Towards the adoption of virtual reality training systems for the self-tuition of industrial robot operators: A case study at KUKA

Original
Towards the adoption of virtual reality training systems for the self-tuition of industrial robot operators: A case study at KUKA / Pratticò, Filippo Gabriele; Lamberti, Fabrizio. - In: COMPUTERS IN INDUSTRY. - ISSN 0166-3615. - STAMPA. - 129:(2021), p. 103446. [10.1016/j.compind.2021.103446]

Availability:
This version is available at: 11583/2874574 since: 2021-03-27T08:24:41Z

Publisher:
Elsevier

Published
DOI:10.1016/j.compind.2021.103446

Towards the adoption of virtual reality training systems for the self-tuition of industrial robot operators: A case study at KUKA

F. Gabriele Pratticò[a], Fabrizio Lamberti[a,*]

[a]*Politecnico di Torino – Dipartimento di Automatica e Informatica, C.so Duca degli Abruzzi 24, 10129, Torino (TO), Italy*

ARTICLE INFO

Keywords:
Virtual Reality Training System
Self-Tuition
Industrial Robot
Maintenance

ABSTRACT

Interest is growing around Virtual Reality training systems (VRTSs), which started to be considered as a credible option to train companies' workforce. Although the efficacy of VRTSs as a fancy alternative to traditional learning material used by trainers in their lectures has already been proved, their effectiveness as self-learning tools not requiring human instructors is still controversial, since experiments carried out within established training programmes are still rare. This paper reports the results of a user study aimed to investigate how an immersive VRTS designed to train industrial robot operators on light maintenance tasks actually compares to the training practice based on in-class and hands-on sessions that is adopted worldwide by an international company. After the training, study participants were evaluated by company's instructor while performing the taught task autonomously on a real robot. Obtained results showed that the effectiveness of the devised VRTS in making the trainees able to successfully complete the task was comparable to that of the traditional training. It was also found that there is still room for improvement, e.g., on virtually training interaction with physical tools and equipment, as well as on emulating and fortifying aspects of human trainer-trainee social dynamics.

1. Introduction

The advancements in the field of eXtended Reality (XR) over the last decade are unprecedented for these media, and the availability of cost-effective hardware is promoting their diffusion at a mass scale. Due to the disruptive potential of this technology, both the industry and the academia are dedicating significant efforts to help it attain maturity in a variety of fields, encompassing engineering, arts, design, architecture, medicine, etc. [5].

Since from the first developments, for both the main media in the XR family, i.e., Mixed Reality (MR) and Virtual Reality (VR), training represented the application attracting most of the interest. Thus, simulation-based training is nowadays a well-established practice, e.g., in the aviation industry, as it was proved that pilots who first trained in simulators required less in-flight training to reach a satisfactory level of proficiency [35].

However, it is just in recent years that XR training systems (XRTSs), especially those supporting self-learning/-assessment, stepped out from research laboratories and reached the industry [24, 7, 12, 18]. In this context, XRTSs are generally used to complement established training practices rather than to replace them [18, 29]. Notwithstanding, the possibility to use an XRTS for transferring the required abilities to the trainee without the intervention of a human instructor could be an extremely valuable option, should this system be proved to be even just as effective as traditional methods in reaching this goal [18].

Providentially, since XR is acknowledged to be a key element in the Industry 4.0 perspective [24], industry is giving a great momentum to the experimentation of this technology. Among the many use cases for training, applications prevalently tackle assembly and maintenance tasks [11, 9]. In this context, MR training systems (MRTSs) have already demonstrated their capability to be equally or even more effective than conventional approaches in passing the intended knowledge to trainees. Even considering the current challenges and limitations of technology, paper-based instructions are easily outperformed by MR solutions offering the possibility to receive timely step-by-step instructions directly overlaid to the equipment or the tool being used [39], without sacrificing the possibility to interact with physical elements.

The above features, coupled with the relatively simpler process of authoring a MR environment [26] with respect to a VR one (considering, e.g., the smaller number of 3D assets required and the lower complexity of the simulation),

---

[*] Corresponding author

✉ `fabrizio.lamberti@polito.it` (F. Lamberti)
ORCID(s): 0000-0001-7606-8552 (F.G. Pratticò); 0000-0001-7703-1372 (F. Lamberti)

led to the immediate success of these systems. However, the need to rely on physical equipment to deploy the MR experience represents at the same time the major weakness of MRTSs, which limits their flexibility and scalability if compared to VR training systems (VRTSs) [18].

Unfortunately, according to recent literature reviews [18, 12, 9], the situation of VRTSs for industrial settings is far from ideal. The body of research is scattered, and results are often controversial or can hardly be generalized to real use cases (because, e.g., of the hardware used, of over-controlled experimental conditions, etc.). This situation is also partly due to the fact that the success of the medium is arguably task-dependent, and it remains to be investigated what kind of learning outcome acquisition is best supported by VRTSs in a given scenario [40].

Motivated by the scarcity of studies that, moving from strategies adopted by companies' training departments, analyzed how VR training could be integrated into current practice, as well as by the growing need for on-field evaluations [18, 12], this paper presents an immersive VRTS featuring state-of-the-art functionalities which has been designed to support self-learning in a specific domain represented by industrial robot (IR) maintenance. More explicitly, this work is driven by the following research questions:

RQ.1 Can an immersive VRTS guarantee a transfer of knowledge and skills comparable to that of traditional training? Specifically, is an individual trained with such a system able to perform a self-learnt procedure on a real IR with a level of performance comparable to that of an individual trained in a traditional way?

RQ.2 From a trainee perspective, is the training experience with a self-learning immersive VRTS less satisfactory compared to a traditional learning session with a human instructor?

The rest of the paper is structured as follows: Section 2 provides a background for the use of VRTSs in IR training; Section 3 presents the selected case study, describing the standard training delivered by KUKA Roboter Italia Spa as well as the system devised to implement it in immersive VR; Section 4 introduces the experiments that have been performed to compare the above training methods, whereas Section 5 discusses obtained results; finally, Section 6 summarizes key findings and shows some possible directions for future research.

2. Background

As confirmed by a decent amount of studies [18], the most recommended way to let a user experience the simulated environment of a VRTS is by leveraging immersive VR. In particular, among the possible alternatives, Head-Mounted Displays (HMDs) represent the most cost-effective technology to adopt. However, although immersion can be sufficient, alone, to boost VRTS knowledge transfer capabilities compared to, e.g., self-taught content acquisition from printed, static material (like an aircraft's safety card [4]), this is not the case of industrial assembly and maintenance (IMA) tasks. In fact, learning systems adopted today to train operators for the above tasks are more complex and structured than (just) paper-based instructions: for instance, they often employ human instructors for pairwise or face-to-face teaching [40], and allow the trainees to gain "hands-on" experience on the physical equipment [9].

One of the reasons why IMA has been acknowledged to be particularly challenging from a didactic point of view is that it falls in the category of *mixed tasks* [18]. Those are tasks were a combination of both cognitive and physical requirements are to be met in order to successfully accomplish the goal, since a trainee should be able to use learned skills while simultaneously recalling applicable procedural information, by also analyzing possible safety risks [18]. Thus, as a matter of example, VRTSs not allowing the trainee to interact with the elements in the virtual environment (VE) were proven to be less effective than a mere slide-supported classroom lecture [21].

Given the key role played by interaction in VRTSs, plenty of methods were proposed to increase the interaction fidelity, e.g., using gloves or other devices able to track hands and fingers [16], or employing elements capable of conveying to the users haptic and force feedback [34]. Nevertheless, most of these components are still immature, costly, or not reliable enough to be adopted in common VR-based training scenarios. For these reasons, tracked hand controllers are today the most viable and robust choice to deploy at scale [16].

