A virtual reality methodology for cardiopulmonary resuscitation training with and without a physical mannequin

Fabio Buttussi,[1,2] Luca Chittaro,[1] and Francesca Valent[2]

[1]Human-Computer Interaction Lab, Department of Mathematics, Computer Science, and Physics, University of Udine, Udine, Italy

[2]SOC Istituto di Igiene ed Epidemiologia Clinica, Azienda Sanitaria Universitaria Friuli Centrale, Udine, Italy

Corresponding Author: Fabio Buttussi, PhD, Human-Computer Interaction Lab, Department of Mathematics, Computer Science, and Physics, University of Udine, Via delle Scienze 206, 33100 Udine, Italy (fabio.buttussi@uniud.it)

Keywords: virtual reality, training, cardiopulmonary resuscitation, medical education, mannequin, experimental evaluation

ABSTRACT

Background: Cardiopulmonary resuscitation (CPR) is an emergency procedure that can increase survival after a cardiac arrest. Performing CPR effectively requires both procedural knowledge and manual skills. Traditional CPR training methodology includes lessons led by instructors and supervised practice on mannequins, thus requiring considerable resources.

Objective: This paper proposes a new methodology for low-cost CPR training based on virtual reality (VR) with and without the addition of a physical mannequin. Moreover, it describes an experimental evaluation of the methodology that assessed gain in manual skills during training, transfer of procedural knowledge and manual skills in a final assessment, and changes in self-efficacy with three measurements over time (pre-training, post-training, and post-assessment).

Methods: We implemented a VR application that supports the proposed methodology, and can thus be used with or without a mannequin. The experimental evaluation involved 30 participants who tried CPR in VR twice, performing two repetitions of 30 chest compressions per trial. Half participants tried the VR application with the mannequin and half without it. Final assessment required all participants to perform CPR on the mannequin without the assistance of VR. To assess self-efficacy, participants filled in a questionnaire at the three times of measurement.

Results: Mixed-design ANOVAs showed effects of repetition, effects of group, or interaction between the two variables on manual skills assessed during training. In the final assessment, participants in both groups correctly remembered most of the steps of the procedure. ANOVAs revealed differences between the two groups only in pressure-related skills (better with mannequin) and in the number of wrong steps added to the procedure (better without mannequin). Mixed-design ANOVA showed a self-efficacy increase in both groups after training, which was maintained after final assessment.

Conclusions: The proposed VR methodology for CPR training has a positive effect on procedural knowledge, manual skills, and self-efficacy, with as well as without the physical mannequin. Trials on a mannequin are required to understand the correct pressure for chest compression. This supports the adoption of the proposed VR methodology to reduce instructor and mannequin time required to teach CPR to trainees.

1. Introduction

The 2020 update of the Heart Disease and Stroke Statistics from the American Heart Association [1] reported about 17.8 million deaths attributed to cardiovascular diseases globally in 2017. Among them, sudden cardiac arrest was mentioned as responsible of nearly 380,000 deaths only in the United States [1]. These numbers motivate Advanced Life Support (ALS) training of emergency medical responders as well as Basic Life Support (BLS) training of health professionals and laypeople to handle both in-hospital and out-of-hospital events, such as those happening at home, workplace or public places. Among BLS procedures, cardiopulmonary resuscitation (CPR) specifically addresses cardiac arrest and could increase survival [1]. The CPR procedure consists of a sequence of steps concerning examination (e.g., checking if airways are open) and treatment (e.g., performing chest compressions). Performing CPR effectively requires both procedural knowledge (e.g., remembering the correct sequence of steps) and manual skills (e.g., performing a chest compression with the correct pressure). Traditional CPR training methodology includes lessons led by instructors who explain the steps of the procedure, and supervised practice on mannequins to master manual skills, thus requiring considerable resources. Therefore, authorities in the field are calling for innovative, cost-effective CPR training methodologies [2][4].

In the last decade, an increasing number of technology-based solutions to teach CPR [5][21], BLS [22][24], and ALS [25][27] have been proposed. These include virtual reality (VR) applications for PCs [5][7][25][27] and mobile devices [9]-[12],[22]. In the last years, the availability of consumer VR headsets with wide field of view (FOV) and the possibility to track head position and rotation with six degrees of freedom (6-DOF) facilitated the development of new applications using VR [16][24]. As reported in detail in Table 1, most of these VR applications were evaluated to assess their effects on knowledge [7][10][12][17][20],[22][25][27], skills [6][9][12],[16][19][20][23], or self-efficacy [5],[23]. However, the evaluation methodologies followed by previous proposals did not take into account knowledge, skills, and self-efficacy all together.

Reference	Topic	Platform	Target	Knowledge assessment	Skill assessment	Self-efficacy assessment

