Impact of avatar representation in a virtual reality-based multi-user tunnel fire simulator for training purposes

(Article begins on next page)

Impact of Avatar Representation in a Virtual Reality-based Multi-User Tunnel Fire Simulator for Training Purposes

Davide Calandra[1][0000-0003-0449-5752], Filippo Gabriele Pratticò[2][0000-0001-7606-8552], Gianmario Lupini[3], and Fabrizio Lamberti[4][0000-0001-7703-1372]

Politecnico di Torino, 10129, Turin, Italy
[1]davide.calandra@polito.it
[2]filippogabriele.prattico@polito.it
[3]gianmario.lupini@studenti.polito.it
[4]fabrizio.lamberti@polito.it

Abstract. Virtual Reality (VR) technology is playing an increasingly important role in the field of training. The emergency domain, in particular, can benefit from various advantages of VR with respect to traditional training approaches. One of the most promising features of VR-based training is the possibility to share the virtual experience with other users. In multi-user training scenarios, the trainees have to be provided with a proper representation of both the other peers and themselves, with the aim of fostering mutual awareness, communication and cooperation. Various techniques for representing avatars in VR have been proposed in the scientific literature and employed in commercial applications. However, the impact of these techniques when deployed to multi-user scenarios for emergency training has not been extensively explored yet. In this work, two techniques for avatar representation in VR, i.e., no avatar (VR Kit only) and Full-Body reconstruction (blending of inverse kinematics and animations), are compared in the context of emergency training. Experiments were carried out in a training scenario simulating a road tunnel fire. The participants were requested to collaborate with a partner (controlled by an experimenter) to cope with the emergency, and aspects concerning perceived embodiment, immersion, and social presence were investigated.

Keywords: Virtual Reality (VR) · multi-user simulation · emergency training · road tunnel fire · avatar representation

1 Introduction

Virtual Reality (VR) technology allows the creation of arbitrarily wide and sophisticated Virtual Environments (VEs), enabling the possibility to develop training scenarios that are very complex, or very expensive, to be deployed in real-life [11]. One of the most prominent fields which has taken large advantage of

VR is that of emergency training, as indicated by the huge number of literature works which investigated the role of this technology in the creation of effective scenarios for managing emergencies [1,13,28,23,20]). Among the various use cases investigated so far, firehttps://www.overleaf.com/project/617a6c84f00075d307ee0385 emergency is indeed among the most representative and widely studied ones [12,33,10,26].

As stated in [11], for near-future developments one of the most promising challenges posed by VR technology for firefighting training is represented by the transfer of findings from other domains regarding VR experiences involving multiple users. In multi-user scenarios, two or more users can engage from different locations and at the same time in a shared virtual experience [7]. Differently than in single-user experiences, in multi-user scenarios the avatar realism is critical for the development of collaborative VEs [2].

VR kits typically include an Head-Mounted Display (HMD) and two hand controllers, providing a synchronized visuomotor feedback to the user [24]. Hence, the only available sensory information is related to the position and orientation of the user's head and hands, which is not sufficient for full-body motion capture [27]. By relying on this information, most of the single-user commercial VR experiences usually show only a virtual representation of the hand controllers (e.g., SteamVR Home[1]) or, in some cases, two floating hands/gloves aligned with the real hands (e.g., Oculus First Steps[2]). A previous work explored different visibility levels for the user's own avatar in single-user experiences, and found no significant differences in terms of perceived embodiment between fully showing, partially hiding or not showing, along with the hand controllers, a virtual body for the VR user [24], confirming the above choices.

Although these reduced avatar representations appeared be sufficient from a first-person perspective, they may not be suitable for representing the other users' avatars in shared experiences. In particular, for emergency training scenarios, these techniques may have a negative impact on the perceived realism of the simulated scenario and, consequently, on the training efficacy.

An alternative technique to represent the avatar of a VR user consists in applying Inverse Kinematics (IK) to operate a body reconstruction targeted from head and hand sensors only [27]. From a first-person point of view, these techniques may increase the user's embodiment thanks to the visuomotor correlation [19], but may also worsen it when the estimated pose is characterized by a low accuracy [32]. When seen from outside in multi-user scenarios, IK techniques also require to correctly manage the users' legs, taking into account the user's motion in the VE. This can be done by procedurally generating the gait (like, e.g., in Dead and Buried[3]), or by blending the IK outcome with animations (like, e.g., in VRChat [4]). Thanks to the possibility to show and manage a full representation of the user's body, these techniques may be effective in guaran-

[1] SteamVR Home: https://store.steampowered.com/app/250820/SteamVR/
[2] Oculus First Steps: https://www.oculus.com/experiences/quest/1863547050392688
[3] Dead and Buried: https://www.oculus.com/experiences/rift/1198491230176054/
[4] VRChat: https://hello.vrchat.com/

teeing an appropriate level of immersion and embodiment, especially in case of realism-oriented, multi-user emergency simulations for training purposes.

The aim of this work is to study the impact of two avatar representation techniques, namely, the VR Kit only (no avatar, hereafter referred to as VK) and the Full-Body reconstruction obtained by blending IK and animations (hereafter referred to as FB), when used to represent VR users in a multi-user, emergency training experience. The scenario, named *FréjusVR* and presented in [7], consists of a VR road tunnel fire simulator provided with multi-user, multi-role and multi-technology capabilities. The scenario embeds a serious game that can be used as a training tool for firefighters and as a means for communicating correct procedures to civilians. This serious game, developed in the context of the PITEM RISK FOR[5] project, provides a good test-bench for the considered avatar representation techniques.

A user study involving 15 users was conducted, by requesting each of them to collaborate with another user (an experimenter) to respond to a fire emergency happening in the VE. All the study participants experienced both the avatar representation techniques, as they were asked to face two slightly different situations in the same simulation. Subjective measures were gathered through standard questionnaires to investigate the level of embodiment [14], social presence [4], as well as immersion and presence [16].

Results show that, even though the FB representation did not significantly increase the embodiment with respect to VK, it was judged as significantly better than the other technique in terms of ability to foster mutual awareness, mutual attention, mutual understanding, and immersion. Moreover, almost all the participants preferred the FB for representing the other user's avatar in terms of aesthetics and multi-player interaction. No clear winner emerged regarding the representation of the user's own avatar, neither in terms of usability and aesthetics, nor regarding the overall experience.