A recent work leveraging consumer-grade hardware mentioned above, that can be taken as an example of VRTSs for IMA evaluated in concrete training settings is reported in [40]. The authors conducted a user study on pump maintenance with the aim to determine how a VRTS would compare with two hands-on teaching methods adopted by the involved company, namely, pairwise tutoring and video-based training. Results showed that, even though the trainees who experimented the VRTS performed decently when asked to replicate the procedure on the physical pump, the effectiveness of VR-based training proved to be significantly lower than that of both the standard methods. Even though special care was taken in designing the VR experience to make content easy to understand (e.g., using 3D animations to explain the correct actions to be executed [3]), due to the numerous tools and the small components to be manipulated, the VRTS failed at conveying the required feedback to the user.

Nevertheless, operating conditions in training scenarios involving IRs may be quite diverse from those considered in [40]. Besides possible differences in the required interactions, in the above work only small-scale movements were considered. The trainees could perform the task on a table or in a seated-only position, which may not be the case when intervention pertains medium-to-large size equipment. Operator performing a maintenance task on an IR may be required to move in the environment; hence, they must constantly be aware of the risks and potentially harmful behavior of a moving machine by exploiting their proprioception and spatial awareness abilities. Unfortunately, although supporting interaction in VR is important to foster proprioception [27], it is not enough to also enforce spatial knowledge transfer, which is affected by other factors like the time spent in the VE [38, 30], the simulation fidelity [18], etc. Works like [28] substantiate the superiority of immersive VR in fostering a spatial knowledge transfer w.r.t. a non-immersive experience (with desktop-based VR). Spatial awareness was specifically investigated by the authors of [25], who used an immersive VRTS to instruct an operator on how to manage the space shared with a cooperative IR in a simulated manufacturing task. According to the reported qualitative results, the VRTS was found to be a good way to firstly approach a robot; however, there was no evaluation of training effectiveness on a real robot.

In [29], the effectiveness of using a VRTS for training robot programming abilities was evaluated through a user study involving also a real robot. Experiments showed that individuals trained with the VRTS achieved better results w.r.t. those trained with standard training. Differently than in the previous work, however, the VRTS was not used for self-tuition, but rather as a support tool to complement the instructional material in a traditional classroom setting with a human instructor. Moreover, the tool was evaluated by considering as control condition individuals who received just a theoretical training. Although this condition was appropriate for the study, this kind of training may not be fully representative of real-world scenarios in which theoretical training is usually complemented by hands-on practice.

Another work in which a VRTS was used to teach robot programming and show trajectories execution is reported in [36]. In this case, the system only allows the user to have an immersive view of a traditional desktop-based programming software, without actually offering the possibility to interact with the VE using VR interfaces (the intention was to create a very low-cost system, and for this reason hand controllers were not considered). Thus, this system could hardly be used to train IR operators on tasks requesting manual operations.

The study reported in [33] also targeted robot programming. In this case, VR interaction was supported; in particular, already mentioned limitations of the hand controllers were acknowledged, and the standard interface to program the robot was simplified and adapted to better exploit the potential of immersive VE. Even though this new interface for programming the robot was judged as promising by the participants of the devised user study, it is unknown how they would have behaved if asked to program a real robot since, like in [25], this latter experience was not foreseen in the study. As a matter of fact, the “encoding specificity” principle suggests that when there is a (not small enough) discrepancy between the learning environment and the environment in which the learning effectiveness is subsequently measured, performance tends to deteriorate [37]. Hence, it can be speculated that a subject trained with such a system would need to undergo at least an accessory training on the real robot programming interface in order to gain proficiency on the given task.

A possible way to tackle the above problem could be to let the user experience the same task under different conditions (environments, configurations, etc.). In this respect, it could be worth recalling that the authors of [22] recently proposed a method to procedurally generate multiple VEs that may be used to train IR operators. Unfortunately, no evaluation was performed to evaluate the impact on training effectiveness of such an approach, as the goal was just to determine the quality of the generated VEs.

To the best of the authors’ knowledge, the few papers cited above are the only published works that used a VRTS in a training scenario encompassing IRs. Besides choices made, e.g., about how to reproduce robot’s interfaces, or hardware used, which may hinder representativeness and applicability/replicability, the major limitation of these works is that no study was performed to validate if the VRTS was actually able to assure an adequate transfer of knowledge and skills in a self-learning scenario, by asking the users to perform, on a physical robot, the particular procedures that were taught them by the system on a virtual replica.

Figure 1: KUKA Smart Pad (SP): this is the device that is used by operators to manually control and program the robot. It incorporates a touch-based graphics interfaces, as well as some physical buttons and switches for safety-critical operations.

3. Case Study

This section reports the activities that were carried out to create an immersive VRTS targeted to self-training on IRs and to validate its effectiveness by means of a user study encompassing a real robot. The use case was selected by mutual agreement with several experienced instructors from the KUKA College[1], who exhibited a genuine interest in exploring possible ways to convert part of their training practice into a VR-based experience. The aim of the study was to evaluate whether such a VRTS could guarantee a transfer of knowledge and skills as well as an experience comparable to that of the standard training programme adopted by the company worldwide.

As a subject for the study, the *mastering procedure* (MP) of an industrial manipulator was selected. This choice was made considering that, even though the procedure is taught in basic robot programming courses, the mastering of a robot is actually a light maintenance task (a mixed task, according to the definition given in Section 2). Moreover, it is an infrequent task, that operators (generally not belonging to the maintenance staff) perform directly on the production line. Hence, a VRTS could let the operators autonomously access training content directly when needed without overloading the technical support (improving retention and reducing delays on the production line), and could support the implementation of asynchronous remote training strategies (limiting the need to move people to training facilities).

3.1. Mastering Procedure

In order to achieve accurate, precise and repeatable movements, an IR must be calibrated. For each robot's axis, the objective is to align the internal references of electrical/software zero-point to the mechanical one [23]. As a good practice, a robot should be calibrated prior to the first use (commissioning), in case of payload or load distribution changes, after/during maintenance, as well as in case of collisions or failures. The calibration of a KUKA robot is achieved through a MP whose key steps are summarized below.

1. All the axes must be manually moved by the operator using the SmartPad (SP) in Fig. 1 to make each of them reach a specific angle named *pre-calibration position* (different for each robot class, and slightly different within the class for each robot). This has to be done by visually aligning specific references which can be found in the proximity of axis joints (visible in Fig. 2); no further feedback is provided to the operator by the robot or the SP. Order is not relevant.

2. An additional external sensors, the *Electronic Mastering Device* (EMD) [23], must be screwed to a specific pawl on the axis being considered (starting from axis #1), and electrically connected to the robot.

3. A semi-automatic mastering program needs to be configured for the given axis using the graphics interface of the SP, and launched. This program controls the *mastering motion* for that axis. During the motion, a physical button on the back of the SP (the *enabling switch*) must be kept pressed for safety reasons. At the end, if operations were performed correctly, axis should be calibrated.

---

[1] The training department of KUKA, a worldwide manufacturer of IRs.

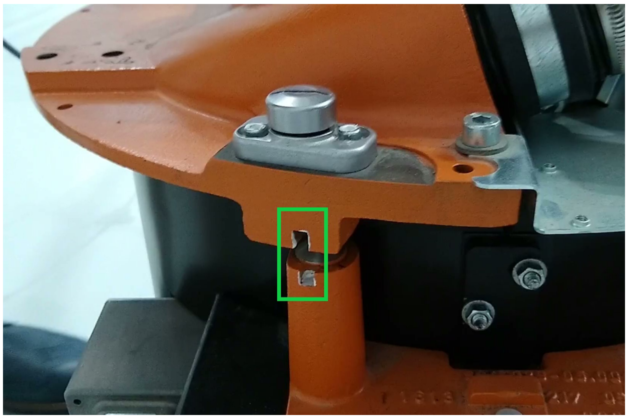


Figure 2: Axis #1 of a KUKA KR-16 robot in pre-calibration position. The green box highlights the visual references to be aligned by the operator in the first step of the MP.