Creutzfeldt et al. [5]	CPR	PC + monitor	High-school students in a group setting	No	No	Before and after training + 6-months follow-up
Wattanasoontorn et al. [6]	CPR	PC + monitor + camera-based tracker	Laypeople in an individual setting	No	During training	No
Latif et al. [7]	CPR	PC + monitor	Laypeople in an individual setting	Before and after training	No	No
Morrison-Smith et al. [9]	CPR	Tablet + camera-based tracker	Laypeople in an individual setting	No	Before and during training	No
Nas et al. [10],[11]	CPR	Smartphone + VR goggles	Laypeople in an individual setting	After training	After training	No
Leary et al. [12]	CPR	Smartphone + VR goggles	Laypeople in an individual setting	After training	After training	No
Varun Durai et al. [16]	CPR	VR headset	Laypeople in an individual setting	No	After training	No
Bucher et al. [17]	CPR	VR headset	Laypeople in an individual setting	Before and after training	No	No
Vaughan et al. [18]	CPR	VR headset	School children in an	No	No	No

			individual setting			
Semeraro et al. [19]	CPR	VR headset	Medicine students in an individual setting	No	During training	No
Gent et al. [20]	CPR	VR headset	Middle / high-school students in an individual setting	After training	Before and after training	No
Almousa et al. [21]	CPR	VR headset	Laypeople in an individual setting	No	No	No
Aksoy [22]	BLS	Tablet and VR headset	Paramedicine students in an individual setting	Before and after training	No	No
Bench et al. [23]	BLS	VR headset	Laypeople in an individual setting	No	After training	Before and after training
Semeraro et al. [24]	BLS	VR headset	Laypeople and health professionals in an individual setting	No	No	No
Buttussi et al. [25]	ALS	PC + monitor	Health professionals in an	Before and after training + 3-months follow-up	No	No

			individual setting			
Vankipuram et al. [26]	ALS	PC + monitor + haptic joystick	Health professionals in a group setting	No	No	No
Khanal et al. [27]	ALS	PC + monitor + haptic joystick	Health professionals in a group setting	Before, during, and after training	No	No

Table 1. VR applications to teach CPR, BLS, or ALS.

This paper aims at advancing knowledge in different directions. First, we introduce a new methodology for low-cost CPR training based on VR with and without the addition of a physical mannequin, and develop a VR application to support it. Second, we propose an evaluation methodology for assessing possible effects of CPR training methodologies on knowledge, skills, and self-efficacy. Third, we apply the proposed evaluation methodology to assess if our VR methodology for CPR training, with or without the addition of a physical mannequin, has a positive effect on gain in manual skills during training. Moreover, we study the effects of the proposed VR methodology on the transfer of CPR procedural knowledge and manual skills in a final assessment, where participants perform CPR on a mannequin without the assistance of VR. Finally, we assess if participants' self-efficacy changes immediately after training and after final assessment. To the best of our knowledge, this paper is the first to evaluate the effects of a VR methodology for CPR training on knowledge, skills, and self-efficacy all together, and both with or without the addition of a physical mannequin.

2. The proposed VR methodology for CPR training

2.1. Foundations

The proposed VR methodology for CPR training lays its foundations on training principles and on evidence from VR research. First, the methodology builds upon constructivist theories, whose application to VR has been discussed in [28]. Following constructivist theories, individuals learn through a direct experience of the world [28], so the methodology proposes highly interactive experiences where trainees physically perform the steps of the CPR procedure in VR. In addition, situated learning theory suggests that it is easier to learn concepts in the same context where they will be applied [28], so the interactive experiences proposed by the methodology will also be set in cardiac arrest scenarios that reproduce real-world settings and events. Since we aim at reducing resources needed for CPR training, we set as a requirement for our methodology that the proposed training experiences should be deliverable on consumer VR devices. Moreover, in those cases where a mannequin is available, the methodology must support the integration of VR with the physical mannequin using low-cost devices. Finally, our methodology applies the training strategy of fading, whose application to VR has been described in [29]. Fading strategy suggests giving trainees major clues and guides at the start of training, and then gradually fading out the amount of clues and guides until the trainee is required to perform the task without support [29].

Considering VR-related aspects, we observed that CPR, BLS, and ALS proposals in the literature differed in terms of display fidelity, i.e. the realism of the output devices [30], and interaction fidelity, i.e., the realism of the input devices [30]. Some proposals displayed VR on a PC monitor [5-7,25-27] or on the screen of a mobile device [9][22], other used mobile devices inside of VR goggles offering 3-DOF head tracking [10][12], and others relied on the higher display and interaction fidelity of VR headsets with wide FOV and 6-DOF head tracking [16-24]. However, only one study [22] compared the effects of two versions of a BLS VR application with different fidelity showing higher knowledge gain with a VR headset than a tablet. Research on the effects of higher fidelity in other domains did not always show higher knowledge gain (see the related work section in [31]) than traditional displays, but VR headsets with wide FOV and 6-DOF head tracking were associated with higher presence in VR and higher engagement than PC monitors [31], leading us to adopt a VR headset with wide FOV and 6-DOF head tracking in our methodology.

To improve the realism of input, some proposals used a robotic 3D controller [26][27], camera-based motion tracking [6][9][13][14],[18], or a custom chest model with a spring [16]. Recent proposals took advantage of consumer VR 6-DOF trackers that can be used to track position and rotation of trainees' hands [19],[21]. In particular, Semeraro et al. [19] found equivalent results in measuring chest compression data with sensors integrated in a training mannequin or with consumer VR 6-DOF trackers. Therefore, we decided to adopt consumer VR 6-DOF trackers in our methodology.

