Immersive virtual reality for procedural training: comparing traditional and learning by teaching approaches

(Article begins on next page)

Immersive Virtual Reality for Procedural Training: Comparing Traditional and Learning by Teaching Approaches

Federico De Lorenzis,[a], Filippo Gabriele Pratticò[a], Maurizio Repetto[a], Enrico Pons[a], Fabrizio Lamberti*,[a]

[a] *Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy*

Abstract

Virtual Reality (VR) has been widely adopted for the creation of training experiences, since it appears to allow overcoming, especially for practical training, some of the limitations of real exercises. Many previous works focused on investigating the effectiveness of VR Training Systems (VRTS) in a variety of fields, but the efficacy of these systems is very task-dependent, and the best way to integrate them into existing learning paths has yet to be thoroughly investigated. The goal of the present paper is to explore the latter aspects considering a case study in the context of energy management in industry, and focusing on a measuring procedure that is part, e.g., of energy audit inspections. To this aim, a VRTS was developed and used to conduct two user studies: a first study designed to investigate the effectiveness of the devised system when used as an alternative or in combination with lecture and laboratory-based teaching experiences, and a second study aimed to compare two different approaches (traditional learning, TL, and learning by teaching, LBT) for exploiting the provided VR-based functionalities in a learning path. Experimental results showed that the use of the VRTS alone improved the participants' performance compared to traditional experiences, and that LBT proved to be more effective that TL for practical learning purposes.

Keywords: virtual reality, learning by teaching, procedural training

*Corresponding author

Email address: `fabrizio.lamberti@polito.it` (Fabrizio Lamberti*,)

1. Introduction

When it comes to delivering procedural learning content, integrating lectures with hands-on exercises is essential to let the learners acquire the practical abilities that are required on the job. However, the possibility to organize efficient practical exercises is heavily task-dependent, and may get compromised when the learning activity can expose the participants to potential risks. Moreover, practical exercises can also require expensive, fragile, or rare equipment, which could be difficult to secure for real-life training experiences. Because of these limitations, practical exercises may often reduce to low-fidelity simulations, or even be excluded from the learning path and get replaced with demonstrations performed directly by an instructor. One of the fields that face this type of problems is training in the context of risk management, since first-responders need to be ready to face dangerous situations in real life, but cannot be directly exposed to risks during exercises (thus, for instance, firefighting courses usually offer practice sessions without real fires). Another field that is particularly affected by these issues is that of industrial training, where operators must be prepared to work with potentially hazardous systems like robotic manipulators or other electrical machinery, but cannot exercise directly on real equipment for safety reasons.

The limitations above could be addressed by leveraging Virtual Reality (VR). In fact, thanks to the developments made over the years, this technology has now the potential to create highly-detailed simulations of real-life scenarios that, given also the widespread availability of low-cost devices, can be easily exploited to safely support training delivery. In particular, VR-based Training Systems (VRTSs) have been the subject of many studies investigating the training efficacy of this technology, and have been adopted in various fields (Lamberti et al. (2021); Conges et al. (2020)).

There are, however, different approaches to integrate a VRTS into a learning path. It is therefore fundamental to know what is the best way to leverage such systems from the point of view of learning effectiveness and learners' experience. Unfortunately, works in the current literature generally focus on presenting a VRTS for a particular domain and, in some cases, perform comparisons with alternative training approaches (rarely more than one), but do not provide results of further analyses on the possible strategies to exploit them for maximizing the pedagogical outcome. For instance, a recent study that reviewed the use of VRTSs for industrial training (Naranjo et al. (2020)) includes many examples of works like Mirauda et al. (2020), which showed the usability and the effectiveness of the developed tools without actually comparing against existing methods. When such a comparison against a real course is performed like in Lacko (2020), the VRTS is used as a standalone tool, and other ways to integrate it in a pedagogical experience (other than replacing it) are not considered.

To investigate these aspects, the present paper considers a case study in the context of energy management, focusing on the measuring procedures performed on an electrical switchboard that are part of an energy audit or power management inspection. The inspection aims to quantify the energy consumption of an electrical system, with the goal of supporting the optimization of its efficiency. The steps involved in the inspection are usually the subject of industrial training, though there are also academic courses on electrical systems that include the said procedures as part of their program. In both the fields, due to the potential risks associated with working directly on an electrical switchboard and to the difficulty in accurately simulating the measuring tools and the switchboard behavior in a real-life simulation, the training is typically delivered through lectures or passive laboratory demonstrations where the learners can observe the instructor, but are not allowed to intervene in the operations. Thus, such a scenario could benefit from the use of VR: for instance, a VRTS could be used to let the learners operate on a virtual switchboard on which they could get familiar with the measuring procedure and be assessed on it without being exposed to any risk.

The present paper focuses on the above training scenario, reporting on a study aimed to address two research questions (RQs):

- **RQ1.** In which way a VRTS could be profitably used (as an alternative, or combined with the existing activities) into a learning path to improve conventional lectures and laboratory experiences?

- **RQ2.** As a follow-up to RQ1, i.e., considering the most valuable approach for leveraging VR in a learning path, is a learning by teaching (LBT) approach based on VR as effective as it proved to be in non-VR contexts?

The answers to the above RQs were obtained by conducting two consecutive user studies. The first study compared a traditional, laboratory-based training experience, the use of a VRTS, and a mixed activity that combined the laboratory experience and the use of a VRTS. The second study compared the best VR-based alternative resulting from the above analysis with a

LBT approach, in which the learners used the VRTS to learn the procedure and then guided other subjects through the simulated experience.

2. Related Works

As anticipated, VR technologies and, especially, immersive VRTSs (Kaplan et al. (2020)) have been largely studied and widely adopted for the creation of various training experiences (Jensen and Konradsen (2018); Checa and Bustillo (2020); Pellas et al. (2020)). In procedural training contexts, these systems proved to be more effective over printed (Buttussi and Chittaro (2021)) and video-based materials (Lovreglio et al. (2021)), and are typically considered when traditional, practical learning activities suffer from logistical limitations or are too dangerous to be performed (for training surgeons, in Kaluschke et al. (2018), pilots, in Rantanen and Talleur (2005), etc.).

As a matter of example, the adoption of VRTSs has been particularly prolific for training first responders like civil protection operators or firefighters. For instance, in De Lorenzis et al. (2022), the authors presented an application for training operators to manage hydrogeological emergencies. The devised system focused on the interactions between the learners and the instrumentation needed to complete the considered procedures, such as tubes, hoses, and heavy machinery (a high-capacity pump), and offered two training modes (one for training and one for assessment) that can be used by the learners to train safely and without the need for an instructor. A user study showed that the use of the VRTS had a positive impact on the learners' performance, since it was associated with a significant learning gain, although no direct comparisons with existing training were conducted to evaluate the efficacy. Cakiroğlu and Gökoğlu (2019) described a VRTS targeted to firefighters that could be used to teach the necessary concepts related to a fire safety procedure, train the subjects on it, and finally assess their performance. After the experience in VR, the learners were evaluated in a real-life scenario with a controlled fire, and it was shown that the use of VR significantly improved the training effectiveness.

Similar solutions have been leveraged also in the context of industrial training, where operators often interact with dangerous or expensive tools and systems, and must be well-prepared to avoid injuries and prevent damages to the instrumentation. For instance, Morillo Tena et al. (2020) performed a study to compare Augmented Reality training methods with video-based ones in the context of equipment maintenance. Although no significant differences were found, both the considered approaches helped the learners to acquire the necessary skills. Focusing on VR, the benefits associated with the use of a VRTS are exemplified in Pratticò and Lamberti (2021), where the authors presented a system for learning maintenance procedures on industrial robots. The system focused on self-learning, and was validated through a comparison with a traditional, instructor-based course. The results of the performed user study showed that the effectiveness of the VR experience was comparable to that of the traditional one.

Despite the above examples, there are also studies where a clear advantage associated with the use of VRTSs over traditional training was not evident. For instance, Gavish et al. (2015) failed to find significant differences between a VR-based and a traditional training, suggesting that this result could have been due to the level of expertise of the selected participants and concluding that further evaluations on the efficacy of VRTSs were needed. Another work casting doubts on the effectiveness of VR-based training is represented by Andaluz et al. (2017), where a VR application is intentionally designed as a complementary tool to existing training at the university in the field of industrial processes, moving from the hypothesis that a VR applications cannot be truly compared to traditional learning modules. Although evidence was not provided and VR technologies have largely advanced since the publication of the work, this is indeed another confirmation that more investigations are needed to confirm or refute positions about the role of VRTSs.