4. The EMD must be disconnect and unscrewed (these operations need to be performed in this exact order to prevent damages).
5. Steps 2–4 must be repeated for all the robot's axes, following their numbering.

Those above are the steps to follow in the no-error case. However, the operator also needs to know several error recovery procedures (ERPs), and must be able to decide when to apply them based on the acquired procedure-dependent analytical and decision-making skills. In fact, no clear feedback is generally provided by the robot when errors occur; this is due to the nature of errors in the considered procedure, which are mainly detectable just at the end of the operations or as a result of an equipment damage. Moreover, differently than in other IMA tasks, errors in the MP cannot be corrected by simply rollbacking or undoing steps. These peculiarities translate into challenging didactic features that any training method shall manage in order to make the operators capable of completing the procedure autonomously.

Therefore, the key aspects an operator should be trained onto can be summarized as follows:

- acquire procedural knowledge about the MP;
- be able to recognize system status changes by relying only on visual and/or auditory cues, which are often rather subtle (e.g., very slow axis movements);
- be able to safely move around the robot by exploiting spatial awareness and proprioception skills;
- properly manage the equipment to prevent damages;
- handle errors by performing specific ERPs, to be selected based on the kind of error and the system status.

3.2. Standard Training

Master instructors from KUKA College arranged a training method that has proven (by internal audits) to assure the acquisition of expected learning outcomes by trainees, and is adopted in every training facility of the company.

As illustrated in Fig. 3, for each salient part of a typical course the trainees follow a three-phase learning process. The first phase is managed using a classroom lecture approach, with small classes of maximum 12 people. In this phase, the instructor may use several kinds of projected supports such as drawings, slides, videos, or SP emulators to back their teaching. The aim of this phase is to teach both theoretical and practical concepts (like procedures), as well as to prepare trainees for the activities on the robot that will be carried out in the next two phases. In the second phase the instructor shows, hands-on, how to perform procedures and practical exercises. A *didactic cell* (DC), i.e., a dedicated teaching area that includes an IR as well as different safety devices and props (depending on the topic taught) is used. In the third phase, trainees practice directly on the robot in teams of three. Every team operates in a different DC under the supervision of an instructor, who can intervene if/when needed. This supervised, "peer-tutoring" approach

(a) In-class learning (b) Hands-on demonstr. learning (c) Peer-tutoring with supervision

Figure 3: The three phases of the standard training (ST) at KUKA College.

was preferred to alternative configurations (e.g., with one trainee per DC) with the aim to benefit from the effects of cooperative learning dynamics [31] and optimize the allocation of limited physical resources (the DC).

For the purpose of setting up the user study reported in this paper, the method described above based on alternate in-class and hands-on phases was used to arrange a so-called *standard training* (ST) for the considered use case. Training consists of two modules, one on the the basic operation of a KUKA IR (teaching how to identify and recognize the robot's main elements, how to use the SP to move its axes, etc.), the other one on the MP. Both modules were extracted "as is" from the KUKA course on basic robot programming. The first module, in particular, was introduced in the experiment since the intention was to involve in the study subjects with scarce familiarity with IRs (as discussed later in Section 4). According to internal statistics of the KUKA College, the expected duration of the ST is estimated to be about 90 min.

3.3. VR Training

A state-of-the-art VRTS was developed, intended to be used as a self-learning tool for operators training on the same modules of the ST. In the following, its implementation will be described in detail; the design process, which included multiple iterations with continuous feedback from KUKA instructors, and the preliminary validation steps (based on several pilot studies) won't be included for brevity.

3.3.1. Technologies and Implementation

The VRTS was developed for being exploited using an immersive VR HMD and its hand controllers. Specifically, the HTC Vive Pro[2] kit and ecosystem were used in the study. This HMD features a display resolution of 1400×1600 pixels per eye, spanning a horizontal 110° FOV with a 90Hz refresh rate. The native positional tracking leverages the infrared lasers emitted by the so-called base stations (built upon the Valve's Lighthouse technology) which, combined with the HMD's built-in sensors, enables a 6DOF outside-in tracking over an area of up to 10m×10m. The user interacts with the VE using the tracked controllers and their physical buttons.

The VRTS application was implemented using the Unity[3] game engine and the SteamVR framework (to enable compatibility with other VR kits). 3D assets used to create the VE (in particular, to recreate a high-fidelity replica of the DC, including a KUKA KR-16 robot) were extracted from KUKA SimPro v3.0.5 and polished with Blender[4]. Appropriate care was taken in reproducing in VR also the audio cues from the real DC; specifically, sound recordings of physical robot movement and of interaction with props (like, e.g., safety doors) were harvested and included in the VE using spatialized 3D audio simulation.

3.3.2. User Experience

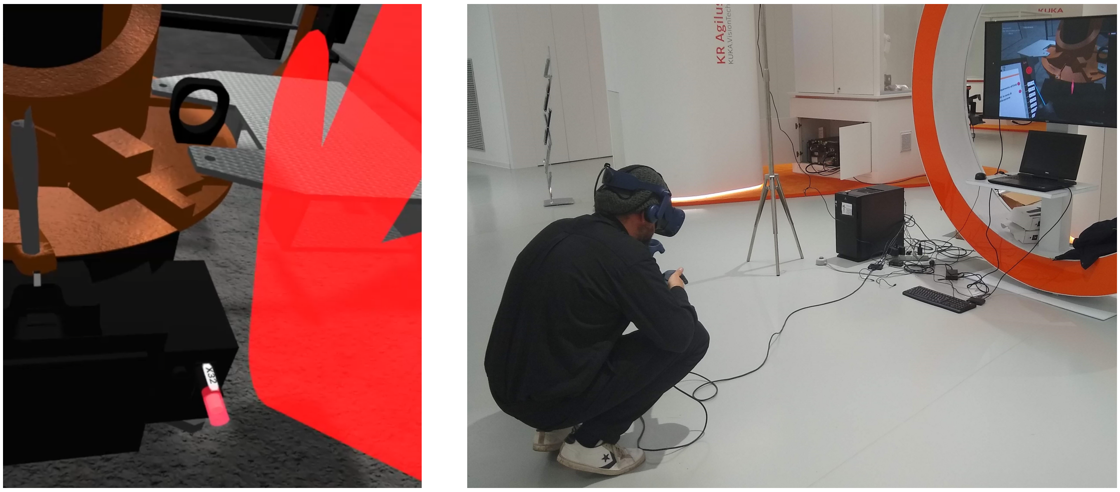
The VRTS was designed targeting a fruition time of approximately 40 min since, according to [20], trainees' performance in VR tends to deteriorate after 55 min. The self-learning user experience (summarized in Fig. 4) was created following the well-known approach, largely adopted in VRTSs [6, 8], which consists in providing step-by-step instructions by means of a guidance system (GS).