2.2. Methodology

Since trainees may not be familiar with VR, the proposed methodology for CPR training suggests starting training with a tutorial on using VR. In the tutorial, trainees can be immersed in the VR environment that will be used for training, and familiarize with VR by looking around, performing movements, and observing their hands moving in VR. Then, the methodology proposes to teach CPR by making trainees perform the different steps of the procedure in VR with different grades of clues and guides as described in the following. The following are the essential steps of a CPR procedure:

1. Safety check. The rescuer must make sure the environment is safe.
2. Shoulder shaking. The rescuer must shake the shoulders of the person who appears to be in cardiac arrest, and verbally call him/her to check if he/she is unconscious.
3. Head rotation. The rescuer must rotate the person's head to open airways (Figure 1A);
4. Airways check. The rescuer must look inside the person's mouth to check that airways are open.
5. Breath check. The rescuer must look at the person's chest and listen to sound from his/her mouth to be sure he/she is not breathing.
6. Phone call. The rescuer must make a phone call to the first response emergency number (unless he/she is an EMS provider in service).
7. Chest compression. The rescuer must perform 30 chest compressions at 100-120 compressions per minute (cpm), with compression depth between 5 cm and 6 cm (Figure 1B);
8. Insufflation. The rescuer must perform two insufflations.
9. Breath check. The rescuer must check that the person has started to breathe again autonomously.

It is important to note that different CPR scenarios may require repeating steps 7 and 8 several times before the person starts to breathe again autonomously, so different VR applications that support our methodology can propose difference scenarios that require one or more repetitions. Different scenarios can also be set in different environments (e.g., at home, in a workplace, in a nursing home, in a public place), and with the possible availability of virtual rescuers that can help the trainee by performing some of the CPR steps.

Applying the training strategy of fading, the methodology makes trainees perform the CPR procedure in three modes: first in a guided VR mode, then in a semi-guided VR mode, and finally in an unguided mannequin-only mode. In the guided VR mode, every step of the procedure is verbally explained (e.g., using a synthetic voice) and written information about the step is provided (e.g., through a panel inside the simulation). In both VR modes, different clues suggest the position and orientation in which trainees have to place their head and hands (e.g., using semi-transparent cues as in Figure 1A), indicate relevant off-screen objects (e.g., using arrows), and show current and correct cpm and compression depth (e.g., using two gauges as in Figure 1B). In both VR modes, a voice provides advice (e.g., saying “press more deeply” or “faster”) when trainees have to increase or decrease cpm or compression depth as well as when they perform incomplete releases, i.e., they are releasing their hands, but start a new chest compression when depth is still higher than 1 cm. In the unguided mannequin-only mode, trainees perform CPR without the assistance of VR, but using sensors to record performance for further review.

The minimum equipment needed to support our methodology is a VR headset with a diagonal FOV of at least 100 degrees (supported by most current consumer headsets) and 6-DOF head tracking (Figure 2A), and two 6-DOF trackers that trainees wear (one for each wrist as in Figure 2B) to track their hands. Figure 2 (A and B) shows an example of hardware that meets the requirements set by the methodology. Since this equipment allows for tracking the position and rotation of trainees’ head and hands, the methodology uses it also to check if trainees actually perform the steps of the procedure.

To provide physical feedback and further increase fidelity, our methodology proposes the addition of a physical mannequin (Figure 2C), if available. When the mannequin is used, two additional 6-DOF trackers (Figure 2D) have to be placed on the mannequin to synchronize its position with that of the virtual person on which CPR is performed in VR.

Figure 1. A trainee in the physical room with the corresponding VR view in the top-left box during A) head rotation and B) chest compression.

The methodology recommends calculating the depth of compressions on the mannequin chest by combining data from two sensors on a low-cost (e.g., Arduino) board inside the mannequin (Figure 2E): an ultrasonic distance sensor and a temperature sensor, because measurements of ultrasonic distance sensors could vary with temperature.