2 Background

Nowadays, the use of VR technology to simulate emergency situations has been widely studied [21,18,22,29]. Studies largely considered also the field of fire simulation. As reported in [22], tunnel fires are among the simulated scenarios that have been most commonly explored in the past, even though most of them were not exploited in multi-user experiences. To cope with this lack, a multi-user road tunnel fire simulator for training purposes was recently presented in [7], supporting multiple roles (civilian or emergency operators), various VR technologies (consumer VR kits, locomotion treadmills, and motion capture suits) that can be arranged in different configurations, as well as a real-time fire spreading logic, with the possibility to integrate Computational Fluid Dynamics (CFD) data for the smoke visualization.

As mentioned before, a critical aspect of multi-user collaborative VEs is represented by the level of realism of employed avatars, as testified by a relevant

[5] https://www.pitem-risk.eu/progetti/risk-for

number of works on the effects of the specific avatar visualization technique being adopted [2]. The authors of [30], for instance, presented an experimental method to investigate the effects of reduced social information and behavioral channels in immersive VEs with non-realistic FB avatars (mannequins). To this purpose, both physical and verbal interactions were executed in both VR and real-life, and then compared in terms of social presence, presence, attentional focus, and task performance. Results showed that the lack of realism of the humanoid avatars hindered the social interactions and possibly reduced the performance, although the authors stated that the lack of behavioral cues such as gaze and facial expressions could be partially compensated.

In [17], the main focus was to find out how useful was the implementation of high-fidelity avatars in a multi-user VR-based learning environment. In particular, both educators and students were endowed with the possibility to access a shared VE by means of avatars. The educator's avatar consisted of a high-fidelity representation (including facial cues and eye motion) and was motion-controlled in real-time by the educator. The student's avatars was implemented as not anthropomorphic in order not to draw attention from the educator. The results of the study suggested that representing with high-fidelity avatars the subjects with important roles in the simulation can enhance the overall user experience for everyone participating in it.

A similar approach was pursued in [3] regarding the use of avatars in rehabilitation scenarios, though with a different set of technologies. By using a Microsoft Kinect sensor coupled with virtual scenarios, the authors concentrated their efforts on reproducing human posture failures by monitoring avatars, with the aim to improve the posture of people in different rehabilitation stages. Results confirmed that the framework created had enough flexibility and precision, thus confirming that avatar representation can play a key role in a wide variety of scenarios.

As reported in [31], several studies underlined that FB and Head and Hands avatar representations are the most widely used in the current body of literature. More broadly, the authors of [25] investigated the realism that an avatar should have to sufficiently appease the user's tastes. Work in the field led to the definition of a standardized Embodiment Questionnaire [14] and to a branch of studies regarding avatar use in multi-user environments, focusing on factors such as social presence and social interactions [4].

Regarding the visualization of the avatar from a first-person point of view in VR, the authors of [24] studied the effect of the visual feedback of various body parts on the user experience and performance in an action-based game. The work considered as visual feedback a completely hidden body (except for the VK), a low visibility body (hands and forearms), and a medium visibility body (head, neck, trunk, forearms, hands, and tail for the lower limbs). Differently than some previous works, no significant differences between the three alternatives were observed in terms of perceived embodiment.

Finally, regarding the research topic about avatar movements, works such as [27,9] showed the promising capabilities of IK techniques for estimating the pose of humanoid avatars, whereas studies like [15] managed to achieve accurate results by using such IK methods on different body parts, in particular on head and hands, leaving the lower body movement to several animations cleverly blended together.

Despite the wide number of works regarding the topic of avatar representation techniques for single and multi-user VR experiences, to the best of the authors' knowledge, investigations about the impact of these techniques in the field of VR simulations for emergency training are still scarce.

3 FrèjusVR

In this section, the fire training scenario used for the experimental activity is described, along with its configuration and the customization introduced for the purpose of the evaluation.

3.1 Devices and Technologies

The application was developed with Unity 2018.4.36f and the SteamVR Software Development Kit (SDK), allowing the deployment to any OpenVR compatible VR system. Due to the wide extension of the depicted scenario (a road tunnel), additional stationary locomotion techniques [8] were included to overcome the limitations of room scale movements. Among them, the arm-swinging technique was selected for the considered experimental activity, since it did not require additional hardware and showed to outperform the other techniques in some previous investigations [5,6].

3.2 Multi-user

Regarding the multi-user capabilities, the Unity legacy high-level network API (U-NET) was used to support a client-server architecture. The host can be either one of the user or a dedicated non-VR machine to lower the computational load of the two VR clients. To complement body-to body communication, a VOIP channel is established between the users though the Dissonance VOIP asset for Unity[6], adding two additional UNET channels (one reliable and one unreliable). Position and rotation updating of network objects is performed at 60Hz, and interpolation is employed to smooth the transition between consecutive updates. The scenario can support different roles, i.e., civilians, firefighters, and truck driver(s), which can be either played by real users, Non-Player Characters (NPCs), or be deactivated.

[6] Dissonance Voice Chat: https://assetstore.unity.com/packages/tools/audio/dissonance-voice-chat-70078

3.3 Avatar Representation Techniques

The training scenario provided two different avatars (male or female) to be chosen as civilian or truck driver, and a generic firefighter avatar for users playing as firefighting operators. Independent of the avatar selection in the main menu, the VR application was modified in order to allow the configuration of one of the avatar representation techniques considered in this study (Fig. 1), which are described below.

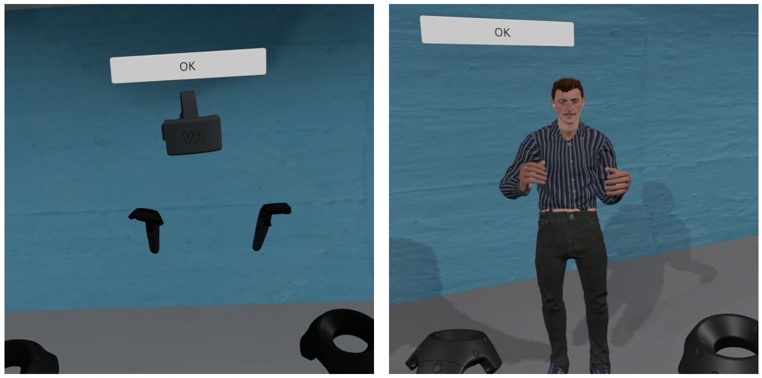
VR Kit (VK) This technique did not require particular modifications to the application. For the own avatar, the SteamVR CameraRig Unity prefab automatically manages the visualization of the VR hand controllers and their real-time synchronization with the real ones. For the remote users, 3D meshes representing a generic VR HMD and the two hand controllers (VIVE wand[7]) were displayed in correspondence of the synchronized position and orientation of the other user's head and hands (Fig. 1a).