---

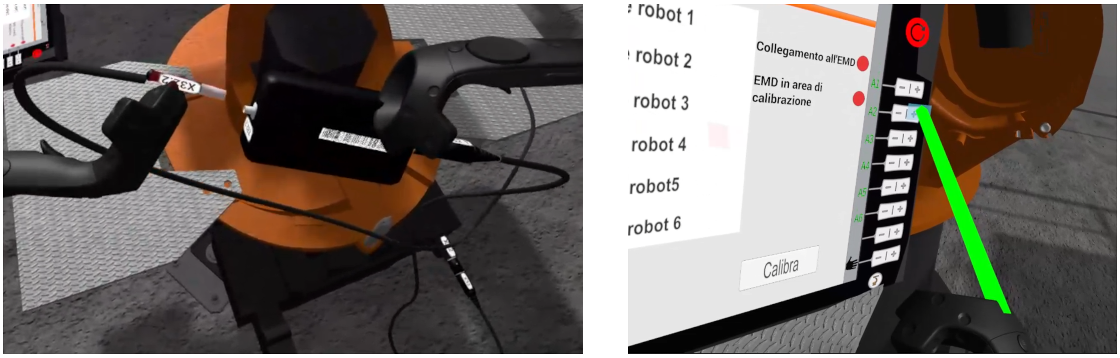
[2] https://www.vive.com/us/product/vive-pro-starter-kit/

[3] https://www.unity.com

[4] https://www.blender.org



(a) Example of MR-like hints (b) Trainee intent on aligning the robot's axis #1



(c) Interaction with virtual tools (d) Interaction with the virtual SP

Figure 4: Selected steps of the devised VRTS experience.

Differently than in the ST, in the VRTS the three phases are mixed in a single, interactive flow of information. The GS includes a voice-over, with pre-recorded tracks of a synthesized female voice (implemented using a state-of-the-art text-to-speech tool) delivering chunks of instructions. To stimulate the user to keep a high attention level, instructions are not delivered also in an alternative, duplicated form, such as using text (as done, e.g., in [40]). However, the user can replay the last chunk by pressing a button on the virtual SP.

The user's focus is shepherded onto so-called *hot-spots* by means of simulated MR-like hints (Fig. 4a) [39] that stimulate pre-attentive visual processing [14]; to this purpose, objects highlighting with blinking outlines, arrows and other graphics symbols are used. Examples of *hot-spots* could be an object to grab, a visual element to look at, etc.

In some steps of the MP, the user must also learn how to look at specific hot-spots from an appropriate point of view. This is enforced by the GS based on the user's head position and gaze direction, since in the various design iterations it was found that users tend to ignore the voice instructions when, e.g., an alternative, "quick and dirty way" of completing the assigned task exists. For instance, moving the robot to make it reach the pre-calibration position on axis #1 requires the user to have a close look at visual elements on the robot's basement, and this is possible only crouching near this *hot-spots* (Fig. 4b); it was observed that users tend to be negligent in this operation, especially if the GS accidentally provides additional feedback (not matched in the real scenario).

It is common for a step-by-step GS to be designed to block on each step. Considering the previous example, the GS could be programmed to invite the user to move on and align the next axis as soon as axis #1 has been correctly brought to its pre-calibration position. However, the experience would be very different from that with a real robot, since in that case the user wouldn't receive any feedback from the robot indicating that the task has been successfully

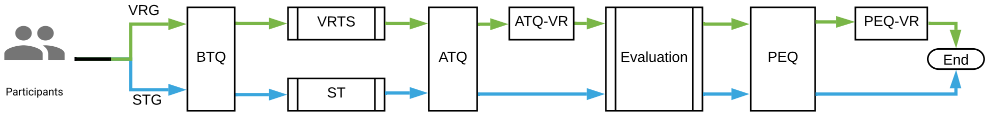


Figure 5: Experiment design and analysis tools.

accomplished. This situation must be accounted appropriately by the GS, postponing the check (e.g., after the next user action) to assure a correct transfer of related skills.

Further skills to acquire, as stated in Section 3.1, encompass being able to handle the involved equipment without damaging it. In particular, the user needs to learn how to manage cables and how to perform the necessary micro-manipulations with provided tools (Fig. 4c); for instance, a given connector may have to be screwed or fixed using a bayonet taking care of the orientations. Given the current limitations of technology for hand tracking and force feedback (at least, outside laboratory settings), these interactions were illustrated by using 3D animations showing the required actions, which are displayed on the given components when, e.g., two connectors get close enough with the correct orientation.

Furthermore, given the fact that high spatial awareness and proprioception are required to safely move around the robot, locomotion in the devised VE was mainly limited to *real walking*. When living the virtual experience in a real environment that matches the size of the virtual one (the DC, in this case), the user is allowed to freely move in VR, and this possibility is expected to preserve the sense of immersion. Additionally, the user is given the chance to use *teleportation*, but only to predefined locations that were carefully defined in the VE far enough from hot-spots, thus forcing anyway the user to reach such locations through real walking.

Finally, since acquiring dexterity with the SP is a critical part of the MP with a real robot, special scrupulousness was taken in implementing its simulated version. In particular, the safety-critical enabling switch was mapped to the gripper button of the hand controller; other interactions were managed through a ray-casting selection using the free hand (Fig. 4d), exploiting the trigger button of the respective controller to simulate interactions with the touch-screen.

A video recording illustrating the VRTS experience is available at http://tiny.cc/s2p6tz.

4. Experimental Settings

This section presents the design of the user study that was run to evaluate the devised VRTS against the two research questions in Section 1 by comparing it to the company's ST. The training was performed in a worst case-like scenario in which trainees are mostly domain agnostic.

4.1. Experiment Design and Metrics

The study's sample included 18 volunteers (16 males, 2 females) aged between 23 and 35 ($\mu = 25.33$, $\sigma = 2.62$), selected among students enrolled in engineering courses at Politecnico di Torino and the authors' network of contacts.

To avoid bias induced by learning effects, a between-subjects design was adopted. Participants were randomly assigned to two equal-sized groups. The first group, referred to as STG, underwent the ST; participants were split in teams of three for the second and third phases, as detailed in Section 3.2. The second group, named VRG, was trained with the VRTS; participants were allowed to practice interaction with objects and locomotion in the VE in a "sandbox" VR environment before actually getting into the VRTS.

After the training, each participant (from both groups) passed through an evaluation step in which the acquisition of the expected learning outcomes was assessed by a company's instructor through a quiz and the analysis of participant's performance in the execution of the MP on a real robot (based on the procedure that will be described in Section 4.2).

In addition to the measures about training effectiveness gathered by the instructor during the above evaluation, participants' feedback on the training experience was collected by means of a questionnaire including 109 statements articulated in three parts; participants were asked to respond by expressing their agreement with each statement on a 1-to-5 scale (from strongly disagree to strongly agree).

The first part, administered before starting the training (BTQ), contained items concerning demographics, previous knowledge and expertise with technology related to the experiment, as well as perceived *self-efficacy* (i.e., expectations from/attitudes towards the training to be experienced) [15]. The second part, administered after the training (ATQ), was made up of items adapted from the Instructional Materials Motivation Survey (IMMS) [19] and of statements concerning post-experience self-efficacy; in order to measure the participants' cognitive load, the NASA-TLX tool [13] was also included. For subjects in the VRG, this part included also a section (ATQ-VR) with items about usability of the VR application based on the SUS tool [2] and the VRUSE [17]. Finally, the third part consisted in a post-evaluation questionnaire (PEQ) that was meant to complement the participant's assessment made by the instructor. To this aim, statements about participants' self-evaluation (regarding the perceived ability to carry out the given task on the real robot) and overall satisfaction were used. Like for the ATQ, a section of this part (PEQ-VR) was reserved to subjects in the VRG and used to specifically investigate participants' satisfaction with the VR application as well as simulation fidelity. It is worth noticing that the results of the instructor's evaluation were provided to the participants only after having administered the PEQ in order to prevent conditioning. At the end of the experiment, open feedback about the experience was additionally collected.