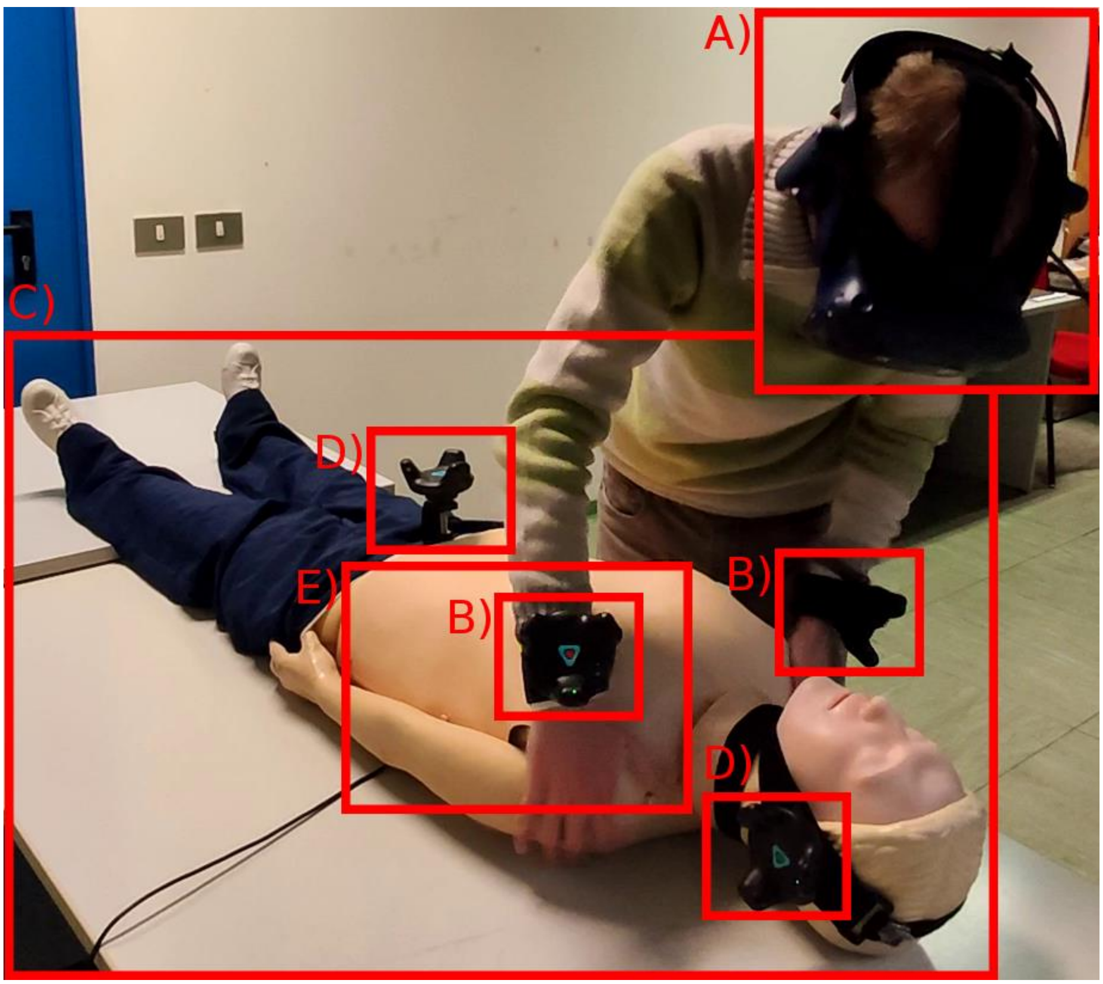


Figure 2. Equipment to support the proposed VR methodology for CPR training: A) VR headset with a diagonal FOV of 110 degrees and 6-DOF head tracking, B) two 6-DOF trackers that trainees wear, one for each wrist, C) physical mannequin, D) two 6-DOF trackers to synchronize mannequin position with position of the virtual person who is in cardiac arrest in VR, E) Arduino board (inside the mannequin) with ultrasonic distance sensor and temperature sensor.

2.3. VR application

We developed a VR application that supports the proposed VR methodology for CPR training. As suggested by the methodology, the VR application starts with a VR tutorial, which immerses trainees in a virtual office environment, where they can familiarize with VR as described in previous section for 2 minutes. When the CPR scenario starts, trainees remain in the same virtual office, but there is a person laid down on a desk in front of them, with his office colleague standing nearby, and a PC with electrical problems that has fallen on the floor. The response procedure consists in eleven steps, because in the considered scenario trainees have to perform two repetitions of the 30 chest compressions with two insufflations before the person starts to breathe again autonomously (the eighth and ninth step of the above described procedure have to be repeated). In the considered scenario, the safety check consists in turning off the power strip the PC is plugged into. Trainees have to perform the other steps of the procedure on the person laid down on the desk. In the considered scenario, an office colleague helps the trainee by performing the two insufflations of the eighth step of the procedure (as well as its repetition at the tenth step) and the trainee has to wait for completion before performing the next step.

Since the methodology makes trainees perform the procedure in VR twice (in guided and semi-guided VR mode), the trainee performs a total of four repetitions (two per mode) of 30 chest compressions each in VR. In the unguided mannequin-only mode, the application detects and logs chest compressions using sensors, but does not display VR on the headset, which is not worn by trainees in this mode.

As proposed by the methodology, the VR application can provide CPR training with or without the addition of a physical mannequin. The VR devices it runs with are the HTC Vive Pro headset with its two Lighthouse base stations, and two HTC trackers for the wrists. The mannequin is a Laerdal Resusci Anne, with two additional HTC trackers to synchronize its position with VR. The two sensors on the Arduino board inside the mannequin are an HC-SR04 ultrasonic distance sensor and a LM35 temperature sensor. The VR application was developed using Unity game engine, and built for Windows PCs.

3. Experimental evaluation

3.1. Evaluation methodology

Before describing the specific experimental evaluation we carried out, we introduce the evaluation methodology that we propose to follow for testing CPR training methodologies. The evaluation methodology adopts a between-groups design in which each participant tries one of the CPR training methodologies to be compared. Since the effectiveness of CPR depends on both procedural knowledge and manual skills, the evaluation methodology aims at assessing both of them in a single experimental evaluation. Moreover, since self-efficacy can be a predictor of actual performance [32], our evaluation methodology includes also this construct.

Regarding manual skills, the evaluation methodology takes inspiration from the experimental evaluations described in [33][34], which assessed if surgical skills improved among repeated trials with VR simulators, and if skills learnt with the simulators can be transferred to the real world.

Regarding procedural knowledge and self-efficacy, the evaluation methodology takes inspiration from the experimental evaluation described in [35], which asked participants to perform a safety procedure in the real world after they learned it in a virtual world, and assessed participants self-efficacy at three times: before training, after training, and after they performed the procedure in the real world.

Combining aspects from those experimental evaluations and adapting them to CPR training, the evaluation methodology we defined consists of the steps illustrated in Figure 3 and described in detail in the following paragraphs.

3.1.1. Introduction and consent.

Participants receive essential information about the experimental evaluation (i.e., they are informed that the scope concerns testing a CPR training methodology, that they can refrain from continuing the experimental evaluation at any time and for any reason, that they will have to fill in questionnaires, that questionnaires will be analyzed in an anonymized form) and give their written consent for participation.

3.1.2. Demographic questionnaire and assignment to groups.

Participants are assigned to groups corresponding to different CPR training methodologies or variants of a training methodology. The evaluation methodology recommends minimizing differences between groups in relevant demographic characteristics. Groups must always be balanced with respect to participants' age, gender, and familiarity with CPR to avoid biased results that can be caused by differences in these characteristics between groups. Additional demographic characteristics may be relevant depending on the CPR training methodologies to be evaluated. For example, to evaluate CPR training methodologies based on VR, it is important to consider familiarity of participants with VR.