Full-Body (FB) To support the FB representation, the Unity asset named FinalIK by RootMotion[8] was acquired and integrated in the scenario. The asset provides a VR-oriented FB IK (VRIK) solution, which supports both procedural locomotion steps, recommended for micro-movements on the spot, as well as an animated locomotion, and was designed for room-scale movements or faster techniques. Since the navigation of the tunnel scenario requires relatively fast virtual movements, the procedural locomotion of VRIK was discarded in favour of the the animated one (Fig. 1b).

For the local user, VRIK was configured to always apply the IK algorithm to the upper limbs, spine and head, whereas animated locomotion was automatically triggered when the HMD is moved on the horizontal (X, Z) plane, and the blending was adjusted on the basis of the movement speed. By default, the blended animation are managed with the Unity Mecanim animation system, through a 2D free-form directional blend-tree integrating a set of standing animations (including idle, directional walking and running). Animations for movements in crouch position were not provided by the assets. For this reason, a second blend-tree implementing crouched movements was added, and the blending between the two trees was obtained on the basis of the user's HMD position, normalized with respect to his or her height. Since the mentioned asset does not apply IK to fingers, hands were managed with an overriding layer in the Unity animator, in order to support hand gestures (i.e., open hand, fist, and pointing) which were activated based on the context. As a matter of example, if the user is pressing a button, the pointing animation will be temporarily showed. Similarly, if the user is walking around with the arm-swinging technique, the relative avatar will be displayed with his or her hands balled in tight fists.

[7] Vive wand controller: https://www.vive.com/eu/accessory/controller/

[8] https://assetstore.unity.com/packages/tools/animation/final-ik-14290



(a) VR Kit (VK). (b) Full-Body (FB).

Fig. 1. The two avatar representation techniques considered in the evaluation, as displayed to the participant in the form of mirrored avatar before starting the simulation.

To manage the representation of the other user, the standard VRIK behaviour had to be modified to maintain a sufficient level of naturalness. In particular, to hide out the unnatural arm-swinging gesture while walking, the upper body IK was temporarily disabled by blending it with the full movement animation. Hence, when the additional locomotion technique is triggered by a user, his or her avatar on the other user's machine will perform a FB walking animation. As soon as the user stops triggering the arm-swinging, the upper body IK is restored, and the animation is again applied to the lower limbs only.

To avoid misalignments between grabbed objects and FB animations, the grabbing position is always adjusted to either coincide with the actual controllers (when the IK is enabled) or with the hands of the other avatar's rig (when the arm-swinging is triggered and the FB animation is displayed).

VRIK allows to manage the synchronization over the network of the FB IK by synchronizing the position and orientation of the user's CameraRig (head and controller). To synchronize the additional functionalities mentioned before, a custom UNET network component was used.

3.4 Selected Procedure

As said, the considered training scenario provides different roles (civilian, firefighter, truck driver), each characterized by its own procedure and interactions (some of them depicted in Fig. 2). For the purpose of the current evaluation, it was decided to focus on the role of the civilian for both the users. The reasons behind this choice are manifold. The civilian role does not require previous knowledge (differently than for firefighters and, to some extent, the truck driver), it maximizes the percentage of time in which the two users can see each other and interact together (the two civilians can start the experience inside the same car), and provides the highest level of flexibility for the execution of the procedure.

In order to guarantee a good level of visibility, the CFD-based smoke simulation was disabled. Moreover, to avoid possibly confounding factors related to the visualization of the other peer, some NPC roles were disabled (the firefighters) or reduced to aesthetic features visible just at the end of the experience (the truck driver in the security shelter). Considering these modifications with respect to the complete experience detailed in [7], the considered procedure can be summarized as follows:

1. Both the civilians start on the same car, travelling from Italy to France, while the car radio is broadcasting the usual messages for tunnel users.
2. After some travel time during which the car occupants can communicate, and the passenger has access to the security brochure of the tunnel, a truck on fire is spotted in the opposite lane. In this situation, the driver can operate the brakes and stop the car at an arbitrary distance from the vehicle on fire.
3. As soon as the car stops, the two users can interact with the car interior, e.g. turning off the engine (Fig. 2a), enabling the hazard lights, or getting out of it using the doors' handles.
4. Once outside the car, the two collaborating users may press one of the many SOS buttons (Fig. 2b) placed inside the tunnel to signal the accident, and then decide whether to head to the closer SOS shelter, which is beyond the truck, or turn back and reach a farther shelter. It should be noted that after getting out of the vehicle, a second car accident involving two civilian cars occurs behind the users' vehicle, starting a second, more contained and less threatening fire. Hence, both ways will be partially occluded by damaged vehicles on fire.
5. If the two users opt for the first choice, they can take advantage of two SOS niches (Fig. 2c) on both sides of the tunnel right before the truck, provided with SOS telephones and extinguishers which can be freely used. The users can either try to extinguish the main fire (Fig. 2d), or directly run towards the selected shelter. Behind the truck on fire, the users will have to walk crouched to walk under a wooden plank (Fig. 2e) which fell, with others, from the truck load over a civilian car blocking the way to the shelter.
6. If the users head back to the other shelter, they will be again forced to crouch to pass under a metal rod (Fig. 2f) placed nearby the second accident. In this case, the closest extinguisher and SOS telephone will be inside the shelter. The users will be allowed to get back to the tunnel to try to extinguish the smaller fire. In this shelter, the users will meet an NPC character sitting on a bench near a locker containing a first aid kit and some water bottles.
7. In both cases, after opening the door one of the shelter (Fig. 2g), getting inside (Fig. 2h) and asking for help with the SOS telephone, the simulation ends (if the call was already done from a SOS niche in the tunnel, this step is not needed and the simulation is quickly terminated).

The possibility to extinguish both fires was disabled; however, the users were not made aware of this aspect, and were left free to choose whether to try using the extinguishers, fail, and thus continue with the evacuation. The layout of the described version of the tunnel scenario is reported in Fig. 3.