In fact, works like those above generally show that training activities involving VRTSs, despite being quite task-dependent and different from each other, can be superior (Cakiroğlu and Gökoğlu (2019)) or comparable (Pratticò and Lamberti (2021) to the traditional ones, at least when the learners are not familiar with the taught contents and are trained and assessed in the application (which can be quite a common situation); when not improving the learning outcome, they might have a positive impact on the overall learning experience (De Lorenzis et al. (2022)).

Previous studies, however, typically performed one-to-one comparisons and did not explore the various configurations in which a VRTS could be possibly exploited in a learning path (e.g., they rarely investigated the efficacy of mixed approaches involving both a traditional learning activity and the use of a VRTS). In particular, the studies that integrate a VR application in existing learning paths generally use the VRTS in a rather traditional way, by letting the learners train and possibly get assessed in the application, but do not investigate the use of VR in combination with other pedagogical approaches.

In non-VR training scenarios, a possible alternative to TL is LBT, which leverages peer-tutoring and lets learners teach each other self-learned contents. Despite being more time consuming and difficult to organize compared with TL, LBT is recognized to be more effective, e.g., in terms of knowledge retention (Roscoe and Chi (2007)). Nevertheless, it has not been studied yet in scenarios in which humans teach other humans in VR. An attempt to study this learning approach in combination with novel technologies was made in Pratticò et al. (2021), where a Mixed Reality training system supporting LBT was presented and tested, confirming its advantages. In particular, by comparing a TL approach with a LBT one, the authors found that, after one week, the subjects in the LBT group had preserved the taught concepts better than those in the TL group, confirming the effectiveness of LBT in terms of knowledge retention. However, since the system was designed to let the learners interact with a teachable agent (Biswas et al. (2005)), namely a robot acting as a substitute of the taught human peer, it is not immediate to conclude that the findings obtained in that setting would be relevant also in a scenario in which immersive technologies are used to support the teaching of human subjects. In fact, interacting with humans involves empathy and other social factors (Duran (2017)) that could not be easily replicated with a robot; therefore, further investigations are needed to evaluate the efficacy of a VR-based LBT pedagogical approach.

3. Methodology

According to the objectives set for this work, a VRTS was used to investigate the best way to integrate VR technologies into a learning path focused on energy inspections and to compare VR-based TL and LBT approaches. The VRTS exploited in this work was developed in the context of the E2DRIVER project.

3.1. Use Case

The considered use case is a measuring procedure performed on electrical switchboards. In this procedure, the operators use a measuring tool and connect it to the cables and conductors of a switchboard to record the current, voltage and power values over a period of time (Fig. 1).

Since the operators may work directly on active power lines, they are required to wear appropriate Personal Protective Equipment (PPE) and per-

(a) (b)

Figure 1: Details of the measuring activity: an operator (a) connects the probes to the electrical switchboard, and (b) configures the measuring instrument.

form the steps in the correct order to avoid possible risks associated with electric shocks and arcs.

Six main steps can be identified, which are described below.

- The operator opens the switchboard door and assesses the situation; he or she checks whether the power line to be measured (or target power line, TPL) is active and, if so, if there is a backup power line (BPL, not connected to an electrical utility) that can be used to measure the voltages. If the TPL is active, the corresponding circuit cannot be opened (otherwise the power supply is interrupted), and the voltage values must be measured directly on live contacts. If a BPL is present, it can be used to safely perform the voltage measurements since the corresponding circuit can be opened. The operator must also check whether the power lines are separated from each other.

- The operator follows the safety measures. If the TPL is active a BPL is not present, he or she must wear insulating gloves to protect from electric shocks, as well as a protective helmet with a visor to protect from electric arcs. He or she must keep the PPE for the rest of the procedure. If a BPL is present, the operator must open the corresponding circuit, wear the PPE and test the power line with a tester to check if the voltage is null; if the value is null, the BPL is safe to use, and the operator can take off the PPE. If the other lines in the switchboard are not separated, he or she must insulate them using insulating sheets.

- The operator connects the monitoring tool to the power supply, places it near the switchboard, and connects all the electrical probes to the tool. A total of 11 probes are used: three clamps to measure the current values and eight clips (also called crocodile clips) to measure the voltage values. The clamps must be chosen based on the cables diameter and the current range of the power line, whereas the crocodile clips are generic.

- The operator connects the probes to the switchboard. The power lines have four contacts (in the considered case named L1, L2, L3 and Neutral) and, if they are connected to an electrical utility, four cables (with the same names) departing the contacts. The operator must measure the voltage values of L1, L2 and L3 with respect to the Neutral, and the voltage values of the Neutral with respect to ground. He or she must also measure the current values for the L1, L2 and L3 cables. If a BPL is present, the operator connects the crocodile clips to the BPL contacts, and the clamps to the TPL cables. If a BPL is not present, the operator connects all the probes to the TPL.

- The operator sets up the monitoring tool by using its configuration interfaces through the touch screen. He or she must select the type of electrical probes used, the monitoring period and the sample interval. He or she can check whether the connections made are correct by looking at the corresponding phasors. The operator can then start the recording. Before leaving the room, he or she must signal the presence of exposed active conductors by placing a warning sign near the monitoring tool.

- The operator returns to the electrical room after the monitoring time has passed, undoes the connections, and collects the data.

3.2. VRTS

The developed VRTS is an immersive VR application that can be used both for training and assessment purposes. The VRTS was implemented using Unity 2019.4 and the SteamVR framework, and was originally designed to be used with an immersive VR system including a headset and a pair of hand controllers. All the 3D elements were modeled in Blender using either real objects or pictures of real objects as references. Even though it supports various consumer VR systems (like the Meta Quest 2), the application was specifically targeted to the HTC Vive Pro.

3.2.1. Implementation

In the VRTS, the learners experience the procedure and interact with an environment representing an electrical room. The learner can exploit the controllers to interact with 3D elements and can move in the virtual world either by walking in the real room or using the Arm Swing locomotion method, i.e., swinging his or her arms back and forth while standing in place and pressing a button on the controllers (Cannavò et al. (2021)). In the desktop, non-immersive version of the application, the learner can move using the keyboard and interact with the environment using the mouse.

A fictional scenario was created by reconstructing in 3D one of the electrical rooms at the authors' university. In the room, the learner can find a series of virtual elements (Fig. 2): an electrical switchboard with a series of power lines, the monitoring tool, two tables, the probes necessary for the procedure, a closet containing the PPE (a pair of insulated gloves and a protective helmet), a tester, and a warning sign.

Other, non-interactable elements (e.g., a transformer) were added to increase the realism of the simulation. All the elements were designed and implemented to replicate the actual behavior of their real counterparts.

- The electrical switchboard has a series of power lines (the TPL, the BPL, plus other lines that are not of interest for the procedure). The learner can freely interact with all the lines, placing and removing insulating elements, connecting and disconnecting probes, as well as turning on and off the associated switches. Interacting with the contacts of an active power line without wearing the PPE can lead to a shock effect inside the application.

- The monitoring tool offers all the functionalities required for the procedure, and provides a touch-based interface that can be navigated by the learner to make the necessary settings (selecting the probe type, the monitoring period, the sample interval, etc.). The learner can also exploit the interface to check the correctness of the connections made and change them acting on the probes to see the real-time updating of the corresponding phasors.