The experiment design is illustrated in Fig. 5, whereas the full questionnaire is available at http://tiny.cc/l2p6tz.

4.2. Evaluation Procedure

As said, in order to compare training effectiveness in terms of knowledge and skill transfer capabilities, an evaluation procedure was specifically designed, structured as follows.

After having completed the training, the trainees were first requested to answer a multiple-choice quiz. The quiz included seven questions (each with five options, only one correct), which shall be answered in maximum five minutes. Questions were either borrowed from the KUKA's robot programming certification exam, or created ad-hoc together with the company's instructors. The quiz was aimed to investigate theoretical aspects of the MP, as well as to preliminarily check the acquisition of procedural knowledge.

Afterwards, the trainees were asked to perform the full MP on a real robot, working autonomously. It is worth recalling that, for subjects in the VRG, this was the first hands-on experience in the physical DC. The DC was prepared as follows:

- all the tools required for the procedure were left outside the cell in their carry-on box;
- the protective cups of the pawls on the robot's axes were removed;
- the safety door was securely closed;
- the robot was setup by giving each axis a $-10°$ offset w.r.t. their pre-calibration position.

During their performance, the trainees were observed and evaluated by a KUKA instructor, but they were not allowed to communicate. Instructor could intervene and stop (temporarily or permanently) the operations only for safety reasons, i.e., to prevent damages to people or equipment.

The evaluation was supported by a structured evaluation sheet, which allowed the instructor to record both objective measures and subjective scores. Objective measures included the overall task completion time, a set of time intervals on some milestones identified in the MP numbered from I to V (safety door unlocked, axis #1 in pre-calibration position, all axes in pre-calibration position, EMD correctly connected the first time, axis #1 mastered, all axes mastered), as well as the number of errors made. A list of the most frequent errors (split in two categories, i.e., major and minor, based on their severity) was included in the evaluation sheet, and the instructor was requested to annotate their possible occurrences. Regarding subjective aspects, the instructor was asked to grade on a 1-to-10 scale the trainee's operations considering the following perspectives: SP management, equipment management (cables, connectors and tools), ability to move around the robot, safety aspects, and overall performance. The evaluation sheet is available for download at http://tiny.cc/i2p6tz.

5. Results and Discussion

Data collected during the experiments were analyzed to spot significant differences between the two groups using two-tailed Mann-Whitney U-tests ($p = 0.05$). Here below, after presenting the population features, a discussion of the most relevant results (including all the statistically significant ones) is reported; discussion is organized along the two major dimensions set by the posed research questions, i.e., effectiveness of the standard and the VR-based training (essentially based on the outcomes of the evaluation step), and participants' experience with the two approaches (based on the questionnaire results). Afterwards, overall considerations and remarks are provided.

5.1. Population (BTQ)

According to information collected through the BTQ, 78% of the participants were not familiar with IRs, whereas 22% had little familiarity with them. Regarding immersive VR, 56% of the subjects in the VRG had never or just rarely used an HMD, whereas 44% of them stated to use this technology quite often. No statistical difference was observed between the two groups; hence, no bias was apparently introduced by the random participant selection.

5.2. Training Effectiveness

In the following, training effectiveness is analyzed by comparing the duration of the two experiences as well as the data collected in the evaluation step (summarized in Fig. 6), i.e., the scores assigned by the instructor, the time needed to reach the various milestones of the MP, the errors made and the acquired knowledge (assessed through the quiz).

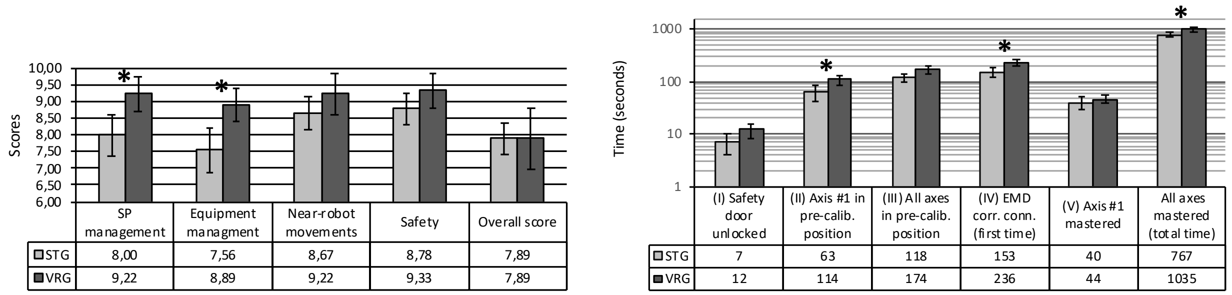
5.2.1. Training Time

For what it concerns the mere training time (i.e., not including the evaluation), on average the VRTS outperformed the ST in terms of efficiency, as training in the VRG was roughly 2.5 times faster than in the STG (36.3 ± 8.2 min compared to 90.7 ± 4.8 min).

5.2.2. Instructor's Evaluation Scores

According to Fig. 6a, the VRG performed particularly well, leveling off to the STG. All the subjects from both the groups were able to complete the procedure autonomously on the real robot, and no significant difference was observed in the *overall score* assigned by the instructor. The designed VRTS proved to be effective in coping with the challenges and requirements of the MP as defined in Section 3.1.

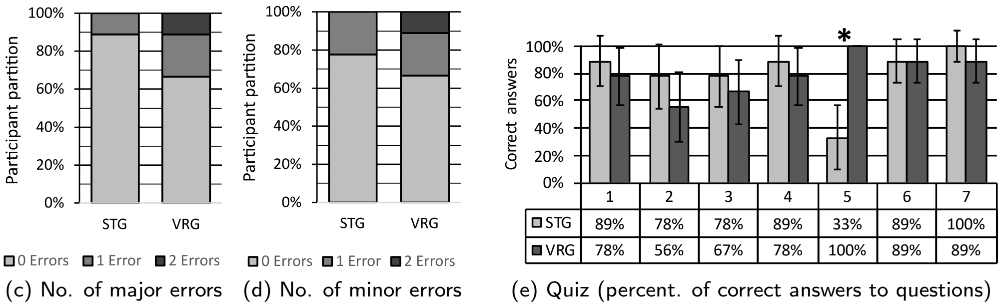
More specifically, for what it concerns *safety* and *near-robot movements*, the subjects in the VRG obtained scores comparable to those in the STG. The VRTS was even more effective than the ST in making the subjects learn how to manage the SP and other equipment. With regards to the *SP management*, better performance for the VRG can be explained by the fact that the evaluation procedure included aspects of the interaction with the SP's touch-based graphics interface which were easier to teach interactively by the GS than with the approach used in first and second phase of the ST (whereas in the third phase of the ST, trainees were generally less inclined to get into the details of the interface). Although based on obtained scores similar conclusions may be drawn also for *equipment management*,



	SP management	Equipment managment	Near-robot movements	Safety	Overall score
▨ STG	8,00	7,56	8,67	8,78	7,89
■ VRG	9,22	8,89	9,22	9,33	7,89

	(I) Safety door unlocked	(II) Axis #1 in pre-calib. position	(III) All axes in pre-calib. position	(IV) EMD corr. conn. (first time)	(V) Axis #1 mastered	All axes mastered (total time)
▨ STG	7	63	118	153	40	767
■ VRG	12	114	174	236	44	1035

(a) Instructor's evaluation scores (b) MP milestones (intervals)



	1	2	3	4	5	6	7
▨ STG	89%	78%	78%	89%	33%	89%	100%
■ VRG	78%	56%	67%	78%	100%	89%	89%

(c) No. of major errors (d) No. of minor errors (e) Quiz (percent. of correct answers to questions)

Figure 6: Data collected in the evaluation. Significant differences are marked with *.

at a closer look the situation appears to be slightly different. The average score of the VRG is indeed high enough to confirm the effectiveness of the VRTS also in this perspective, but this result should be probably ascribed to the fact that subjects in this group actually used the physical equipment for the first time during the evaluation; hence, they were probably way too careful in operating it.