(a) Driver turning off the car engine. (b) User pressing a SOS button. (c) Both users interacting with a SOS niche. (d) Users collaborating in the extinguishing of the fire.

(e) Users crawling under the wooden board to get to the closer shelter. (f) Users crawling under the metal rod heading to the farther shelter. (g) Users interacting with the shelter door. (h) Users entering a SOS shelter.

Fig. 2. Pictures of the the FrèjusVR scenario taken during the experimental phase.

4 Experimental Setup

In this section, the setup used for the experimental activity and the adopted evaluation criteria are thoroughly explained.

4.1 Participants

The 15 participants (14 males, 1 female) were aged between 24 and 67. Most of them reported medium to high experience with video-games, VR and multiplayer applications, but almost all of them had little to no experience with serious games for emergency training.

4.2 Hardware

For the experiment, two HTC Vive Pro[9] kit were employed, one worn by the participant, one by the experimenter playing the part of the second user. The VR scenario was run on two Intel i9-9820X machines, each equipped with 32GB of RAM and a NVIDIA GeForce RTX 2080 Ti video card.

[9] HTC Vive Pro: https://www.vive.com/eu/product/vive-pro/

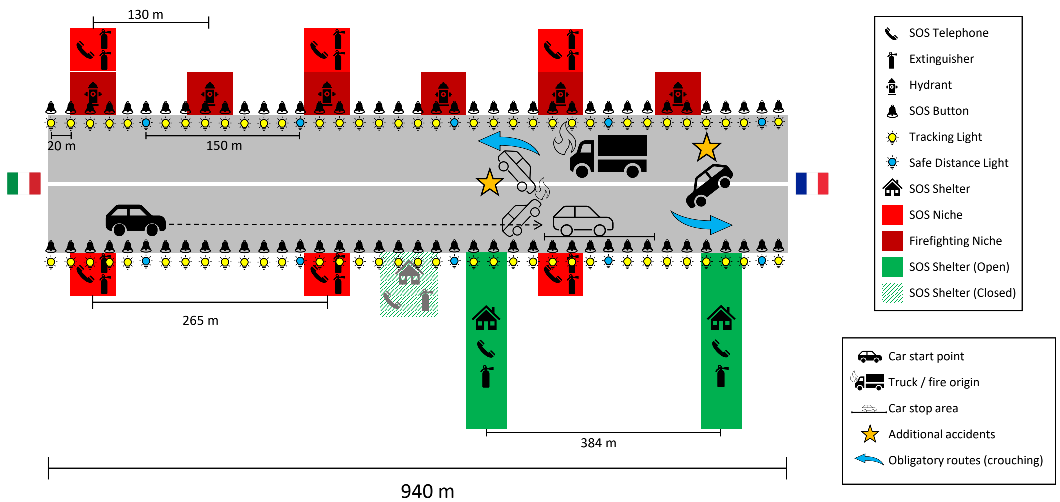


Fig. 3. Layout of the modified version of the tunnel scenario with respect to the original one detailed in [7].

4.3 Methodology

The experiment was designed as a within-subjects study, with the avatar representation technique as independent variable. The participants were initially asked to fill in a demographic questionnaire aimed to assess their previous experience with the involved technologies. Then, they were introduced to the experiment, and told what they were supposed to do in the simulation. A sample footage of the experiments with both modalities is available to download[10]:

Each participant experienced the two representation techniques in a random order. Before starting each simulation run, they were asked to select, inside the VR application, one of the two available avatars for the civilians. After that, a mirrored version of the avatar with the currently used technique was displayed to the user to show how he or she was going to be seen from the other user's point of view. With the VK, the choice of the avatar was not relevant, since it was not going to be displayed as a virtual body, but only as 3D meshes representing the VR equipment.

Then, for both the techniques, the participant was requested to choose the role of the driver. At this point, a second user, controlled by one of the experimenters, connected to the multi-user session and spawned as a second civilian sitting beside the driver, automatically starting the experience.

[10] http://tiny.cc/zmnnuz

During the simulation, the experimenter tried to follow a predefined set of actions, in order to force the other peer to perform all the intended actions and interactions, as well as to uniform the experience among the various participants. To limit learning effects, the experimenter followed two distinct scripts for the first and the second experience, whose main difference concerned the selection of the evacuation route. In particular, in the first run, the experimenter:

1. Performed interactions while observed by the participant.
2. Suggested to head to the closer shelter beyond the truck.
3. Ensured that the participant notices the SOS niches and interacts with them.
4. Showed the participant how to surpass the main fire and the obstacles requiring to crouch.
5. Should the participant try to extinguish the fire, he or she encouraged him or her to give up after a while and reach the shelter.

During the second run, the experimenter:

1. Suggested to turn back and head to the farther shelter beyond the car accident.
2. Showed the participant how to surpass the main fire and the obstacles requiring to crouch.
3. After reaching the shelter, he or she suggested to equip the extinguisher and try to deal with the smaller fire encountered on the way.
4. After a brief try, he or she suggested to head back to the shelter and wait there for the rescue team to arrive.

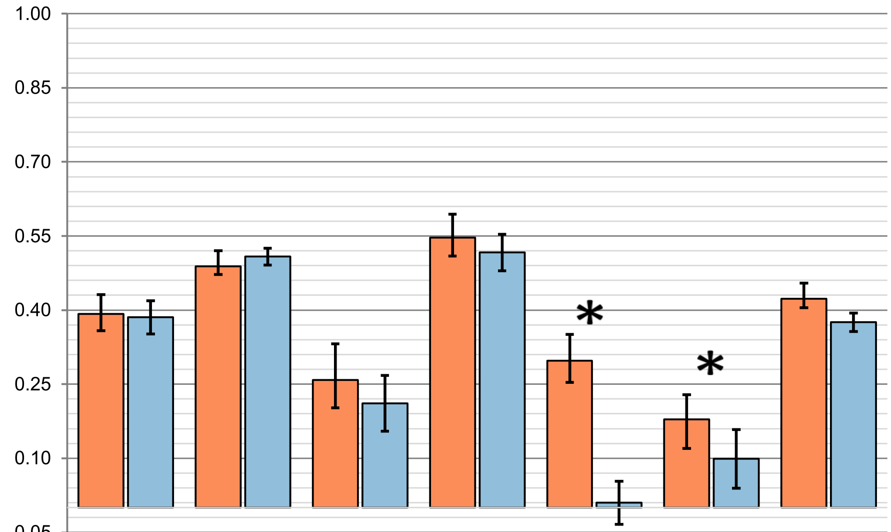
It should be noted that these guidelines were not considered in a very strict way, as some interactions may be repeated, omitted or executed in a different order, depending of the participant's collaborativeness during the simulation. After each run, the participant was asked to fill in the evaluation questionnaire detailed in the following.