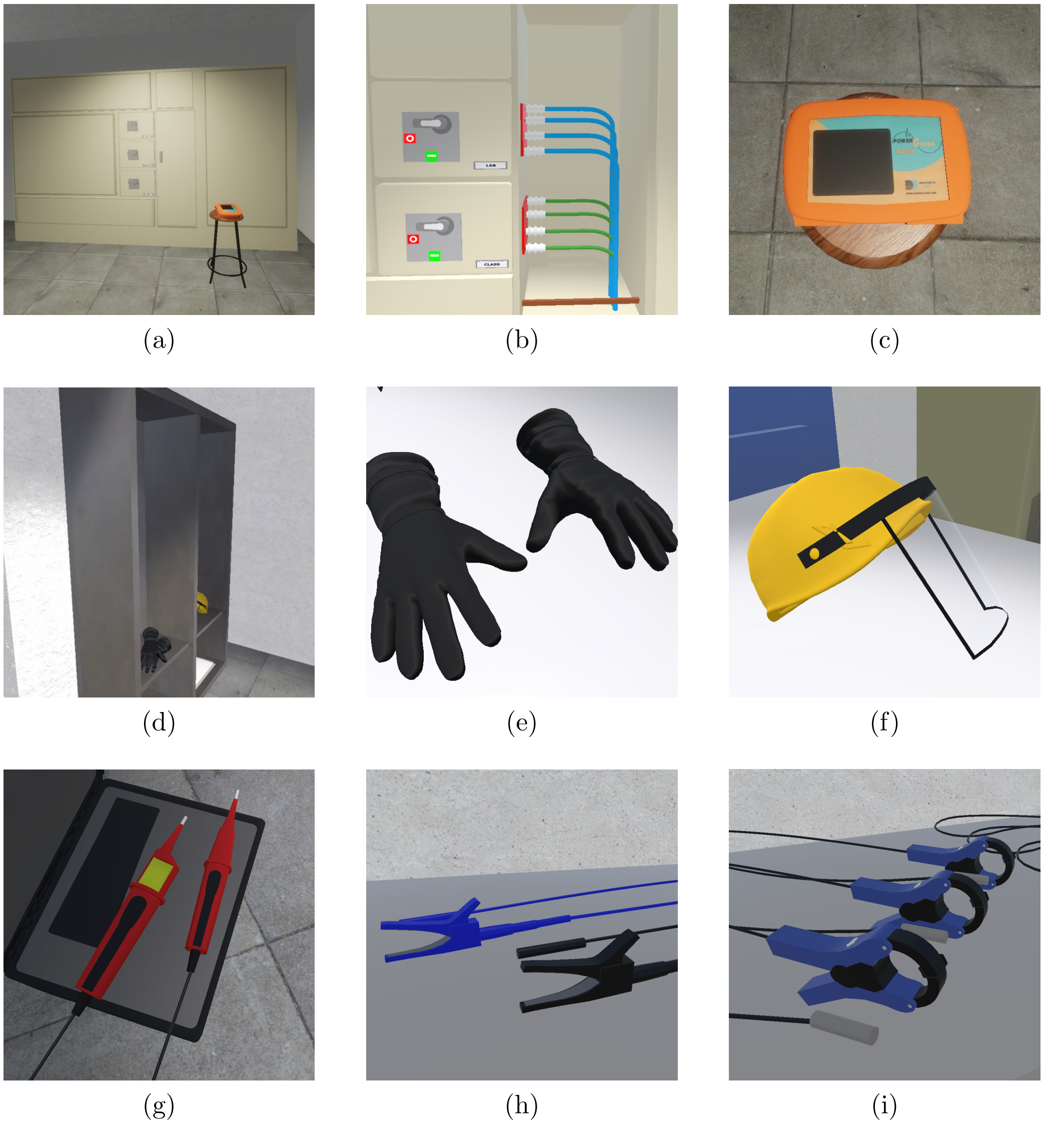


Figure 2: Virtual elements: (a) electrical switchboard, (b) power lines, (c) monitoring tool, (d) closet with the PPE, (e) insulated gloves, (f) protective helmet, (g) tester, (h) crocodile clips, and (i) clamps.

- The probes can be freely connected to the monitoring tool and to the switchboard. Eight crocodile clips are available on one of the tables and can be connected to the contacts of the power lines to perform voltage measurements; moreover, six clamps (of two different types) can be found on the other table and can be connected to the switchboard cables to perform current measurements. With the aim to improve the realism of the experience, the probes were implemented using an asset for the physical simulation of cables; this asset was further improved in this work to simulate different strength levels and add the possibility for a probe to detach from the point it was connected to if the learner stretches the cable excessively, thus replicating the real behavior in a better way.

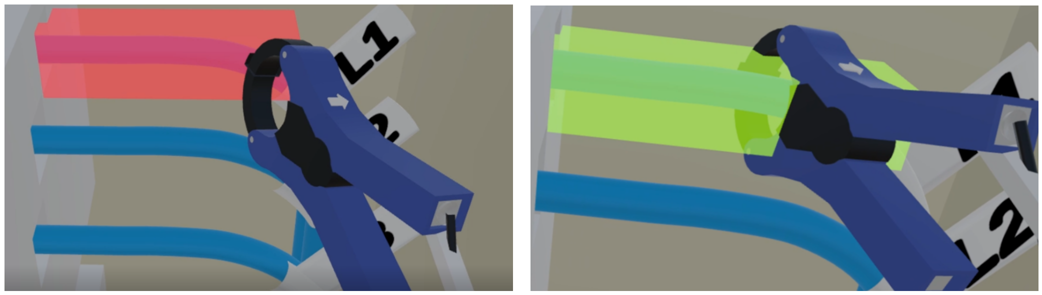
When the application is launched, the learner can choose between two possible scenarios.

- In the first scenario, the switchboard has three power lines (a BPL, a TPL and a third line that is not of interest for the procedure). All the power lines are insulated, and the learner can decide whether to follow the procedure working only on the TPL or using the BPL.
- In the second scenario, the switchboard has only two power lines, the TPL and a second line that is not of interest for the procedure. The power lines are not insulated, and the learner shall wear the PPE to perform the procedure.

3.2.2. Training Modes

The VRTS offers two training modes to teach the procedure and assess the learner's performance: a Guided Mode (GM) and an Evaluated Mode (EM).

The GM features a scaffolding system that guides the learner through the entire experience. The procedure is split in a series of micro-tasks, and the learner is instructed step-by-step on how to carry them out by a voice-over (available in five different languages) and several highlights emphasizing the virtual elements that are involved in the execution of the task. The highlights react to the learner's actions and change color if he or she is performing the correct action (Fig. 3). For each task, the learner is asked by the voice-over to perform a specific action (e.g., "place the monitoring tool near the switchboard"), and the simulation does not continue until the said task has been completed; if the learner makes a mistake, the voice-over repeats the last message and invites the learner to repeat the action. In the GM, the virtual elements are locked until they are mentioned by the voice-over in



(a) (b)

Figure 3: Highlights in the Guided Mode: (a) the red highlight shows where to place the clamp, (b) the green highlight shows that the clamp is in the correct position and has the correct orientation, thus can be connected to the cable.

order to reduce the chances of errors. The possible tasks to perform are moving an object, interacting with an object, finding an object in the scene, and making a connection. At any time, if the learner does not remember the current task to perform, he or she can ask the voice-over to repeat the last message by pressing a button on the controller. The GM was designed based on a real laboratory activity performed at the authors' university, with the support of professors of electrical systems.

The EM is a sandbox mode in which the learner can freely interact with all the elements in the virtual environment to replicate the steps of the procedure. Voice-over messages and corresponding highlights are disabled. Errors are not explicitly signaled, but the learner can spot incorrect actions by using the interface of the monitoring tool (e.g., if he or she makes a wrong connection, the tool will show a wrong value for a voltage or a current, or a phasor with wrong orientation) or other diegetic clues (e.g., if the learner tests the voltage of an active power line, the tester will signal a non-null value). When the learner leaves the virtual room, an integrated evaluation system uses the information collected during the experience to produce a report on his or her performance. The report is provided directly in the application at the end of the experience, and it is also saved in the application folder for future use. The assessment (and the report) is organized in five categories:

- Safety, to evaluate if the learner followed all the prescribed safety measures;
- Voltage connections, to evaluate if the crocodile clips were used on the correct power line;

- Current connections, to evaluate if all the electrical probes, of the correct type, were used on the correct power line;
- Connections logic, to evaluate if the crocodile clips and the electrical probes were used on the correct contacts and cables;
- Monitoring tool configuration, to evaluate if the learner configured the tool in the appropriate way.

A video showing the GM for the first scenario is available as supplemental material.

3.3. LBT

Besides a more traditional pedagogical approach (that, as said, in the present paper is referred to as TL) in which the learners are asked to "linearly" experience the GM and the EM of the VR application, there are other ways to integrate a VRTS like the devised one in a learning path. One of the possible alternatives is represented by LBT. With the TL approach, the learners would typically experience first the GM to learn, guided step-by-step by the voice-over, how to perform the measuring procedure; afterwards, they would be asked to go through the EM and replicate all the steps of the procedure testing their knowledge and ability with it. With the LBT approach, in turn, the learners act as teachers, lecturing other peers. According to the guidelines provided in the literature (Fiorella and Mayer (2014); Okita and Schwartz (2013)), the LBT approach encompasses the steps described below.

- The learner studies the teaching material, takes notes (if necessary) and prepares for a lecture ("expectation to teach", Fiorella and Mayer (2014)).
- The learner gives the lecture teaching one or more peers; the peers may ask questions to get clarifications on the lecture contents and solve doubts. This process is expected to enhance the learner's understanding of the teaching material.
- The peers spend the acquired knowledge performing some task while observed by the learner, who can verify the outcomes of the teaching phase ("recursive feedback", Okita and Schwartz (2013)).

In this work, the LBT approach is implemented as follows (breaking the linearity of TL).

- The learner experiences the GM to learn the measuring procedure and prepares for a lecture; if needed, he or she can pause the application and take notes on paper.
- The learner leads another subject (with no previous knowledge of the procedure) who experiences the EM; in doing that, the learner can assess his or her own knowledge of the procedure acquired in the GM (hence the name, LBT).
- The other subject experiences again the EM while observed by the learner, who can verify the outcomes of his or her teaching without intervening in the operations or providing any help.