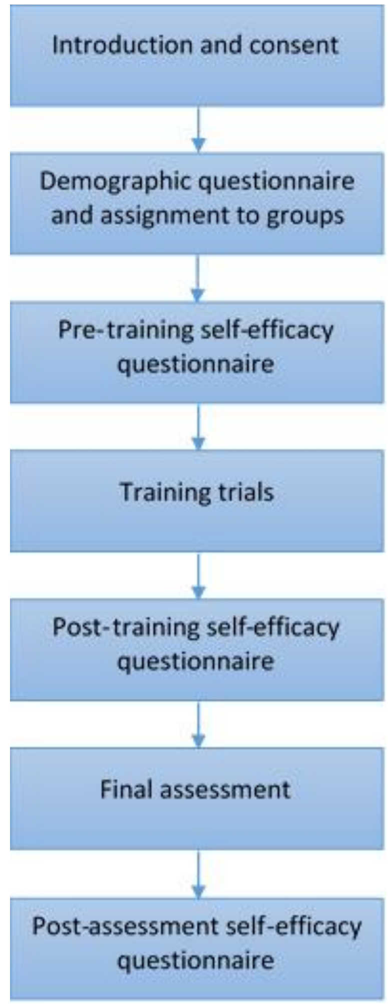


Figure 3. Steps of the evaluation methodology.

3.1.3. Pre-training self-efficacy questionnaire.

Before training, participants fill in a self-efficacy questionnaire to assess their baseline score. We designed a self-efficacy questionnaire for CPR by adapting items from well-known self-efficacy questionnaires and previous work in training for emergencies [31],[35],[36]. The self-efficacy questionnaire contains five items: i) I feel confident of my ability to perform CPR, ii) I would be able to check if a person can breathe autonomously, iii) I can practice a cardiac massage correctly, iv) I would be able to understand when a person has regained vital functions, and v) I can practice a cardiac massage without losing time. The questionnaire asks participants to rate their level of agreement with the items using a 7-value Likert scale (1 to 7, with 7 indicating highest agreement).

3.1.4. Training trials.

Participants train in performing CPR following the CPR training methodology or methodology variant they were assigned to. The experimenter informs participants that they cannot talk with him/her during training, and invites participants to ask him/her possible questions before starting training. To assess if repeating training leads to a gain in manual skills, the evaluation methodology proposes to make participants repeat training trials as in the experimental evaluations described in [33],[34]. For example, the number of trials was set to four in [33] and to six in [34]. Since different CPR scenarios can require repeating a particular step (e.g., chest compression) two or more times within the same trial of the procedure, the number of trials can be adjusted accordingly. For example, the VR application we built asks trainees to perform two trials with two repetitions of 30 compressions per trial. As a result, manual skills related to chest compression can be assessed four times.

The evaluation methodology proposes to assess the following measures for each repetition:

i) Compression depth. It is calculated as the average depth of compressions performed during the repetition.

ii) Cpm. It is calculated as the number of compressions performed by participants divided by the time they spent to perform them.

iii) Error in cpm. It is set to: 0, if cpm was in the correct range (100–120); 100 minus cpm, if cpm was below 100; cpm minus 120, if cpm was above 120.

iv) Correct compressions. It is calculated as the percentage of compressions with correct pressure, complete release, and correct hand position.

v) Too deep compressions. It is calculated as the percentage of compressions with depth higher than 6 cm.

vi) Too shallow compressions. It is calculated as the percentage of compressions with depth lower than 5 cm.

vii) Compressions with incomplete release. It is calculated as the percentage of compressions in which the depth at the time of release was higher than 1 cm.

viii) Error in the number of chest compressions. It is calculated as the absolute value of 30 minus the number of chest compressions performed by participants.

Comparing the values of these measures in the different repetitions, the proposed evaluation methodology allows assessing gain in manual skills during training.

3.1.5. Post-training self-efficacy questionnaire.

Participants fill in the self-efficacy questionnaire for a second time, immediately after training. In this way, the evaluation methodology allows to assess a possible immediate increase in self-efficacy due to training.

3.1.6. Final assessment.

To assess possible transfer of manual skills and procedural knowledge, participants perform CPR on a mannequin without any training support, and describe each step before performing it. More precisely, to assess possible transfer of manual skills, the evaluation methodology proposes to assess the same measures considered during training and described in Section 3.1.4. To assess possible transfer of procedural knowledge, the evaluation methodology recommends to record video of participants while they described and performed the procedure, so that the experimenter can later assign one of the following codes to each expected correct step of the procedure:

i) “Correct”, if participants showed they understood that step correctly and completely;
ii) “Partial”, if participants included that step, but it was incomplete or partially wrong, e.g., participants remembered the step concerning insufflations, but they did not say how many times insufflations should be performed in that step;
iii) “Omitted”, if participants did not include that step.

By summing up all the steps with the same code, the evaluation methodology allows calculating three measures respectively called correct steps, partial steps, and omitted steps. In addition, the evaluation methodology considers the following measures:

i) Misplaced steps. For each correct and partial step, the step is considered misplaced if it was included before (respectively after) at least one step that should have preceded (respectively followed) the considered step. The measure is the number of misplaced steps in the procedure.
ii) Wrong steps. It is the number of steps that should not have appeared in the procedure at all, e.g., the participant says that another rescuer has to check if the person starts to breathe again.