4.4 Evaluation Criteria

As mentioned before, the participants were evaluated from a subjective perspective by means of a post-test questionnaire[11].

The questionnaire was organized in four sections. In the first section, the participant was asked to fill in the Embodiment Questionnaire[14], in order to evaluate his or her level of embodiment in terms of body ownership, agency and motor control, tactile sensations, location of the body, external appearance and response to external stimuli (as in [24]). The second section corresponded to the Networked Minds Social Presence Questionnaire [4], and it was aimed to assess the virtual representation of the other user's avatar in terms of mutual awareness, attentional allocation, mutual understanding, behavioural interdependence, mutual assistance, and dependent actions (similarly to [30]); the only

[11] http://tiny.cc/1nnnuz



	Ownership	Agency	Tactile Sensation	Location	Appearence	Response	Total Embodiment
FB	0.39	0.49	0.26	0.55	0.30	0.18	0.42
VK	0.39	0.51	0.21	0.52	0.01	0.10	0.38

Fig. 4. Average results for the Embodiment Questionnaire[14] sub-scales (values normalized between 0 and 1). Statistically significant differences (p-value < 0.05) marked with a star (*) symbol.

exception was the empathy category, which was not included, being the relative items not suitable for the considered use case. The third section included the immersion and presence section of the VRUSE[16] questionnaire. Finally, the last section, filled in after having completed both the runs, asked the participant to express his or her preference between the two representation techniques in terms of usability, aesthetics, multi-player interactions, and overall.

5 Results and Discussion

The results obtained for the subjective metrics presented in the previous section were used to compare the VK and FB techniques.

The Shapiro-Wilk test was used to analyze the normality of data. Since data were found to be non-normally distributed, the non-parametric Wilcoxon signed-rank test with 5% significance ($p < 0.05$) was used for studying statistical differences.

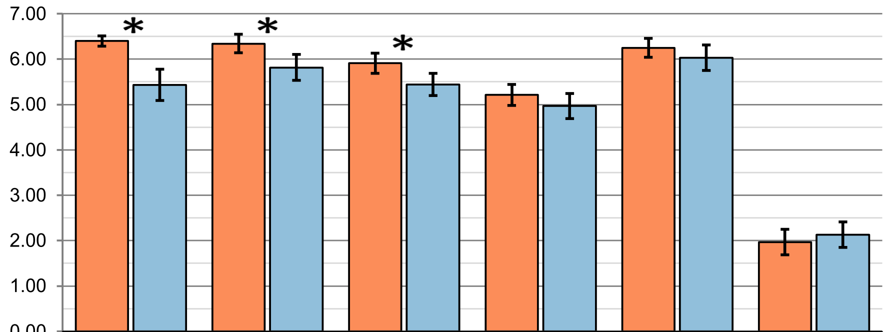
5.1 Embodiment

For what it concerns the perceived embodiment, reported in Fig. 4, no significant differences were observed in most of the sub-scales, as well as for the total embodiment. For the sake of readability, the sub-scales and the total embodiment, calculated as suggested in [14], have been normalized in a range between 0 and 1, being each of them originally characterized by different minimum and maximum values. The only exceptions are represented by the appearance (0.3 vs 0, p-value < 0.001) and response to external stimuli (0.18 vs 0.1, p-value $= 0.003$) sub scales, for which the FB was perceived as better than the VK. This result is in line with some previous literature works mentioned before, which did not find significant differences between showing or not showing the user's avatar body [24].

Going into the details of scores assigned to the individual items, expressed on a 7-point Likert scale from -3 to 3 (from strongly disagree to strongly agree), the participants felt the FB as a representation of their body more than the VK (1.4 vs 0.27, p-value $= 0.042$); however, they also perceived a higher sense of having more than one body (0.73 vs -0.6, $p = 0.01$), confirming that the FB could have either positive or negative effects in terms of embodiment based on its accuracy [19,32]. This is in agreement with the results got from the item, according to which, with the FB, the participants felt as if movements of the virtual representation were influencing their real movements more than with the VK (-0.13 vs -1.13, $p = 0.0488$). Moreover, concerning the initial part of the experience in which a mirrored version of the user's avatar is displayed to the participant, FB appeared to be felt as the own body more than VK (0.73 vs -0.6, $p = 0.0412$). Regarding external appearance, with the FB the participants felt as if their real body was becoming an "avatar" body more than with the VK (0.8 vs -0.6, $p = 0.003$), that their real body was starting to take the posture of the avatar body (0.2 vs -1.13, $p = 0.009$), that at some point the virtual representation started to resembled more their real body in terms of shape and other visual features (0.13 vs -2.0, $p = 0.003$), and that they felt more like they were wearing different clothes than when they came to the laboratory (0 vs -2.26, $p = 0.005$) than with VK. For what it concerns the response to external stimuli, with the FB the participants felt as if virtual elements (fire, objects) could affect them more than with the VK (1.2 vs 0.33, $p = 0.015$), and had an higher feeling of being harmed by the fire (-0.07 vs -0.53, $p = 0.015$).

5.2 Social Presence

As for social presence, whose metrics are depicted in Fig. 5, the difference between the two techniques was more marked. Responses were given on a 7-point Likert scale in a range between 1 and 7 (from strongly disagree to strongly agree). Considering the various sub-scales, the participants perceived the VK as better than the FB in terms of mutual awareness (6.4 vs 5.43, $p = 0.003$), mutual attention (6.34 vs 5.81, $p = 0.003$), and mutual understanding (5.91 vs 5.44, $p = 0.047$). These results indicate that the use of a body to represent the other user's avatar significantly improves multi-user cooperation. In particular, with the FB, the participants noticed more the presence of the other peer (1.33 vs 3.0, $p = 0.001$), they were more aware of themselves inside the VE (6.4 vs 4.93, $p = 0.002$), and they perceived the other peer as more aware of them (6.33 vs



	Mutual Awareness	Mutual Attention	Mutual Understanding	Behavioural Interaction	Mutual Assistance	Dependent Action
FB	6.40	6.34	5.91	5.21	6.25	1.97
VK	5.43	5.82	5.44	4.97	6.03	2.13
	Co-Presence	*Psychological Involvement*		*Behavioural Engagement*		

Fig. 5. Average results for the Networked Minds Social Presence Questionnaire [4] subscales (7-point Likert scale from strong disagreement to strong agreement). Statistically significant differences ($p < 0.05$) marked with a star (*) symbol.