4. Experiments

This section presents the two user studies that were performed using the developed VRTS, first to investigate the efficacy of the devised tool as a substitute or in combination with a laboratory activity, then to analyze its use in the context of two possible pedagogical approaches (TL and LBT). All the questionnaires used in the two studies are available as supplemental material. The population of the studies encompassed students enrolled in engineering degrees at the authors' university, with diverse backgrounds and aged between 21 and 27 ($M = 23.63$, $SD = 1.60$). The students had no previous knowledge of the measuring procedure.

In order to test the representativeness of the study population, a preliminary experiment was performed by involving both the above students and industry operators who participated in the E2DRIVER project. Subjects in the two groups were invited to go through the GM and the EM of the VRTS, and the reports on their performance produced by the application were analyzed. The scores of the two groups were shown to be equivalent (using a TOST test and setting the equivalence margin to $(-1, +1)$ on a 1-to-10 scale), thus indicating that the devised VRTS could be successfully used to train subjects with different levels of expertise. The above reports were not leveraged in the two main studies described in the following, since they could not be produced for all the experimental conditions.

4.1. First Experiment Design

The first experiment was designed to compare a traditional education activity performed in a laboratory, a VR-only training, and a third approach that combined the traditional activity and the use of the VRTS. To avoid possible bias, the laboratory activity was the same used to design the GM of the VR application. For this experiment, the user sample included 36 students (24 males, 12 females). A between-subjects design was adopted, by randomly assigning the participants to three equal-sized groups.

At the beginning of the experiment, each participant was asked to fill in a pre-training questionnaire. Afterwards, each group experienced the training:

- the first group (VR) experienced only the devised VRTS (using the GM to learn the procedure and the EM for assessment);
- the second group (L+VR) experienced, in addition to the laboratory activity (focused on the same topics taught in the GM), the devised VRTS (using the GM to learn the procedure and the EM for assessment);
- the third group (L) experienced only the laboratory activity on the same topics taught in the GM, and received no additional training.

After the experience, the participants were asked to take a post-training quiz on the procedure, designed to compare the knowledge gain associated with the three training experiences; they were also invited to fill in a post-training questionnaire evaluating the experienced training and collecting general feedback as well as possible suggestions for future improvements.

The post-training quiz consisted in a series of questions and exercises concerning safety measures and procedural details. The items of the quizzes were grouped in two parts, organized as follows:

- the first part (or *ordering part*) consisted of an exercise asking to put the various steps of the procedure ("connecting the electrical probes to the measuring instrument", "connecting the electrical probes to the switchboard", etc.) in the correct order;
- the second part (or *connections part*) consisted of a more detailed exercise that asked to connect the different parts of the electrical switchboard (i.e., each contact of the power line) to the corresponding sockets on the measuring tool.

The quizzes and exercises were created ad hoc with the help of the professors who supported the design of the VRTS.

With respect to the questionnaires used in this experiment, the pre-training one was designed to collect demographic information, and included questions on background knowledge (e.g., about prior experience with VR, with measuring procedures, etc.), together with a series of questions aimed to investigate the participants' perceived self-efficacy (i.e., expectations from or attitude towards the training to be experienced, Heyne et al. (1998)). The post-training questionnaire was composed of several sections designed to investigate different aspects of the pedagogical approach and of the developed VRTS, which are described below.

- The first section investigated the participants' self-efficacy after the learning experience.

- The second section consisted of the Instructional Materials Motivation Survey (Keller (2010)), or IMMS. It included 36 statements investigating the participants' motivations for learning; the participants were asked to score each statement on a 1-to-5 scale (with one corresponding to "not true" and five to "very true"). The statements can be grouped in four categories (attention, confidence, relevance, and satisfaction), and individual scores can be used to calculate an overall score for each category.

- The third section consisted of a subset of the AttrakDiff questionnaire (Hassenzahl et al. (2008)) evaluating the attractiveness and hedonic quality stimulation of the pedagogical experience (Jost et al. (2020)). In particular, the participants were asked to assign a score to 14 pairs of terms on a 1-to-7 scale. Low values corresponded to a positive evaluation.

- The fourth section corresponded to the System Usability Scale (SUS, Brooke (1996)) . The participants were asked to score 10 statements on a 1-to-5 scale (with one corresponding to "total disagreement" and five to "total agreement"), evaluating the general usability of the developed VRTS. It was not used with the L group.

- The fifth section encompassed general usability questions extracted from the VRUSE questionnaire (Kalawsky (1999)). Selected questions investigated the fidelity of the simulated procedure, the level of presence and immersion, the participants' satisfaction, and the overall usability of the developed VRTS. It was not used with the L group.

4.2. First Experiment Results

All the data were collected and analyzed using MS Excel and the Real-Statistics add-on. Cronbach's alpha (α_C) was calculated for each section of the questionnaires. Comparisons were performed using the Kruskal-Wallis Test, possibly followed by the Pairwise Mann-Whitney Test to detect significant differences between the three groups (with a significant threshold of $p \leq 0.05$) where not stated otherwise.

Regarding the pre-training questionnaire, no significant differences were found. The participants of the three groups rarely used VR technologies and were not familiar with the measuring procedure (on a 1-to-5 scale, with one corresponding to "Not familiar" and five to "Very familiar": $M = 1.36$ and $SD = 0.48$ for the L group, $M = 1.09$ and $SD = 0.28$ for the VR group, $M = 1.38$ and $SD = 0.62$ for the L+VR group, $p = 0.310$).

For what it concerns the post-training questionnaire, the results of the IMMS indicate a significant appreciation for the L+VR experience in terms of overall motivation for learning with respect to the L one and the VR one, with no significant differences between the L and the VR groups. In particular, the L+VR was able to hold the attention ($\alpha_C = 0.78$) more than the other two experiences, was judged as more satisfying ($\alpha_C = 0.84$) and more relevant ($\alpha_C = 0.75$). The participants of the L+VR group were also more confident ($\alpha_C = 0.76$) compared to those of the L group, whereas no significant differences were found with respect to the VR group. Considering the VR and L groups, no differences were observed over the four categories. The results for the four IMMS categories and the total score are given in Table 1.

These results can be explained by the fact that the L+VR experience is the most complete in terms of contents and pedagogical offer, combining an explanation of the procedure by a professor with a practical exercise. By experiencing the traditional activity together with the VRTS, the participants of the L+VR group were able to notice the limitations of the former (the lack of a practical session) and the advantages brought by the latter (the possibility to work on a simulated electrical switchboard with no risks), and this put them in the condition to appreciate the relevance of this mixed approach, resulting in a more satisfying training; moreover, by using the VRTS

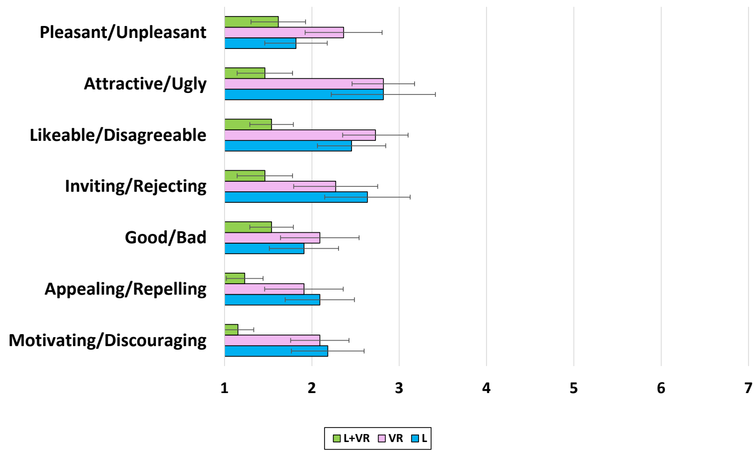
Category	L+VR	VR	L	K-W	L+VR vs L	L+VR vs VR	VR vs L
Attention	4.47 (0.29)	3.85 (0.50)	4.08 (0.42)	**0.006**	**0.030**	**0.001**	0.478
Satisfaction	4.13 (0.56)	3.56 (0.60)	3.48 (0.64)	**0.042**	**0.026**	**0.041**	0.949
Relevance	3.97 (0.47)	3.41 (0.67)	3.60 (0.32)	**0.027**	**0.040**	**0.021**	0.217
Confidence	3.89 (0.41)	3.72 (0.46)	3.39 (0.53)	**0.049**	**0.015**	0.392	0.151
Total	4.14 (0.28)	3.66 (0.34)	3.72 (0.44)	**0.013**	**0.035**	**0.003**	0.846

Table 1: Subjective results based on the IMMS: average scores and standard deviations for the four categories and total score. Bold font is used to highlight significant differences (p-value < 0.05). When the three-way comparison performed with the Kruskal–Wallis Test (K-W column) is significant, the significant pairwise p-values of the follow-up test are provided.