5.2.3. MP Milestones Completion Time

Rationale for the latter results is provided also by the completion times measured for the MP milestones (Fig. 6b). Significant differences between the two groups, in particular with subjects in the VRG being slower than those in the STG, were spotted for milestones II and IV, i.e., those in which the majority of critical equipment management actions are required (the first milestone includes the time for getting familiar with the physical DC, whereas the second one encompasses the connection of several small props and the handling of delicate connectors). These differences were not observed for the other milestones, which mainly required the trainees to manage the SP (III and V).

The latter consideration is particularly relevant also for what it concerns another requirement stated in Section 3.1 about the ability of the VRTS to successfully and faithfully convey the visual and/or auditory cues required to recognize changes in system (mainly IR) status. Milestones III and V are actually those in which this procedural knowledge is particularly relevant.

Considering *overall completion time*, the STG was faster than the VRG (12.78 ± 1.37 min compared to 17.24 ± 2.20 min), but this difference can be considered as acceptable considering the scenarios in which the MP needs to be performed. Moreover, it is also worth observing that this slightly-less-than-five-minute time additionally needed, on average, by subjects in the VRG has to be compared with the approximately 50 min longer duration of the ST with respect to the VR-based training.

5.2.4. Errors and ERPs

Differences in completion times were also due to the fact that subjects in the VRG performed more ERPs than subjects in the STG. As shown in Fig. 6c and Fig. 6d, the average number of errors recorded for the former group was larger than for the latter group, although difference was not significant. Most of the errors made in both the groups were related to equipment management, and even when marked as major, they did not require the instructor's intervention since trainees were able to recover autonomously.

5.2.5. Quiz

Finally, concerning the quiz (Fig. 6e), significant differences between the two groups were found only for *question #5*. This question pertained procedural knowledge and asked trainees to sort the operations required to calibrate a single axis. It is interesting to observe that the VRG scored substantially better than the STG. However, it is also worth noting that in 67% of the cases, subjects in the STG who provided a wrong answer chose the second-best alternative. The difference between the two options was about the right time to press the enabling switch. The interpretations for this phenomenon could be manifold. One possibility could be that some of the subjects in the STG may have experienced an over-learning effect. These subjects may have learnt so well how to perform the operations that they started to carry out some of them without actually realizing it, hence supporting the interpretation that the transfer of skills was higher with the ST. A second (opposite) explanation could be that during the third phase of the ST, in which the instructor supervision is less prominent, subjects who had developed the bad habit of continuously pressing the enabling switch even when not needed erroneously transferred the safety property of this device to other functionalities of the SP. This interpretation would reinforce the already discussed superiority of VRTSs in backing the knowledge transfer for cyber-physical devices that are characterized by hybrid digital and haptic interfaces.

5.3. Training Experience

In the following, subjective feedback is analyzed by referring to the various parts of the administered questionnaire.

5.3.1. ATQ, ATQ-VR

Based on the NASA-TLX (Fig. 7), statistically significant higher scores were reported for the VRG than the STG regarding the *mental demand*, *effort*, *frustration dimensions*, and *overall score* dimensions, whereas comparable scores were associated to *physical demand*, *temporal demand*, and *performance*.

Among the factors analyzed through the ATQ (Table 1), significant differences in favor of the ST were found for what it concerns frustration (item #8), boredom (items #5 and #7), attention difficulties (items #3 and #6), perceived

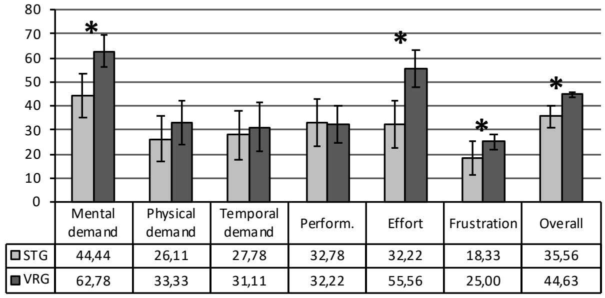


Figure 7: Cognitive load (based on NASA-TLX). Statistically significant differences are marked with *.

Table 1
Scores collected through the ATQ. Significantly better options (based on statement phrasing) are highlighted. Mean (SD) values for both the groups are reported, with *p*-values.

#	Statement/Item	STG	VRG	*p*-value
1	There was something interesting since from the beginning of this training experience that attracted my attention	3.89(0.31)	4.56(0.25)	.050
2	The approach used by the teacher was captivating	4.00(0.44)	4.00(0.41)	.931
3	The quality of information/instructional material helped to hold my attention	4.67(0.40)	3.56(0.25)	.004
4	The way information was provided helped keep my attention	4.33(0.33)	3.67(0.24)	.063
5	The amount of repetition in this training caused me to get bored sometimes	1.33(0.67)	3.44(0.53)	.001
6	The variety of information helped keep my attention on the training	4.22(0.43)	3.33(0.41)	.050
7	The information provided by the teacher was boring	1.56(0.46)	2.56(0.42)	.031
8	The amount of information provided was so large that the training was irritating	1.22(0.46)	2.33(0.47)	.019
9	The teacher showed the relevance of taught content	4.33(0.31)	3.56(0.25)	.024
10	It is evident how this information should/could be used after the training	3.89(0.42)	4.22(0.31)	.605
11	The amount of information provided was so large that it was difficult to identify the most relevant to remember	1.44(0.33)	1.89(0.37)	.258
12	During the training, I was confident that I would have been able to learn the taught content	4.44(0.37)	4.22(0.39)	.605
13	The training activities planned in the experience were too difficult	1.00(0.29)	1.67(0.43)	.067
14	During the training, I was confident that I would have been able to pass a test on it	4.11(0.31)	3.44(0.25)	.050
15	The training topics are relevant to my interests	2.78(0.72)	2.78(0.52)	.796
16	Completing this training successfully was important to me	3.00(0.60)	4.22(0.46)	.050
17	I am overall satisfied with the training experience	4.06(0.34)	4.00(0.28)	.605
18	I feel confident that I will correctly perform the MP on the real robot on the first try and without further help	3.78(0.37)	3.56(0.42)	.667

confidence to pass a test (item #14), and ability to stress the relevance of taught content (item #9). Conversely, compared to the ST, the VRTS was found to be more captivating (item #1) and able to better motivate the trainees (item #16).

Despite the differences reported above, the ST and the VRTS were considered comparable for all the other dimensions analyzed through the ATQ. This is the case of key aspects such as information clarity (item #10), overall satisfaction with the training experience (item #17), and perceived training effectiveness (item #18).

For the ATQ-VR (Table 2), worth noting results concern the usability of the VRTS, which was analyzed with the SUS tool (item #1) and was judged as remarkably high ("excellent", according to [1]), as well as the fact that it was found to be easy to learn (item #5) and capable of offering a high sense of presence (item #7).

5.3.2. PEQ, PEQ-VR

Consistent trends were observed also after the trainees had gone through the evaluation on the real robot. In fact, the paired items in the PEQ (Table 3) and the ATQ, i.e., the *overall satisfaction* (ATQ.#17, PEQ.#12) and *perceived training effectiveness* (ATQ.#18, PEQ.#1), still received comparable scores by the two groups.