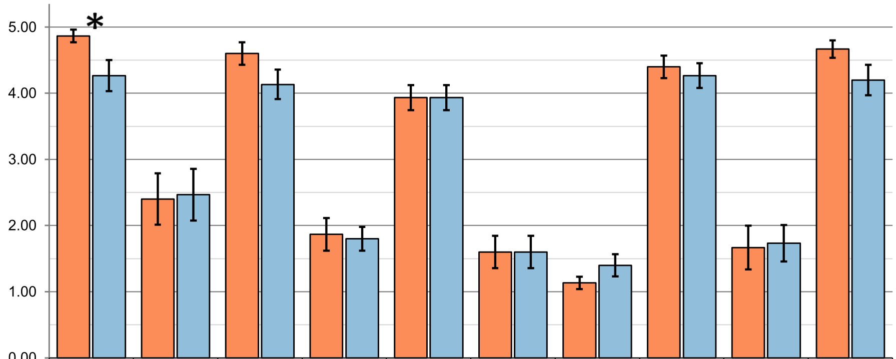
5.4, $p = 0.023$) than with the VK. Moreover, with the FB, the participants felt less alone (1.2 vs 1.93, $p = 0.046$), and perceived the other peer as less lonely (1.4 vs 2.2, $p = 0.039$). The FB also helped the participants to pay higher attention to the other peer with respect to the VK (6.33 vs 5.47, $p = 0.0312$). Furthermore, with the VK, the participants tended to ignore the other individual more than with the FB (1.67 vs 2.73, $p = 0.0117$). Finally, FB allowed participants to express their opinions (6.00 vs 5.27, $p = 0.0027$), as well as to understand the other peers' ones (5.93 vs 5.13, $p = 0.04888$) more than VK.

5.3 Immersion and Presence

The results for the immersion and presence section, derived from [16], are reported in Fig. 6. According to the scores, expressed on a 5-point Likert scale between 1 and 5 (from strongly agree to strongly disagree), the FB was judged as better than the VK only in terms of immersion (4.86 vs 4.26, $p = 0.031$), but nothing can be said in terms of presence. This outcome can be related to the fact that the training experience, being oriented towards realism, benefited of a more realistic-looking avatar representation, which was provided by the FB.

5.4 Direct Comparison

In the final section of the questionnaire, whose results are provided in Fig. 7, the participants were asked to express their preference between the VK and the FB for various aspects.



	I felt a sense of being immersed in the virtual environment	I did not need to feel immersed in the virtual environment to complete my task	I got a sense of presence (i.e. being there)	The quality of the image reduced my feeling of presence	I thought that the field of view enhanced my sense of presence	The display resolution reduced my sense of immersion	I felt isolated and not part of the virtual environment	I had a good sense of scale in the virtual environment	I often did not know where I was in the virtual environment	The overall system's given sense of presence was satisfactory
FB	4.87	2.40	4.60	1.87	3.93	1.60	1.13	4.40	1.67	4.67
VK	4.27	2.47	4.13	1.80	3.93	1.60	1.40	4.27	1.73	4.20

Fig. 6. Average results for the immersion and presence section of the VRUSE [16] questionnaire (5-point Likert scale from strong disagreement to strong agreement). Statistically significant differences (p-value < 0.05) marked with a star (*) symbol.

Interestingly, the FB was judged as significantly better than the VK for the other user representation for the aesthetics (93.33% vs 6.66%, $p = 0.001$), regarding multi-player interactions (86.66% vs 13.33%, $p = 0.010$) and also overall (86.66% vs 13.33%, $p = 0.010$). However, none of the two techniques prevailed for what it concerned the own avatar representation. These mixed results for the own avatar are anyway in line with those regarding the embodiment, suggesting that the two representations of the own avatar may have a similar impact in multi-user experiences too.

6 Conclusions and Future Work

In this work, two different techniques to represent users' avatars in multi-user VR experiences were evaluated in the context of an emergency training simulation. In particular, a road tunnel fire simulation, presented in [7], was used as a test-bench for the considered evaluation. The first considered technique, widely used for the user's representation in commercial, single-user VR applications, consists in not providing an avatar body, but only displaying the VR Kit (VK) equipment, i.e., the hand controllers and, in case of the other user, the HMD. The second technique, referred to as Full-Body (FB), makes use of a combination of IK algorithms and animation blending to operate a full humanoid reconstruction for the user's avatar, targeted to the position and orientation of the user's head and hands. A within-subject user study, involving 15 participants, showed different outcomes regarding how the own avatar and the other user's avatar