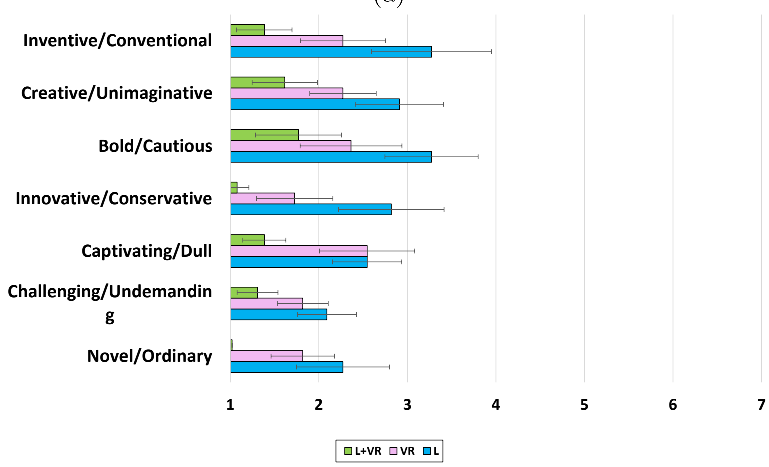
the participants were able to focus on the elements of the procedure that they deemed more critical during the first activity, keeping a high attention throughout the whole VR experience.

A similar situation can be observed when considering the results of the AttrakDiff, investigating the Attractiveness ($\alpha_C = 0.90$) and the Hedonic Quality Stimulation ($\alpha_C = 0.82$) dimensions of the experience (Fig. 4).

Looking at the average scores, the L+VR group performed better than the other two groups on every item of the questionnaire, whereas the L and VR groups followed similar trends and assigned similar scores. The L+VR experience scored significantly better than the L one over 12 of the 14 items, except for "Good/Bad" and "Pleasant/Unpleasant" for which no significant differences were found. Moreover, by comparing the L+VR and the VR experiences, the former was judged as significantly more motivating, likable and overall attractive with respect to the latter in terms of the Attractiveness dimension, whereas it was considered as significantly more captivating, more inventive, and better in terms of novelty for what it concerns the Hedonic Quality Stimulation dimension. Finally, by comparing the L and VR groups, only one significant difference was found for the "Innovative/Conservative" item, where the VR experience was considered better than the L one. These results can be ascribed again to the completeness of L+VR approach, where the students were able to experience the traditional activity, notice its lim-



(a)



(b)

Figure 4: First experiment, results of the AttrakDiff in terms of (a) Attractiveness, and (b) Hedonic Quality Stimulation.

itations, and then use the VRTS to overcome these limitations, therefore resulting in a more appealing and stimulating experience. In contrast, the L approach was characterized by logistic limitations and the lack of a practical session performed on the electrical switchboard, which undermined the overall quality of the experience, whereas the VR consisted of a simulated activity that was unfamiliar to the participants, which may explain the lower scores.

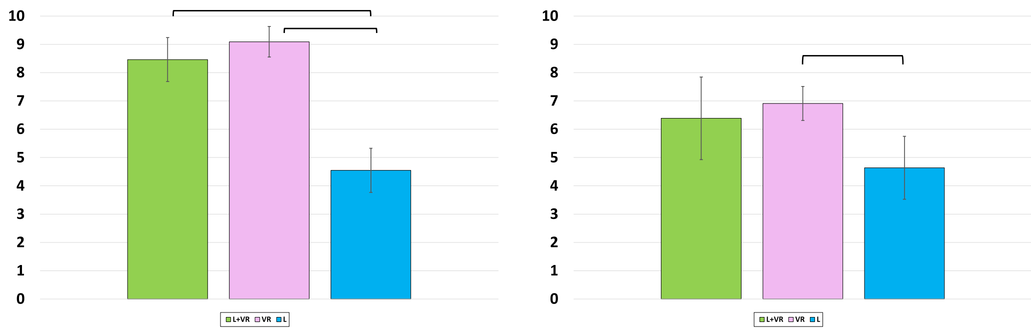
To conclude the subjective analysis, the VRTS was evaluated using the

SUS and VRUSE questionnaires. Based on the overall SUS score, the VR application was considered as "good" by all the participants ($M = 85.03$, $SD = 8.08$). As for the VRUSE, some differences were found between the two groups in terms of simulation fidelity ($\alpha_C = 0.76$). In particular, the participants of the L+VR group were significantly more impressed by the interactions offered by the simulation ($M = 3.54$ and $SD = 0.78$ for the VR group, $M = 4.31$ and $SD = 0.91$ for the L+VR group, $p = 0.044$), probably due to the fact that the VR group did not experience the traditional laboratory activity and could not make a direct connection between the functionalities offered by the simulation and the actions performed in the real procedure.

For what it concerns the objective evaluation, the answers to quizzes were rated on a 1-to-10 scale (Fig. 5). The results for the ordering part (Fig. 5a) showed that the VR and L+VR groups performed significantly better than the L one ($p < 0.001$ for the comparison between L and VR, and the one between L and L+VR), confirming the limitations of a laboratory experience where the students can only observe the professor. No significant differences were found between the VR and L+VR groups, showing that both the approaches are suited for communicating the overall structure of the considered procedure. As for the connections part (Fig. 5b), the VR group performed significantly better than the L one ($M = 6.91$ and $SD = 1.90$ for the VR group, $M = 4.64$ and $SD = 2.23$ for the L group, $p = 0.023$), confirming the advantages associated with the use of a VRTS in procedural training; the participants of the VR group, in fact, exercised directly on a simulated version of the electrical switchboard and were able to learn also the operational details of the measuring procedure.

Considering the L+VR group, however, no significant differences were found with respect to the other two groups, meaning that the efficacy of the VR technologies was somehow undermined by the traditional laboratory experience (even though the mixed approach was judged as more complete and was appreciated more than the other two in the subjective evaluation). The traditional laboratory activity, which was not effective for the purpose of learning the procedure, was detrimental since introduced a bias in the participants and led them to be less accurate while using the VRTS. In fact, the connections part of the quiz was designed to evaluate the operational details of the procedure, which were also the most difficult aspects to see and follow while observing the professor.

Finally, regarding the time spent on the pedagogical contents, the par-



(a) (b)

Figure 5: First experiment, quiz results: (a) ordering part of the quiz, and (b) connections part. Brackets indicate significant comparisons (p-value≤ 0.05) between two groups.

ticipants of the L group were exposed to the procedure for 90 minutes (corresponding to the length of the laboratory activity). The participants of the VR group spent, on average, 25 minutes in the VR application (15 minutes in the GM, 10 minutes in the EM). The participants of the L+VR group experienced both these activities, for an average time of 115 minutes. Therefore, the VRTS alone is also the most efficient solution in terms of time economy, since it requires no preparation and can be easily experienced multiple times.

4.3. Second Experiment Design

The second experiment was designed to investigate the possible ways in which VR can be leveraged in a learning path; in particular, a comparison was conducted between a TL scenario (where the VRTS was used to train and evaluate the participants) and a LBT scenario (where the VRTS was used to train the participants, who were then asked to guide other subjects). Based on the results of the first experiment, the traditional laboratory activity was not considered in this second experiment. A new sample was selected involving 36 students (22 males, 14 females). A between-subjects design was adopted, by randomly assigning each participant to two equal-sized groups (TL group and LBT group). For the LBT approach, a third group (LBT Trainees, LBTT) was involved. This third group was composed of 18 students (11 males and 7 females), one for each participant of the LBT group. All the groups were composed by students with no knowledge of the measuring procedure.

Similarly to the first experiment, the participants were first asked to fill in a pre-training questionnaire to collect demographic information; in addition, they were asked also to take a pre-training quiz on the measuring procedure aimed to assess previous knowledge of the procedure. Afterwards, each participant of the TL group was invited to experience the GM and the EM of the VRTS, whereas each participant of the LBT group was asked to experience the GM, train one subject from the LBTT group using the EM (i.e., the subject experiences the VRTS while the study participant guides him or her from outside), and then observe the same subject operating in the EM without intervening (as described in Section 3). The organization of the two experiments is depicted in Fig. 6.

After the experience, the participants of the TL and LBT groups were asked to take a post-training quiz on the procedure with the aim to compare the knowledge gain associated with the two pedagogical approaches; they were also invited to fill in a post-training questionnaire evaluating the two training experiences and collecting general feedback as well as possible suggestions for future improvements. Finally, the participants were asked to answer the same post-training quiz after one-week (retention quiz), to test the knowledge retention associated with the two pedagogical approaches; during this period, the participants were not exposed to further information on the topic.