Most of the remaining PEQ dimensions, i.e., the perceived capability of the considered training approaches to transfer required spatial knowledge and to support the acquisition of necessary proprioception skills (items #3, #4, #7, and #9) were also judged as comparable by the two groups. Furthermore, the VRTS and the ST were judged as equally effective in transferring procedural knowledge and equipment management abilities (items #2, #5, #6, and #10), although the subjects in the VRG struggled a bit more than those in the STG with operations involving small components (item #11) and felt more anxious when handling the provided equipment (item #8).

Finally, according to the PEQ-VR (Table 2), the simulation fidelity was considered as remarkably good.

Table 2
Section of the questionnaire on simulation perception dedicated only to subjects in the VRG. Statements with negative phrasing are marked with °.

#	Statement/Item	Mean (SD)
	ATQ-VR	
1	*SUS* Score	81.39(3.64)
2	I can see a real benefit in this kind of didactic experience	4.56(0.25)
3	The didactic experience was too simplistic to be useful°	1.78(0.52)
4	I wasn't aware of making mistakes°	1.78(0.31)
5	It was difficult to learn how to use the VR system°	1.44(0.25)
6	It was not so important to feel present to successfully complete the assigned tasks°	2.11(0.60)
7	In the VE I had a sense of "being there" (presence)	4.56(0.34)
8	The information was clearly presented	4.33(0.33)
9	The voice-over explanations (audio) were clear enough	4.33(0.33)
10	The MR-like visual cues helped understanding	4.78(0.21)
	PEQ-VR	
1	Overall, I think that simulation fidelity was very high	4.44(0.25)
2	I wished I had more time to spend free-practicing in the VRTS	3.00(0.62)

Table 3
Scores collected through the PEQ. Significantly better options (based on statement phrasing) are highlighted. Mean (SD) values for both the groups are reported, with p-values.

#	Statement/Item	STG	VRG	p-value
1	I felt confident while performing the MP on the real robot	4.33(1.00)	4.00(0.33)	.258
2	I felt the need to frequently stop throughout the MP to remember what should have been done next	1.67(0.87)	2.00(0.41)	.436
3	It was easy to safely move around the robot	4.44(0.73)	4.44(0.34)	.931
4	It was easy to locate the pawls on the robot's axes	4.00(1.12)	4.22(0.31)	.863
5	I struggled to use the SP	1.44(1.01)	1.11(0.16)	.666
6	The cable management and the connectors coupling/fastening was as expected	4.44(0.73)	4.11(0.28)	.297
7	Working on a real robot made me feel anxious	1.11(0.33)	1.33(0.24)	.436
8	I believe that I managed the equipment with the appropriate care	4.78(0.44)	4.22(0.21)	.050
9	It was easy to locate the visual references close to axis joints	4.44(0.73)	4.11(0.16)	.222
10	It was difficult to move all the axes in their pre-calibration position	1.78(0.83)	1.78(0.39)	.931
11	It was difficult to couple/fasten the connectors without risking to damage them	1.56(0.73)	2.33(0.33)	.050
12	I am still overall satisfied with the training experience	4.56(0.53)	4.44(0.25)	.730

5.4. Considerations & Remarks

For what it pertain training effectiveness, from results reported in Section 5.2 the VRTS emerged as a compelling alternative to the ST in terms of backing appropriately the transfer of knowledge (including spatial one) and the acquisition of necessary skills (also related to proprioception). In particular, all the subjects in the VRG were able to successfully complete the procedure on the real IR, obtaining scores on par or even superior w.r.t. subjects in the STG. Although these scores indicate, on the one side, the effectiveness of VR for training experiences involving the management of cyber-physical devices (the SP, in this case), on the other side they may be misleading for what it concerns the handling of small, physical equipment, as subjects trained in VR were too careful with it and, hence, slower in completing the MP than subjects who underwent the traditional training. Longer completion times for VR trainees are nonetheless compensated by much shorter training times.

As for the training experience, based on results discussed in Section 5.3 the ST and the VRTS were judged as comparable for most of the investigated aspects, including crucial ones such as the overall satisfaction with the received training and its perceived effectiveness.

Differences were nevertheless noted for some aspects. One of these differences pertains the cognitive load of the VRTS, which was found to be higher than that of ST. This result is in line with findings reported in the literature, as it is well known that immersive VR experiences, especially those in which rich interactivity and alluring audio-visual stimuli are involved, are more prone to induce a high extraneous cognitive load [32]. In the reported case study, the higher *mental demand* and *effort* scores of the VRG were probably affected also by the higher pace and amount of instructional content of the VRTS compared to the ST, whereas the usability of the VR-based system apparently did not played a significant role.

Although the workload was regarded as more than adequate and not overwhelming with both the ST and the VRTS (the *overall score* was below the 50th percentile for the categories *Robot operation* and *Video-games* according to

[10]), it should be considered as a factor to be further improved in the future, e.g., by dilating the virtual experience provisioning it in a sort of multi-session flow with out-of-VR breaks.

Another aspect that could be further ameliorated pertains the higher frustration levels reported by the VRG. Although ascribable also to collateral factors (e.g., the higher boredom and attention difficulties signaled for the VRTS compared to the ST), the open feedback collected from subjects in the VRG indicated that their scores for these aspects were mainly due to their experience with the voice-over (they would have preferred a real human recording, instead of text-to-speech) and to the fact that they judged the GS as not adapting to their training needs.

Social elements of the ST are probably the most interesting aspect that should be considered in the future to elevate the experience with the VRTS and make it match that with ST. An element supporting this claim could be the fact that, during the training, subjects in the STG perceived to be more confident that they would have been able to pass an evaluation on tackled subjects, probably because they were reassured and encouraged by the instructor when needed; another supporting element could be the fact that the VRTS was judged as less effective on stressing the importance of the different content being taught (as this result can be explained by the fact that human instructor in the ST integrated the lecture with examples based on their past-experience, stimulating rich Q&A interactive sessions).

It is interesting to note that, even though both the groups were equally interested in the training topics, the VRTS was capable of better motivating the subjects to complete the training compared to the ST. Moreover, subjects in the VRG were more captivated by the training experience than those in the STG, which is quite noteworthy considering that also subjects in the latter group should have remained impressed by their first approach to an IR.

6. Conclusions and Future Work

This paper presented a user study in which a state-of-the-art, immersive VRTS supporting self-learning of a light maintenance task on IRs was validated by comparing it to the established training practice based on in-class and hands-on sessions of an international company.

After being trained with the two different learning methods, study participants were asked to autonomously perform the learnt task on a real robot. Results obtained using subjective and objective metrics outlined that the efficacy of the VRTS was overall comparable to that of the ST in terms of making the trainees able to successfully complete the task. Moreover, the VRTS and ST proved to be equally capable of fostering the development of spatial awareness and proprioception skills, as well to support the transfer of the necessary procedural knowledge. Some design choices, though, made the VRTS induce a higher mental workload and frustration on the trainees. Despite that, trainees who went through the VR training were adequately instructed on how to make decisions based on the system (IR) status and to perform ERPs when required; they also acquired a superior ability in managing hybrid digital-physical devices. Task completion time on the real robot was longer for subjects trained in VR, but training dispensed by the VRTS was significantly shorter than the ST. Finally, VR-based training was judged as a pleasant and time-effective way to learn, though it fell behind the ST in aspects that are usually fortified by trainer-trainee social interactions.

Albeit the generalizability of these results is somehow limited by the modest sample size, the findings of this study could represent an important stepping-stone for the future experimentation of VRTSs for IR-related training and industrial applications in general. They also enabled the identification of areas requesting further research. For instance, attention should be paid to improving VRTSs' skill transfer efficacy in critical areas regarding interaction, particularly with small and delicate objects. Furthermore, efforts should be devoted to make VRTSs capable of better adapting to the trainees' learning pace (e.g., by adjusting the explanation flow to their needs), as well as of encouraging and motivating them throughout the training (e.g., by reacting appropriately to both errors and achievements).