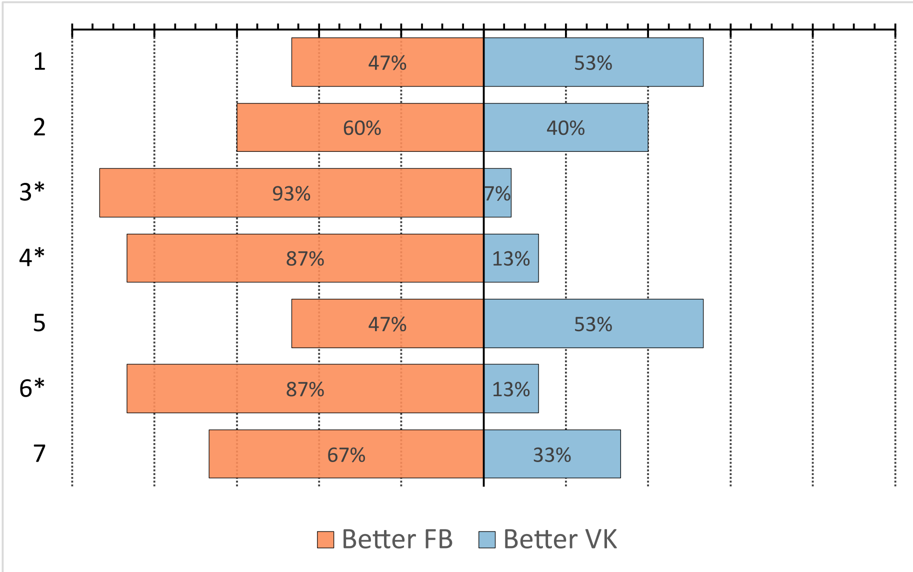


Fig. 7. Average results for the direct comparison section of the questionnaire. Statistically significant differences ($p < 0.05$) marked with a star (*) symbol.
Indicate for each of the following aspect which version you preferred: #01. Regarding the usability (own avatar); #02. Regarding the aesthetics (own avatar); #03. Regarding the aesthetics (other avatar); #04. Regarding the multi-player interactions; #05. Regarding own avatar; #06. Regarding the other avatar; #07. Regarding the overall user experience

are considered. For what it concerns the own representation, no clear winners emerged in terms of embodiment or preference, although the FB was perceived significantly better than the VK in terms of appearance, response to external stimuli, and immersion. Regarding the other user's avatar, the FB appeared to improve various social presence aspects, such as mutual awareness, mutual attention and mutual understanding. Furthermore, it was also preferred over the VK for the other user's representation (in general, but also for the aesthetics and multi-player interactions).

These findings suggest that multi-user training simulations would greatly benefit of the employment of a FB approach to represent the avatar of the other user inside the shared experience, as long as the outcome of the combined use of IK and animations produces a sufficiently believable and realistic result. The situation is different for what it concerns the own representation of the user, as the investigation produced mixed results. Comments provided by some participants at the end of the experience offer possible interpretations. In particular, some participants reported of being distracted by the FB avatar whenever the estimated pose of the body differed too much from the real one. In those cases, they would have preferred not to see the avatar at all, like with the VK. For this reason some users may still perceive VK as better than FB in these particular situations.

Future developments will be devoted to widen the investigation, considering other avatar representation techniques, as well as including the other technologies (e.g., motion capture suits and leg sensors) and locomotion modalities (e.g., locomotion treadmills, walk-in-place, etc.) already supported by the considered training tool. Moreover, more complex situations could be developed, for example requiring the users to fulfill more collaboration-oriented tasks (e.g., interaction with objects that have to be handled by two users at the same time, tasks which can be only completed by working together, presence of more than two users in the simulation, etc.). To support the above scenarios, the set of animations used by the FB implementation may need to be extended, e.g., by introducing also prone crawling, jumping, climbing, and any other movement which may have to be displayed in a collaborative emergency scenario. Furthermore, machine learning-based avatar representation techniques may be added to the comparison, also estimating their impact in terms of computational load with respect to the IK algorithms. Finally, the FB implementation may be integrated with facial and eye tracking capabilities, to further improve communication and expressiveness of the other user's avatar, and to increase the perceived realism of the simulation.

Acknowledgements

This work has been carried out in the frame of the VR@POLITO initiative.

References

1. Andrade, M., Souto Maior, C., Silva, E., Moura, M., Lins, I.: Serious games & human reliability. The use of game-engine-based simulator data for studies of evacuation under toxic cloud scenario. In: Proc. of Probabilistic Safety Assessment and Management (PSAM 14). pp. 1–12 (09 2018)
2. Bailenson, J.N., Yee, N., Merget, D., Schroeder, R.: The effect of behavioral realism and form realism of real-time avatar faces on verbal disclosure, nonverbal disclosure, emotion recognition, and copresence in dyadic interaction. Presence **15**(4), 359–372 (2006). https://doi.org/10.1162/pres.15.4.359
3. Benrachou, D.E., Masmoudi, M., Djekoune, O., Zenati, N., Ousmer, M.: Avatar-facilitated therapy and virtual reality: Next-generation of functional rehabilitation methods. In: 2020 1st International Conference on Communications, Control Systems and Signal Processing (CCSSP). pp. 298–304 (2020). https://doi.org/10.1109/CCSSP49278.2020.9151528
4. Biocca, F., Harms, C., L. Gregg, J.: The networked minds measure of social presence : Pilot test of the factor structure and concurrent validity. In: International Workshop on Presence, Philadelphia (2001)
5. Calandra, D., Billi, M., Lamberti, F., Sanna, A., Borchiellini, R.: Arm Swinging vs Treadmill: A Comparison Between Two Techniques for Locomotion in Virtual Reality. In: Diamanti, O., Vaxman, A. (eds.) EG 2018 - Short Papers. pp. 53–56. The Eurographics Association (2018). https://doi.org/10.2312/egs.20181043