Regarding the quizzes, the pre-training one was the same used after the training for the first experiment, whereas the post-training one consisted in the same questions and exercises, together with an additional third part on the use of the probes and the structure of the measuring procedure (e.g., the participants were asked to select the actions required to complete the experience from a set of plausible choices). The additional questions and exercises were also created ad hoc with the help of a professor of electrical systems and safety from the authors' university. With respect to the subjective assessment, the pre-training and post-training questionnaires used for the first experiment (IMMS, AttrakDiff, SUS, VRUSE and custom questions) were exploited also for this second experiment, with the addition of some custom questions regarding, e.g., previous teaching experiences or the possible interest in functionalities for taking notes directly in the VR application that could be implemented in the future.

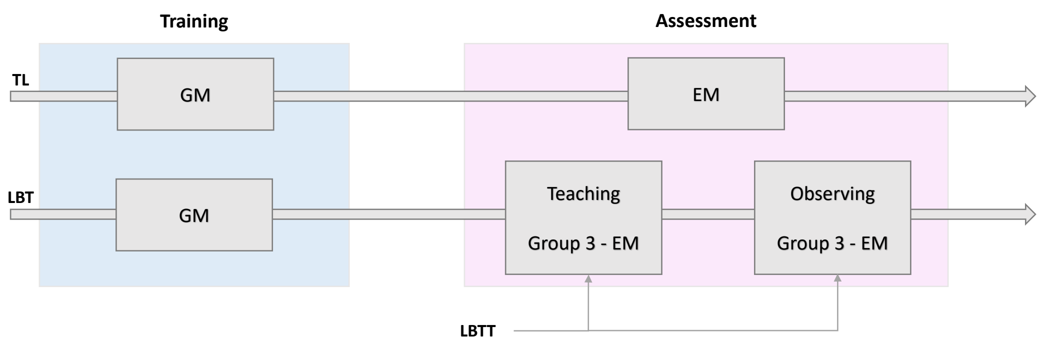


Figure 6: Organization of the TL and LBT approaches.

4.4. Second Experiment Results

Similarly to the first experiment, all the data were collected and analyzed using MS Excel and the Real-Statistics add-on. Cronbach's alpha was calculated for each section of the questionnaires. Comparisons were performed using the two-tailed Mann-Whitney U-test (with a significant threshold of $p \leq 0.05$) where not stated otherwise.

Regarding the pre-training questionnaire, no significant differences were found. The participants of the two groups rarely used VR technologies and were not familiar with the considered measuring procedure (on a 1-to-5 scale, with one corresponding to "Not familiar" and five to "Very familiar": $M = 1.26$ and $SD = 0.44$ for the LBT group, $M = 1.27$ and $SD = 0.55$ for the TL group, $p = 0.89$). Moreover, both the groups were moderately used to teach other people (on a 1-to-5 scale, with one corresponding to "I never teach" and five to "I teach on a daily basis": $M = 2.13$ and $SD = 1.02$ for the LBT group, $M = 2.55$ and $SD = 1.16$ for the TL group, $p = 0.32$); in particular, six participants (one from the LBT group) reported to be teaching others once a week, whereas none reported to be teaching on a daily basis.

With respect to the post-training questionnaire, no significant differences were observed between the two groups for the IMMS (Attention $\alpha_C = 0.82$, Satisfaction $\alpha_C = 0.80$, Relevancy $\alpha_C = 0.84$, Confidence $\alpha_C = 0.77$) and the AttrakDiff (Attractiveness $\alpha_C = 0.92$ and Hedonic Quality Simulation $\alpha_C = 0.86$). These results support the fact that the authors designed the two pedagogical approaches to be similar in terms of motivation for learning, attractiveness and hedonic quality stimulation. The results for the four categories and the total score of the IMMS are reported in Table 2, whereas the results of the AttrakDiff are given in Fig. 7.

Category	TL	LBT	*p*-value
Attention	4.35 (0.43)	4.19 (0.51)	0.330
Satisfaction	4.08 (0.38)	3.83 (0.65)	0.163
Relevance	3.81 (0.61)	3.61 (0.69)	0.329
Confidence	3.72 (0.34)	3.79 (0.46)	0.576
Total	4.01 (0.33)	3.89 (0.41)	0.314

Table 2: Subjective results of the second experiment based on the IMMS: average scores and standard deviations for the four categories and total score. Bold font is used to highlight significant differences (p-value < 0.05).

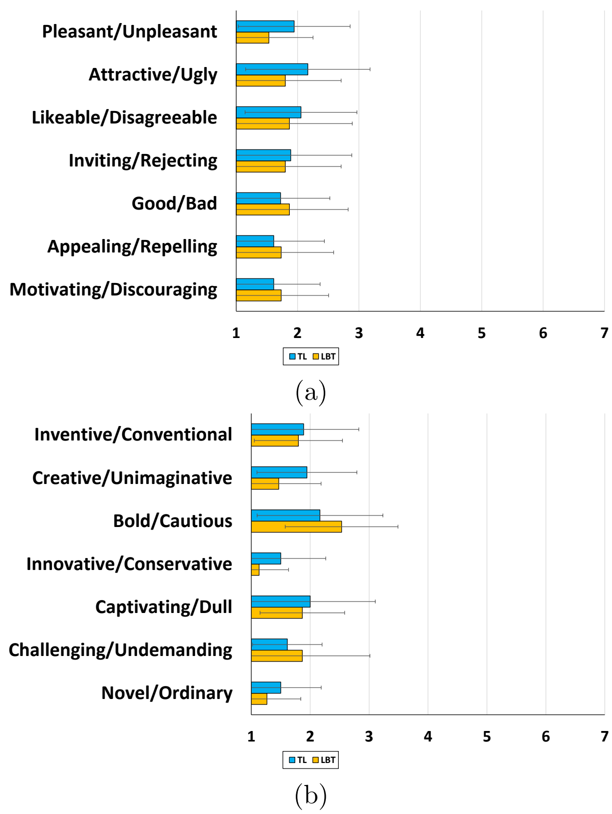


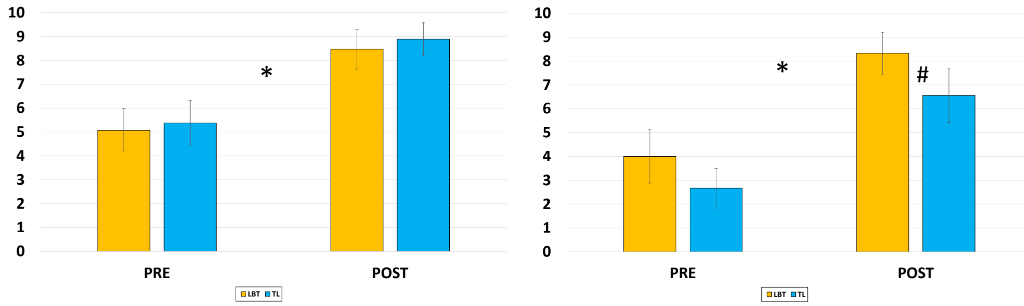
Figure 7: Second experiment, results of the AttrakDiff in terms of (a) Attractiveness, and (b) and Hedonic Quality Stimulation.

Moving to the SUS and VRUSE results, no significant differences were found; this outcome was expected, since the two groups experienced the same VRTS. Overall, based on the SUS scores, the usability of the application was judged as "excellent" by both the groups ($M = 83.67$ and $SD = 8.90$). About the VRUSE, looking at the individual pairs of terms it can be observed that the participants of both the groups praised the adequateness of the functionality offered by the system ($M = 4.53$ and $SD = 0.49$ for the LBT group, $M = 4.50$ and $SD = 0.50$ for the TL group), the potential benefit associated with the use of this VRTS ($M = 4.73$ and $SD = 0.44$ for the LBT group, $M = 4.67$ and $SD = 0.60$ for the TL group), and the overall simulation fidelity ($M = 4.73$ and $SD = 0.44$ for the LBT group, $M = 4.61$ and $SD = 0.50$ for the TL group).

For what it concerns the objective evaluation, the quizzes were rated on a 1-to-10 scale. The results of the pre-training quiz confirmed that the participants were unfamiliar with the considered measuring procedure. In both the ordering part ($M = 5.07$ and $SD = 1.80$ for the LBT group, $M = 5.39$ and $SD = 1.86$ for the TL group, with no significant differences) and the connections part ($M = 4.0$ and $SD = 2.23$ for the LBT group, $M = 2.67$ and $SD = 1.67$ for the TL group, with no significant differences), the average scores obtained by the two groups were not sufficient.