Acknowledgements

This work has been developed in the frame of the VR@POLITO initiative. The authors want to acknowledge the help provided by KUKA Roboter Italia Spa and Anna Francavilla in the design and evaluation steps.

References

[1] Aaron, B., Philip, K., James, M., 2009. Determining what individual SUS scores mean: Adding an adjective rating scale. Journal of Usability Studies 4, 114–123.

[2] Brooke, J., 1996. SUS: A 'quick and dirty' usability scale. Usability Evaluation in Industry , 189.

[3] Brough, J.E., Schwartz, M., Gupta, S.K., Anand, D.K., Kavetsky, R., Pettersen, R., 2007. Towards the development of a virtual environment-based training system for mechanical assembly operations. Virtual reality 11, 189–206.

[4] Chittaro, L., Buttussi, F., 2015. Assessing knowledge retention of an immersive serious game vs. a traditional education method in aviation safety. IEEE Transactions on Visualization and Computer Graphics 21, 529–538.

[5] Chuah, S.H.W., 2018. Why and who will adopt extended reality technology? literature review, synthesis, and future research agenda. Literature Review, Synthesis, and Future Research Agenda , (in press).

[6] Dhalmahapatra, K., Das, S., Maiti, J., 2020. On accident causation models, safety training and virtual reality. International Journal of Occupational Safety and Ergonomics , (in press).

[7] Egger, J., Masood, T., 2020. Augmented reality in support of intelligent manufacturing–a systematic literature review. Computers & Industrial Engineering 140, 106195.

[8] Feng, Z., González, V.A., Amor, R., Lovreglio, R., Cabrera-Guerrero, G., 2018. Immersive virtual reality serious games for evacuation training and research: A systematic literature review. Computers & Education 127, 252–266.

[9] 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.

[10] Grier, R.A., 2015. How high is high? a meta-analysis of NASA–TLX global workload scores, in: Proc. Human Factors and Ergonomics Society Annual Meeting, pp. 1727–1731.

[11] Guo, Z., Zhou, D., Zhou, Q., Zhang, X., Geng, J., Zeng, S., Lv, C., Hao, A., 2020. Applications of virtual reality in maintenance during the industrial product lifecycle: A systematic review. Journal of Manufacturing Systems 56, 525–538.

[12] Hamilton, D., McKechnie, J., Edgerton, E., Wilson, C., 2020. Immersive virtual reality as a pedagogical tool in education: A systematic literature review of quantitative learning outcomes and experimental design. Journal of Computers in Education , 1–32.

[13] Hart, S.G., Staveland, L.E., 1988. Development of NASA–TLX (Task Load Index): Results of empirical and theoretical research. Advances in Psychology 52, 139–183.

[14] Healey, C., Enns, J., 2011. Attention and visual memory in visualization and computer graphics. IEEE Transactions on Visualization and Computer Graphics 18, 1170–1188.

[15] 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.

[16] Ho, N., Wong, P.M., Chua, M., Chui, C.K., 2018. Virtual reality training for assembly of hybrid medical devices. Multimedia Tools and Applications 77, 30651–30682.

[17] Kalawsky, R.S., 1999. VRUSE – A computerised diagnostic tool: For usability evaluation of virtual/synthetic environment systems. Applied Ergonomics 30, 11–25.

[18] 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).

[19] Keller, J.M., 1987. Development and use of the arcs model of instructional design. Journal of Instructional Development 10, 2.

[20] Kourtesis, P., Collina, S., Doumas, L.A., MacPherson, S.E., 2019. Validation of the virtual reality neuroscience questionnaire: maximum duration of immersive virtual reality sessions without the presence of pertinent adverse symptomatology. Frontiers in human neuroscience 13, 417.

[21] Leder, J., Horlitz, T., Puschmann, P., Wittstock, V., Schütz, A., 2019. Comparing immersive virtual reality and powerpoint as methods for delivering safety training: Impacts on risk perception, learning, and decision making. Safety Science 111, 271–286.

[22] López, C.E., Cunningham, J., Ashour, O., Tucker, C.S., 2020. Deep reinforcement learning for procedural content generation of 3D virtual environments. Journal of Computing and Information Science in Engineering 20, (in press).

[23] Mako, P., 2017. Principles of mastering at KUKA robots. International Scientific Journal about Machatronics 2, 7–12.

[24] Masood, T., Egger, J., 2019. Augmented reality in support of industry 4.0 – Implementation challenges and success factors. Robotics and Computer-Integrated Manufacturing 58, 181–195.

[25] Matsas, E., Vosniakos, G.C., 2017. Design of a virtual reality training system for human–robot collaboration in manufacturing tasks. International Journal on Interactive Design and Manufacturing 11, 139–153.

[26] Microsoft, . Dynamics 365. URL: `https://dynamics.microsoft.com/it-it/mixed-reality/guides/`. (Accessed: 10-10-2020).

[27] Mine, M.R., Brooks Jr, F.P., Sequin, C.H., 1997. Moving objects in space: Exploiting proprioception in virtual-environment interaction, in: Proc. 24th Annual Conference on Computer Graphics and Interactive Techniques, pp. 19–26.

[28] Molina-Carmona, R., Pertegal-Felices, M.L., Jimeno-Morenilla, A., Mora-Mora, H., 2018. Virtual reality learning activities for multimedia students to enhance spatial ability. Sustainability 10, 1074.

[29] Nathanael, D., Mosialos, S., Vosniakos, G.C., 2016. Development and evaluation of a virtual training environment for on-line robot programming. International Journal of Industrial Ergonomics 53, 274–283.

[30] Nazir, S., Sorensen, L.J., Øvergård, K.I., Manca, D., 2015. Impact of training methods on distributed situation awareness of industrial operators. Safety Science 73, 136–145.

[31] O'Donnell, A.M., 2006. The role of peers and group learning. Handbook of Educational Psychology , 781–802.

[32] Parong, J., Mayer, R.E., 2020. Cognitive and affective processes for learning science in immersive virtual reality. Journal of Computer Assisted Learning , (in press).

[33] Pérez, L., Diez, E., Usamentiaga, R., García, D.F., 2019. Industrial robot control and operator training using virtual reality interfaces. Computers in Industry 109, 114–120.

[34] Perret, J., Vander Poorten, E., 2018. Commercial haptic gloves, in: Proc. 15th Annual EuroVR Conference, pp. 39–48.

[35] 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.

[36] Román-Ibáñez, V., Pujol-López, F.A., Mora-Mora, H., Pertegal-Felices, M.L., Jimeno-Morenilla, A., 2018. A low-cost immersive virtual

reality system for teaching robotic manipulators programming. Sustainability 10, 1102.

[37] Tulving, E., Thomson, D.M., 1973. Encoding specificity and retrieval processes in episodic memory. Psychological Review 80, 352.

[38] Waller, D., Hunt, E., Knapp, D., 1998. The transfer of spatial knowledge in virtual environment training. Presence 7, 129–143.

[39] Werrlich, S., Daniel, A., Ginger, A., Nguyen, P.A., Notni, G., 2018. Comparing HMD-based and paper-based training, in: IEEE International Symposium on Mixed and Augmented Reality, pp. 134–142.

[40] Winther, F., Ravindran, L., Svendsen, K.P., Feuchtner, T., 2020. Design and evaluation of a VR training simulation for pump maintenance, in: Extended Abstracts Conference on Human Factors in Computing Systems, pp. 1–8.