6. Calandra, D., Lamberti, F., Migliorini, M.: On the usability of consumer locomotion techniques in serious games: Comparing arm swinging, treadmills and walk-in-place. In: Proc. of 2019 IEEE 9th International Conference on Consumer Electronics (ICCE-Berlin). pp. 348–352 (2019). https://doi.org/10.1109/ICCE-Berlin47944.2019.8966165
7. Calandra, D., Pratticò, F.G., Migliorini, M., Verda, V., Lamberti, F.: A multi-role, multi-user, multi-technology virtual reality-based road tunnel fire simulator for training purposes. In: Proc. of 16th International Conference on Computer Graphics Theory and Applications (GRAPP 2021). pp. 96–105 (2021). https://doi.org/10.5220/0010319400960105
8. Cannavò, A., Calandra, D., Pratticò, F.G., Gatteschi, V., Lamberti, F.: An evaluation testbed for locomotion in virtual reality. IEEE Transactions on Visualization and Computer Graphics **27**(3), 1871–1889 (2021). https://doi.org/10.1109/TVCG.2020.3032440
9. Caserman, P., Achenbach, P., Göbel, S.: Analysis of inverse kinematics solutions for full-body reconstruction in virtual reality. In: 2019 IEEE 7th International Conference on Serious Games and Applications for Health (SeGAH). pp. 1–8 (2019). https://doi.org/10.1109/SeGAH.2019.8882429
10. Corelli, F., Battegazzorre, E., Strada, F., Bottino, A., Cimellaro, G.P.: Assessing the usability of different virtual reality systems for firefighter training. In: Proc. of 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (HUCAPP 2020). pp. 146–153 (2020). https://doi.org/10.5220/0008962401460153
11. Engelbrecht, H., Lindeman, R.W., Hoermann, S.: A SWOT analysis of the field of virtual reality for firefighter training. Frontiers in Robotics and AI **6**, 101 (2019). https://doi.org/10.3389/frobt.2019.00101
12. Fathima S J, S.A., Aroma, R.J.: Simulation of fire safety training environment using immersive virtual reality. International Journal of Recent Technology and Engineering (IJRTE) **7**(4S), 347–350 (01 2019)
13. Feng, Z., González, V.A., Amor, R., Lovreglio, R., Cabrera-Guerrero, G.: Immersive virtual reality serious games for evacuation training and research: A systematic literature review. Computers & Education **127**, 252–266 (2018). https://doi.org/10.1016/j.compedu.2018.09.002
14. Gonzalez-Franco, M., Peck, T.C.: Avatar embodiment. towards a standardized questionnaire. Frontiers in Robotics and AI **5** (2018). https://doi.org/10.3389/frobt.2018.00074
15. Gu, L., Yin, L., Li, J., Wu, D.: A real-time full-body motion capture and reconstruction system for vr basic set. In: 2021 IEEE 5th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC). vol. 5, pp. 2087–2091 (2021). https://doi.org/10.1109/IAEAC50856.2021.9390617
16. Kalawsky, R.S.: VRUSE – A computerised diagnostic tool: For usability evaluation of virtual/synthetic environment systems. Applied Ergonomics **30**(1), 11–25 (1999). https://doi.org/10.1016/S0003-6870(98)00047-7
17. Kasapakis, V., Dzardanova, E.: Using high fidelity avatars to enhance learning experience in virtual learning environments. In: 2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW). pp. 645–646 (2021). https://doi.org/10.1109/VRW52623.2021.00205
18. Kinateder, M., Ronchi, E., Nilsson, D., Kobes, M., Müller, M., Pauli, P., Mühlberger, A.: Virtual reality for fire evacuation research. In: Proc. of Federated Conference on Computer Science and Information Systems. pp. 313–321 (01 2014). https://doi.org/10.13140/2.1.3380.9284

19. Kokkinara, E., Slater, M.: Measuring the effects through time of the influence of visuomotor and visuotactile synchronous stimulation on a virtual body ownership illusion. Perception **43**(1), 43–58 (2014). https://doi.org/10.1068/p7545, pMID: 24689131
20. Lamberti, F., De Lorenzis, F., Pratticò, F.G., Migliorini, M.: 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 (2021). https://doi.org/10.1109/COMPSAC51774.2021.00030
21. Louka, M.N., Balducelli, C.: Virtual reality tools for emergency operation support and training. In: Proc. International Conference on Emergency Management Towards Co-operation and Global Harmonization (TIEMS 2001). pp. 1–10 (06 2001)
22. Lovreglio, R.: Virtual and augmented reality for human behaviour in disasters: A review. In: Proc. of Fire and Evacuation Modeling Technical Conference (FEMTC 2020). pp. 1–14 (08 2020)
23. Lu, X., Yang, Z., Xu, Z., Xiong, C.: Scenario simulation of indoor post-earthquake fire rescue based on building information model and virtual reality. Advances in Engineering Software **143**, 102792 (2020). https://doi.org/10.1016/j.advengsoft.2020.102792
24. Lugrin, J.L., Ertl, M., Krop, P., Klüpfel, R., Stierstorfer, S., Weisz, B., Rück, M., Schmitt, J., Schmidt, N., Latoschik, M.E.: Any "body" there? avatar visibility effects in a virtual reality game. In: 2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR). pp. 17–24 (2018). https://doi.org/10.1109/VR.2018.8446229
25. Molina, E., Jerez, A.R., Gómez, N.P.: Avatars rendering and its effect on perceived realism in virtual reality. In: 2020 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR). pp. 222–225 (2020). https://doi.org/10.1109/AIVR50618.2020.00046
26. Morélot, S., Garrigou, A., Dedieu, J., N'Kaoua, B.: Virtual reality for fire safety training: Influence of immersion and sense of presence on conceptual and procedural acquisition. Computers & Education **166**, 104145 (2021). https://doi.org/10.1016/j.compedu.2021.104145
27. Parger, M., Mueller, J.H., Schmalstieg, D., Steinberger, M.: Human upper-body inverse kinematics for increased embodiment in consumer-grade virtual reality. In: Proc. of the 24th ACM Symposium on Virtual Reality Software and Technology. VRST '18, Association for Computing Machinery, New York, NY, USA (2018). https://doi.org/10.1145/3281505.3281529
28. Pedram, S., Palmisano, S., Skarbez, R., Perez, P., Farrelly, M.: Investigating the process of mine rescuers' safety training with immersive virtual reality: A structural equation modelling approach. Computers & Education **153**, 103891 (2020). https://doi.org/10.1016/j.compedu.2020.103891
29. Pratticò, F.G., De Lorenzis, F., Calandra, D., Cannavò, A., Lamberti, F.: Exploring simulation-based virtual reality as a mock-up tool to support the design of first responders training. Applied Sciences **11**(16), 1–13 (2021). https://doi.org/10.3390/app11167527
30. Roth, D., Lugrin, J.L., Galakhov, D., Hofmann, A., Bente, G., Latoschik, M.E., Fuhrmann, A.: Avatar realism and social interaction quality in virtual reality. In: 2016 IEEE Virtual Reality (VR). pp. 277–278 (2016). https://doi.org/10.1109/VR.2016.7504761
31. Schäfer, A., Reis, G., Stricker, D.: A survey on synchronous augmented, virtual and mixed reality remote collaboration systems (02 2021)

32. Steed, A., Frlston, S., Lopez, M.M., Drummond, J., Pan, Y., Swapp, D.: An ‘in the wild’ experiment on presence and embodiment using consumer virtual reality equipment. IEEE Transactions on Visualization and Computer Graphics **22**(4), 1406–1414 (2016). https://doi.org/10.1109/TVCG.2016.2518135
33. Çakiroğlu, Ü., Gökoğlu, S.: Development of fire safety behavioral skills via virtual reality. Computers & Education **133**, 56–68 (2019). https://doi.org/10.1016/j.compedu.2019.01.014