards gamification and create company-wide leaderboards with which different company branches can compete. The participants further proposed some quality improvements for the existing VLE. There were mainly two different aspects to that, which are that they would prefer to access the layout plan directly next to the final construction instead of on the desk with the components. Further, they would appreciate some feedback before misplacing a component.

From a long-term perspective, we will improve the study environment in close collaboration with the industrial partner. Here, it is envisioned to compare the efficiency of the VLE against a real-life learning setup. For this, a similar user study will be conducted, however, a wireless VR solution will then be utilized. In addition, the tracking space will be expanded with multiple lighthouses to cover at least 10x10m allowing for natural free walking in the whole environment. Furthermore, the efficiency of this VLE will be assessed by introducing a control group which will be trained with conventional instructional material. Another major improvement will be the introduction of a second user. This can either be realized as a second person actually learning in the given VLE, but also by creating a teacher-student relationship between two users.

Simultaneously, different long-term approaches will feature applications related to the prototype presented in this paper. A first development will aim at a sandbox approach which includes a bigger variety of assembly components and an automatic layout conversion. This will allow a designer to construct any installation system in VR which will then be translated into a layout plan. In this particular way, planning errors can be eliminated since the physics-based virtual environment refuses to put components together, which do not match. A second approach will be to use a similar virtual environment as a simple visualization tool. Automatically generating and visualizing a designed installation system from a layout plan will immediately help designers and planners to understand, whether this particular system can be assembled given the constraints (e.g. space) of the construction site. These accessibility issues would otherwise only be identified on the construction site and may lead to additional costs. Further, if such a virtual environment was used in early planning phases, it would additionally deliver a common ground for communication between different departments or companies, which will build upon such an installation system. Otherwise, if these planning steps are realized in parallel, there may be communication issues leading to unnecessary planning iterations. Accordingly, eliminating such communication issues early on minimizes the number of planning iterations and thus brings an immense financial benefit.

ACKNOWLEDGMENTS
This work was supported by a project financed by the Hilti Family Foundation in Schaan, Liechtenstein and RhySearch in Buchs SG, Switzerland. Additionally, the authors would like to thank Hilti Schweiz AG for providing the opportunity and the staff to conduct a user study in an authentic setting.

REFERENCES