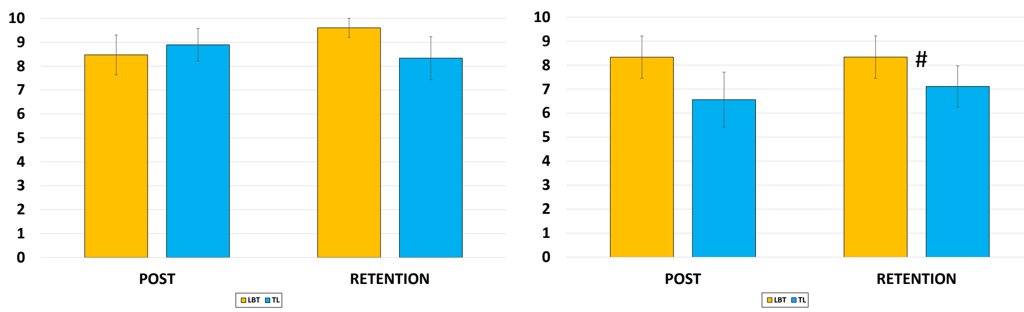
Comparing the pre-training and post-training scores of the ordering part using the Scheirer-Ray-Hare Test, a significant gain ($p < 0.001$) was found for both the groups, confirming the effectiveness of the two approaches for learning the overall structure of the considered procedure, with no significant differences between the two groups (Fig. 8a). Considering the connections part (Fig. 8b), however, although both the groups experienced a significant increase in terms of knowledge gain ($p < 0.001$), using the Scheirer-Ray-Hare Test it was possible to highlight that the gain associated with the use of the LBT approach was significantly better than the TL one ($p = 0.019$) for this part, indicating that the TL approach is less effective for learning the technical details of the considered procedure.

Finally, regarding the retention quiz, in the ordering part the LBT group managed to improve the average score (from $M = 8.47$ and $SD = 1.66$ of the post-training quiz to $M = 9.60$ and $SD = 0.99$), whereas for the TL group the score got worse (from $M = 8.89$ and $SD = 1.37$ of the post-training quiz to $M = 8.33$ and $SD = 1.80$), but no significant differences were found between the two groups (Fig. 9a). Regarding the connections part of the quiz, the average score of the TL group increased (from $M = 6.56$ and $SD = 2.29$



(a) (b)

Figure 8: Second experiment, pre-Post quiz results: learning gain for the (a) ordering part, and (b) connections part. The asterisk marks a significant comparison ($p \leq 0.05$) due to the use of VR, whereas the hash symbol indicates a significant comparison due to the use of different pedagogical approaches.



(a) (b)

Figure 9: Second experiment, results regarding retention: (a) learning gain for the (a) ordering part, and (b) connections part. The hash symbol marks a significant comparison ($p \leq 0.05$) due to the use of different pedagogical approaches.

of the post-training quiz to $M = 7.11$ and $SD = 1.71$), whereas it remained the same for the LBT group ($M = 8.33$ and $SD = 1.76$ for the post-training and retention quizzes), and a significant difference was found in favor of the LBT approach (Fig. 9b). Using the Scheirer-Ray-Hare Test it was possible to confirm that this difference was mainly due to the diverse pedagogical approach ($p = 0.023$), whereas no significant interaction effect or influence of the exposure time were found. This result showed the superiority of the VR-supported LBT approach compared to the TL one, surrogating what it has been stated in the literature for non-VR LBT experiences. No differences were found on the additional parts of the post-training and retention quizzes.

About the exposure to pedagogical contents, the participants of the TL group experienced the VRTS for an average time of 25 minutes (15 minutes in the GM, 10 minutes in the EM), whereas the LBT group was exposed for a total of 40 minutes (15 minutes in the GM, 15 minutes guiding the peers in the EM and answering questions, and 10 minutes observing them experiencing the EM). Even though the exposure of the LBT group was higher than that of the TL group, it is important to note that the additional time was mostly spent observing the peers. According to Okita and Schwartz (2013), this process is comparable to an assessment session in which the learners passively receive feedback on the previously performed teaching. In other words, the participants of the LBT group were not exposed to fresh contents about the considered procedure; furthermore, during this phase they could have been potentially exposed also to incorrect information if their own teaching presented errors or was not clear enough for the peers.

5. Conclusions and Future Work

VR technologies are widely used for the creation of training systems, mainly for experiences focused on procedural operations and hands-on tasks that are common in industrial scenarios. However, further investigations are needed to study the various ways in which VRTSs can be integrated in a learning path to complement, or possibly even replace, traditionally adopted education approaches.

The present paper contributes to advancing the research in this field by focusing on two RQs, which concern how a VRTS could be used in a learning path to improve over non-VR approaches, and how an effective, VR-based TL approach would compare with a VR-based LBT approach. To cope with these RQs, the paper focuses on a case study in the context of energy management and leverages a VRTS developed to support training in a measuring procedure on electrical switchboards.

Two different user studies were conducted, one per RQ. The first study aimed to introduce the VRTS in an existing learning module taught at the authors' university and to investigate the best way to do that by evaluating and comparing the performance of three groups of participants: a first group that experienced only the VRTS, a second group that experienced only the traditional education activity, and a third group that experienced a mixed approach combining both the traditional activity and the use of the VRTS. The results of this first study demonstrated that, although the mixed approach was better overall in terms of motivation for learning, attractiveness and hedonic quality stimulation, the participants that experienced only the VRTS outperformed the other two groups when tested on the measuring procedure. This finding suggests not only that traditional teaching activities may not be fully effective to support procedural training (e.g., due to safety reasons, like in the considered case), but also that these activities, when combined with a VRTS, may introduce a bias which could even hinder the effectiveness of the overall experience. Thus, in scenarios like the considered one, a VRTS could indeed be introduced in an existing learning path, but it might be most effective when used to replace the original experience.

In the second study, the use of VR in combination with a LBT pedagogical approach was compared with a TL approach leveraging VR in a more conventional way. The two training modes offered by the VRTS (GM and EM) were leveraged to design two different learning experiences: in the TL experience, the participants were exposed to the GM mode of the VRTS and, afterwards, to the EM mode (i.e., in the configuration that was found to be the optimal one in the first study); in the LBT experience, after being exposed to the GM, the participants were asked to lead other subjects in the EM. The experiment showed that the integration of VR technologies with the LBT approach was not detrimental. Moreover, the VR-based LBT approach was found to be superior to the VR-based TL one in terms of learning gain and retention for what it concerns the technical details of the procedure, in accordance with the non-VR examples found in the literature. Therefore, the LBT approach could be considered as a viable alternative in the design of future training experiences in VR.

The findings of this work can be summarized as reported below.

- When introducing the VRTS in a learning path, the subjective results showed that the mixed approach combining the use of the application and the laboratory activity was preferred to the experiences taken in isolation, in which only one of the two teaching approaches was leveraged. In particular, the L+VR experience was significantly better than the other two in terms of motivation for learning, likeability, and overall attractiveness.

- In terms of learners' performance, the considered laboratory activity may not be effective to support procedural training and, when combined with the VRTS, it could hinder the effectiveness of the VRTS.

Overall, the use of the VRTS alone was significantly better than the laboratory activity alone, whereas it was not possible to found a clear advantage associated with the use of the the VRTS combined with the laboratory activity.

- When comparing the use of the VRTS alone declined along the two pedagogical approaches, i.e., TL and LBT, no significant differences were found in subjective terms.
- The TL and LBT approaches were both effective in terms of knowledge gain, but the LBT approach proved to be superior with respect to the TL one in supporting the learning of the procedural details of the activity.
- The LBT approach was significantly better then the TL one in terms of knowledge retention after one week concerning the procedural details of the activity.

One of the possible limitations of this work is that, as widely stated in the literature, VR-based training solutions are very scenario-dependent. Hence, further experiments could be needed to confirm the applicability of the findings of this work to different use cases. Another potential limitation could be the fact that the VRTS used in the experiments was not designed to be integrated in a LBT pedagogical approach; for instance, it did not offer a note-taking feature that could have been used by the learners to prepare the lesson while in the GM. A possible future development could consist in adding this feature to the application and investigate the benefits for both the LBT and TL approaches. Moreover, it could be helpful to add a multiplayer component to the application in order to let the teacher and the taught subject in the LBT share the same virtual environment with the aim to investigate the impact of their interactions on the effectiveness of the training experience. In addition, it could be interesting to introduce Non-Player Characters (NPCs) into the VRTS, either as teachers or taught peers for the LBT approach. These characters could be scripted or controlled by AI, and their impact on learning experimentally investigated. In particular, NPC acting as teachers could be compared with real trainers (operating outside the application, like those considered in the second study, or inside the application, leveraging the above mentioned multiplayer functionalities), whereas NPC acting as peers could be considered as teachable agents and could be leveraged to investigate whether the findings of works like, e.g., Pratticò et al. (2021) can be directly transferred from a MR to a VR scenario.