3.1.7. Post-assessment self-efficacy questionnaire.

Participants fill in the self-efficacy questionnaire for a third time, after final assessment. In this way, the evaluation methodology allows to assess a possible increase or decrease in self-efficacy.

3.2. Experimental design and hypotheses

We applied the proposed evaluation methodology to evaluate our VR methodology for CPR training with and without the addition of the physical mannequin. Following the above described between-groups design, half participants trained using VR with the mannequin (VRMA group) and the other half trained using VR only (VRON group).

We formulated the following hypotheses for the experimental evaluation. Considering gain in manual skills during training, we hypothesized an improvement between the first and the last (the fourth) repetition as suggested by the results of [33],[34], which concerned manual skills in surgery. Since only the VRMA group receives physical feedback due to the presence of the mannequin chest during training, we hypothesized that improvement in the two groups might differ on measures concerning pressure. By considering the results of previous experimental evaluations that assessed CPR and ALS procedural knowledge after VR training [20][27] as well as transfer of procedural knowledge in other domains [35][37][39], we hypothesized a transfer of procedural knowledge about CPR in final assessment. The experimental evaluation was explorative with respect to possible differences due to the presence or absence of the physical mannequin on transfer of procedural knowledge. The exploration was meant to assess the possible merits of the flexibility of the proposed VR methodology, which supports CPR training both with and without a mannequin. Considering transfer of manual skills, we hypothesized a successful transfer at least for the VRMA group, because in [20] there was an improvement in cpm and compression depth after training in VR using a physical mannequin. We also hypothesized that measures concerning pressure might be better in the VRMA group because of the physical feedback due to the presence of the mannequin chest during training. Finally, we hypothesized that self-efficacy would increase after training as shown in previous experimental evaluations concerning CPR [5] and procedural training in another emergency-related domain [31],[35]. The experimental evaluation was explorative with respect to possible differences due to the presence or absence of the physical mannequin on self-efficacy.

3.3. Materials

The experimental evaluation employed the VR application with the VR headset, the sensors, and the mannequin described in Section 2.3. The application was run on a Windows PC equipped with a 3.60 GHz Intel i7-3820 processor, 16 GB RAM, and an NVidia GTX 1070 graphic card. The PC was also equipped with a 1280x720 webcam that was used to record trainees' final assessment.

3.4. Participants

The experimental evaluation involved 30 participants (undergraduate computer science students). The demographic questionnaire asked them their gender, their age, if they had ever performed CPR (e.g., in a first response course), and if so, how many months ago they performed CPR. Moreover, since familiarity with VR related technologies can play a role in the evaluation of a VR methodology, the demographic questionnaire also asked participants about how many hours they had previously used VR headsets, and how many hours a week they spent on videogames.

There were 28 male and 2 female participants. Age ranged from 21 to 25 (M=22.00, SD=1.11). Eleven participants had performed CPR, and the remaining 19 had not. Among those who performed CPR, the number of months since they did it ranged from 1 to 40 (M=26.00, SD=14.87). Hours of previous use of VR headsets ranged from 0 to 30 (M=1.62, SD=5.48), and half participants had never used a VR headset before. Weekly hours spent on videogames ranged from 0 to 20 (M=7.12, SD=6.39), and four participants never played videogames.

Participants were assigned to the two groups in such a way that i) both groups had 1 female and 14 male participants; ii) one group (VRMA) had 6 participants who had performed CPR and the other had 5; iii) the groups were similar in terms of age, number of months since they performed CPR (among participants who did it), hours of VR headset previous use, and weekly hours spent on videogames. Each of these variables was submitted to a one-way ANOVA that confirmed the lack of significant differences between the groups.

3.5. Procedure

The procedure of the experimental evaluation consisted in the application of the steps of the proposed evaluation methodology illustrated in Figure 3 and described in Section 3.1.

Before the evaluation step concerning training trials, the experimenter helped participants to wear the VR headset and the 6-DOF trackers on the wrists. After the 2-minute VR tutorial, the experimenter explained participants that they were about to try the CPR procedure twice (one in the guided VR mode, and one in the semi-guided VR mode). When participants completed the two trials, the experimenter helped them in removing the VR headset.

During final assessment, participants kept wearing the 6-DOF trackers on the wrists. We used a webcam (focused on the mannequin) to record audio and video.

The experimental evaluation was carried out following the ethical code of our institution.

3.6. Measures

To assess gain in manual skills during training, for each of the four repetitions of 30 chest compressions (two repetitions for each of the two trials), we assessed all the measures suggested by the evaluation methodology (Section 3.1.4), except the error in the number of chest compressions, because the correct count was monitored and guided by the VR application so participants could not make errors in this measure.

The measures concerning manual skills were assessed using the sensors placed inside the mannequin for the VRMA group, and using 6-DOF trackers on the wrists for the VRON group. As mentioned in Section 2.1, suitability of 6-DOF trackers to measure compression depth and cpm was evaluated by Semeraro et al., who found equivalent results using sensors in a training mannequin or 6-DOF trackers [19].

During final assessment, as prescribed by the evaluation methodology, we assessed all the measures concerning manual skills described in Section 3.1.4 (including error in the number of chest compressions counted through the video recording), as well as the measures concerning procedural knowledge described in Section 3.1.6.

When the depth of one or more compressions was lower than the sensor 0.3 cm detection threshold, there could be a discrepancy between the number of compressions detected from the video and the number detected by the sensor. In these cases, we adjusted compression depth to consider undetected compressions as if they were performed at threshold value (0.3 cm) to reduce the error in the measure.