Acknowledgments

This work has been developed in the context of the E2DRIVER project, which received funding from the European Union Horizon 2020 programme under grant agreement No 847038, and the VR@POLITO initiative.

Author's Contributions

Federico De Lorenzis: Conceptualization, Methodology, Software, Investigation, Data Curation, Writing - Original Draft. Filippo Gabriele Pratticò: Conceptualization, Methodology, Investigation, Data Curation, Writing - Original Draft. Maurizio Repetto: Conceptualization, Methodology, Investigation. Enrico Pons: Conceptualization, Methodology, Investigation. Fabrizio Lamberti: Conceptualization, Methodology, Investigation, Writing - Review Editing, Supervision.

Competing Interests

The authors declare that they have no competing interests.

References

Andaluz, V.H., Castillo-Carrión, D., Miranda, R.J., Alulema, J.C., 2017. Virtual reality applied to industrial processes, in: De Paolis, L.T., Bourdot, P., Mongelli, A. (Eds.), Augmented Reality, Virtual Reality, and Computer Graphics, Springer International Publishing, Cham. pp. 59–74.

Biswas, G., Leelawong, K., Schwartz, D., Vye, N., at Vanderbilt, T.T.A.G., 2005. Learning by teaching: A new agent paradigm for educational software. Applied Artificial Intelligence 19, 363–392.

Brooke, J., 1996. SUS: A 'quick and dirty' usability scale. Usability Evaluation in Industry , 189doi:10.1201/9781498710411-35.

Buttussi, F., Chittaro, L., 2021. A comparison of procedural safety training in three conditions: Virtual reality headset, smartphone, and printed materials. IEEE Transactions on Learning Technologies 14, 1–15. doi:10.1109/TLT.2020.3033766.

Cakiroğlu, Ü., Gökoğlu, S., 2019. Development of fire safety behavioral skills via virtual reality. Computers & Education 133, 56–68. doi:10.1016/j.compedu.2019.01.014.

Cannavò, A., Calandra, D., Pratticò, F.G., Gatteschi, V., Lamberti, F., 2021. An evaluation testbed for locomotion in virtual reality. IEEE Transactions on Visualization and Computer Graphics 27, 1871–1889. doi:10.1109/TVCG.2020.3032440.

Checa, D., Bustillo, A., 2020. A review of immersive virtual reality serious games to enhance learning and training. Multimedia Tools and Applications 79, 5501–5527. doi:10.1007/s11042-019-08348-9.

Conges, A., Evain, A., Benaben, F., Chabiron, O., Rebiere, S., 2020. Crisis management exercises in virtual reality, in: 2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), pp. 87–92. doi:10.1109/VRW50115.2020.00022.

De Lorenzis, F., Pratticò, F.G., Lamberti, F., 2022. Hcp–vr: Training first responders through a virtual reality application for hydrogeological risk management, in: HUCAPP 2022, pp. 273–280. doi:10.5220/0011007800003124.

Duran, D., 2017. Learning-by-teaching. evidence and implications as a pedagogical mechanism. Innovations in Education and Teaching International 54, 476–484.

Fiorella, L., Mayer, R.E., 2014. Role of expectations and explanations in learning by teaching. Contemporary Educational Psychology 39, 75–85.

Gavish, N., Gutiérrez, T., Webel, S., Rodríguez, J., Peveri, M., Bockholt, U., Tecchia, F., 2015. Evaluating virtual reality and augmented reality training for industrial maintenance and assembly tasks. Interactive Learning Environments 23, 778–798. doi:10.1080/10494820.2013.815221.

Hassenzahl, M., Koller, F., Burmester, M., 2008. Der user experience (UX) auf der spur: Zum einsatz von www.attrakdiff.de, in: Brau, H., Diefenbach, S., Hassenzahl, M., Koller, F., Peissner, M., Röse, K. (Eds.), Tagungsband UP08, Fraunhofer Verlag, Stuttgart. pp. 78–82.

Heyne, D., King, N., Tonge, B., Rollings, S., Pritchard, M., Young, D., Myerson, N., 1998. The self-efficacy questionnaire for school situations: Development and psychometric evaluation. Behaviour Change 15, 31–40.

Jensen, L., Konradsen, F., 2018. A review of the use of virtual reality head-mounted displays in education and training. Education and Information Technologies 23, 1515–1529. doi:10.1007/s10639-017-9676-0.

Jost, P., Cobb, S., Hämmerle, I., 2020. Reality-based interaction affecting mental workload in virtual reality mental arithmetic training. Behaviour & Information Technology 39, 1062–1078. doi:10.1080/0144929X.2019.1641228.

Kalawsky, R.S., 1999. VRUSE – A computerised diagnostic tool: For usability evaluation of virtual/synthetic environment systems. Applied Ergonomics 30, 11–25. doi:10.1016/S0003-6870(98)00047-7.

Kaluschke, M., Weller, R., Zachmann, G., Pelliccia, L., Lorenz, M., Klimant, P., Knopp, S., Atze, J.P.G., Móckel, F., 2018. Hips - a virtual reality hip prosthesis implantation simulator, in: 2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), pp. 591–592. doi:10.1109/VR.2018.8446370.

Kaplan, A.D., Cruit, J., Endsley, M., Beers, S.M., Sawyer, B.D., Hancock, P., 2020. The effects of virtual reality, augmented reality, and mixed reality as training enhancement methods: A meta-analysis. Human Factors , (in press).

Keller, J., 2010. Motivational Design for Learning and Performance: The ARCS Model Approach. doi:10.1007/978-1-4419-1250-3.

Lacko, J., 2020. Health safety training for industry in virtual reality, in: 2020 Cybernetics & Informatics (K&I), pp. 1–5. doi:10.1109/KI48306.2020.9039854.

Lamberti, F., De Lorenzis, F., Pratticò, F.G., Migliorini, M., 2021. An immersive virtual reality platform for training CBRN operators, in: Proc. of 2021 IEEE 45th Annual Computers, Software, and Applications Conference (COMPSAC), pp. 133–137. doi:10.1109/COMPSAC51774.2021.00030.

Lovreglio, R., Duan, X., Rahouti, A., Phipps, R., Nilsson, D., 2021. Comparing the effectiveness of fire extinguisher virtual reality and video training. Virtual Reality 25, 133–145. doi:10.1007/s10055-020-00447-5.

Mirauda, D., Capece, N., Erra, U., 2020. Sustainable water management: Virtual reality training for open-channel flow monitoring. Sustainability 12. URL: `https://www.mdpi.com/2071-1050/12/3/757`, doi:10.3390/su12030757.

Morillo Tena, P., García, I., Orduña, J., Fernández, M., Juan, M.C., 2020. Comparative study of ar versus video tutorials for minor maintenance operations. Multimedia Tools and Applications 79. doi:10.1007/s11042-019-08437-9.

Naranjo, J.E., Sanchez, D.G., Robalino-Lopez, A., Robalino-Lopez, P., Alarcon-Ortiz, A., Garcia, M.V., 2020. A scoping review on virtual reality-based industrial training. Applied Sciences 10. URL: `https://www.mdpi.com/2076-3417/10/22/8224`, doi:10.3390/app10228224.

Okita, S.Y., Schwartz, D.L., 2013. Learning by teaching human pupils and teachable agents: The importance of recursive feedback. Journal of the Learning Sciences 22, 375–412.

Pellas, N., Dengel, A., Christopoulos, A., 2020. A scoping review of immersive virtual reality in STEM education. IEEE Transactions on Learning Technologies 13, 748–761. doi:10.1109/TLT.2020.3019405.

Pratticò, F.G., Lamberti, F., 2021. Towards the adoption of virtual reality training systems for the self-tuition of industrial robot operators: A case study at KUKA. Computers in Industry 129. Cited by: 5.

Pratticò, F.G., Merino, F., Lamberti, F., 2021. Is learning by teaching an effective approach in mixed-reality robotic training systems? Intelligent Technologies for Interactive Entertainment , 177–190.

Rantanen, E.M., Talleur, D.A., 2005. Incremental transfer and cost effectiveness of groundbased flight trainers in university aviation programs, in: Proc. Human Factors and Ergonomics Society Annual Meeting, pp. 764–768.

Roscoe, R.D., Chi, M.T., 2007. Understanding tutor learning: Knowledge-building and knowledge-telling in peer tutors' explanations and questions. Review of educational research 77, 534–574.