Since there were eleven steps in the CPR procedure for the considered scenario, correct steps, partial steps, omitted steps, and misplaced steps can range from 0 to 11.

To assess self-efficacy, we administered the self-efficacy questionnaire described in Section 3.1.3. Answers to the five questionnaire items were averaged to form a reliable scale (Cronbach's alpha pre-training = 0.89, post-training = 0.88, post-assessment = 0.91), ranging from 1 to 7.

3.7. Statistical analysis

To analyze the measures concerning gain in manual skills during training, we used a 2×4 mixed design ANOVA, in which group served as the between-subjects variable, and repetition served as the within-subjects variable. To compare the two groups on measures concerning transfer of manual skills in the final assessment as well as on measures concerning transfer of procedural knowledge, we used a one-way ANOVA. To analyze self-efficacy scores, we used a 2x3 mixed design ANOVA, in which group served as the between-subjects variable, and time of measurement (pre-training, post-training, and post-assessment) served as the within-subjects variable.

For all mixed design ANOVAs, if Mauchly's test indicated that the assumption of sphericity was violated, degrees of freedom were corrected using Greenhouse-Geisser. In case of statistically significant main effects of independent variables with more than two levels (repetition and time of measurement), we proceeded with pairwise comparisons using Bonferroni. In case of a statistically significant interaction between group and the within-subjects variable, we proceeded with the analysis of simple main effects with Bonferroni correction by testing the effects of the within-subjects variable separately for each group, and the effects of group separately for each level of the within-subjects variable. We report all effect sizes as partial eta squared (η_p^2).

4. Results

4.1. Gain in manual skills during training

Figure 4 shows the evolution over repetitions of all measures concerning manual skills during training in VR with and without the mannequin. To help understand main effects, Table 2 reports mean and standard deviation of all these measures for each group and repetition as well as cumulative values for all groups and all repetitions. Mixed-design ANOVAs showed a main effect of group on compression depth, $F(1,28)=10.58$, $p<0.005$, $\eta_p^2=0.27$, too deep compressions, $F(1,28)=15.56$, $p<0.001$, $\eta_p^2=0.36$, and too shallow compressions, $F(1,28)=13.00$, $p<0.005$, $\eta_p^2=0.32$. Compression depth (Figure 4A) and too deep compressions (Figure 4E) were higher with VRON than VRMA, while too shallow compressions (Figure 4F) were higher with VRMA than VRON. For the other measures, the main effect of group was not statistically significant ($p>0.05$ for all).

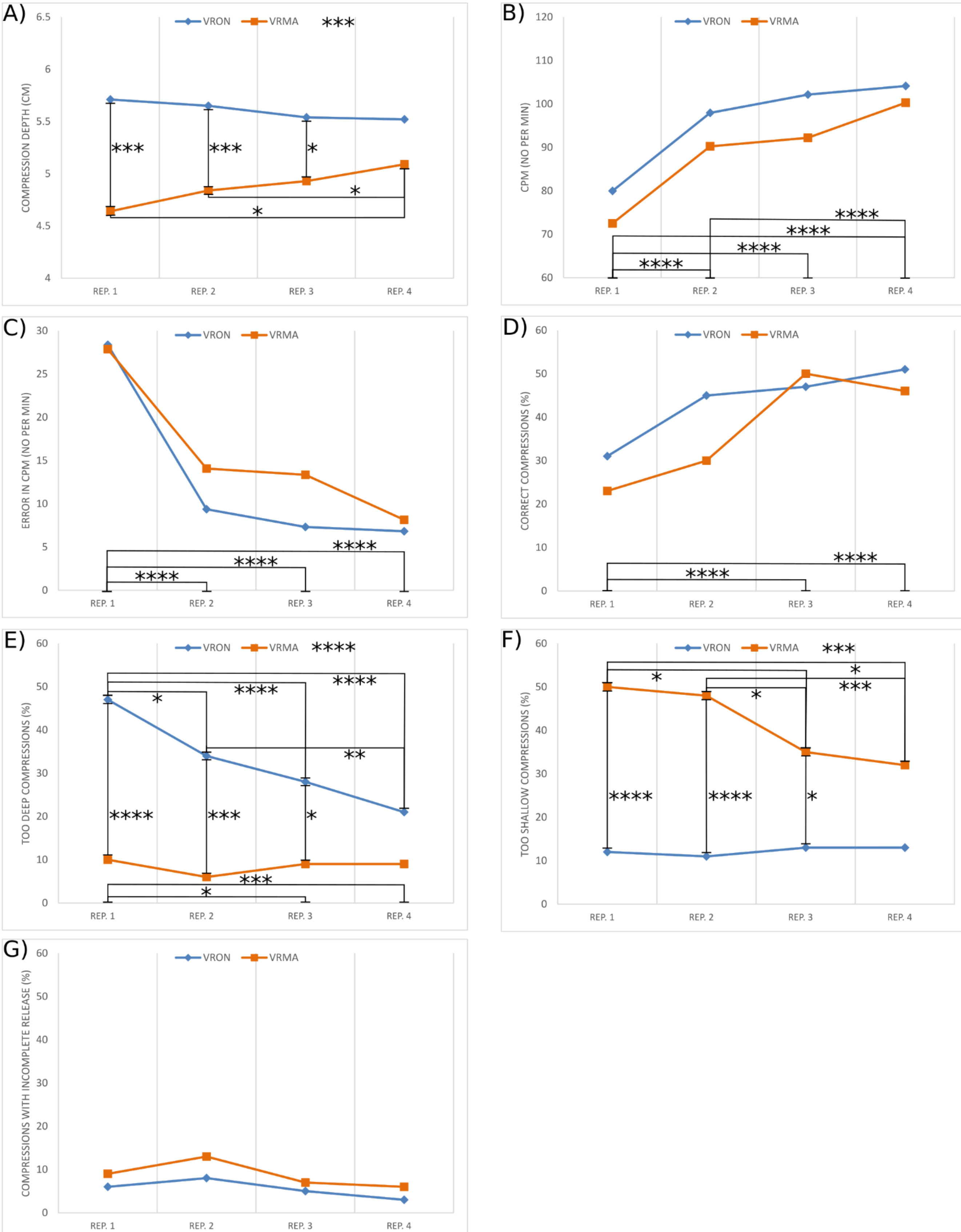


Figure 4. All measures concerning manual skills during training. The *, **, ***, **** signs indicate

statistically significant differences with p-values respectively <0.05, <0.01, <0.005, <0.001. A sign at the top of the chart (at the right of the legend) indicates a main effect of group. Signs at the bottom of each chart indicate significant pairwise comparisons between different repetitions after main effect of repetition. The other signs indicate significant pairwise comparisons between groups (respectively repetitions) for a repetition (respectively group) after a significant interaction was detected.

Compression depth (cm)										
	Repetition 1		Repetition 2		Repetition 3		Repetition 4		All repetitions	
	M	SD	**M**	SD	**M**	SD	**M**	SD	**M**	SD
VRON	**5.71**	0.55	**5.65**	0.37	**5.54**	0.30	**5.52**	0.28	**5.60**	0.25
VRMA	**4.64**	1.01	**4.84**	0.79	**4.93**	0.97	**5.09**	0.81	**4.88**	0.83
VRON+VRMA	**5.18**	0.97	**5.24**	0.73	**5.24**	0.77	**5.30**	0.64	**5.24**	0.71
Cpm (No per min)										
	Repetition 1		Repetition 2		Repetition 3		Repetition 4		All repetitions	
	M	SD	**M**	SD	**M**	SD	**M**	SD	**M**	SD
VRON	**79.97**	32.97	**97.96**	20.07	**102.14**	18.75	**104.12**	18.05	**96.05**	21.02
VRMA	**72.50**	23.44	**90.25**	21.28	**92.21**	25.42	**100.27**	18.78	**88.80**	19.18
VRON+VRMA	**76.23**	28.37	**94.10**	20.70	**97.17**	22.52	**102.19**	18.21	**92.43**	20.11
Error in cpm (No per min)										
	Repetition 1		Repetition 2		Repetition 3		Repetition 4		All repetitions	
	M	SD	**M**	SD	**M**	SD	**M**	SD	**M**	SD
VRON	**28.38**	21.97	**9.37**	13.93	**7.31**	11.73	**6.82**	9.69	**12.97**	13.50
VRMA	**27.87**	22.93	**14.07**	16.48	**13.34**	20.46	**8.14**	12.04	**15.85**	15.22
VRON+VRMA	**28.12**	22.06	**11.72**	15.18	**10.32**	16.67	**7.48**	10.76	**14.41**	14.21
Correct compressions (%)										
	Repetition 1		Repetition 2		Repetition 3		Repetition 4		All repetitions	
	M	SD	**M**	SD	**M**	SD	**M**	SD	**M**	SD
VRON	**31**	24	**45**	29	**47**	27	**51**	30	**44**	23
VRMA	**23**	21	**30**	29	**50**	31	**46**	30	**37**	23
VRON+VRMA	**27**	22	**38**	30	**48**	28	**49**	29	**40**	23

Too deep compressions (%)										
	Repetition 1		Repetition 2		Repetition 3		Repetition 4		All repetitions	
	M	SD	**M**	SD	**M**	SD	**M**	SD	**M**	SD
VRON	**47**	24	**34**	30	**28**	24	**21**	23	**32**	22
VRMA	**10**	11	**6**	7	**9**	15	**9**	9	**9**	8
VRON+VRMA	**28**	26	**20**	26	**18**	22	**15**	18	**20**	20
Too shallow compressions (%)										
	Repetition 1		Repetition 2		Repetition 3		Repetition 4		All repetitions	
	M	SD	**M**	SD	**M**	SD	**M**	SD	**M**	SD
VRON	**12**	16	**11**	9	**13**	8	**13**	8	**13**	7
VRMA	**50**	31	**48**	36	**35**	35	**32**	36	**41**	30
VRON+VRMA	**31**	31	**30**	32	**24**	28	**23**	27	**27**	26
Compressions with incomplete release (%)										
	Repetition 1		Repetition 2		Repetition 3		Repetition 4		All repetitions	
	M	SD	**M**	SD	**M**	SD	**M**	SD	**M**	SD
VRON	**6**	6	**8**	9	**5**	5	**3**	3	**6**	3
VRMA	**9**	17	**13**	21	**7**	15	**6**	13	**9**	15
VRON+VRMA	**8**	13	**11**	16	**6**	11	**5**	9	**7**	11

Table 2. Mean and standard deviation of all measures concerning manual skills during training.

The analyses showed a main effect of repetition on cpm, $F(1.73,48.38)=24.73$, $p<0.001$, $\eta_p^2=0.47$, error in cpm, $F(1.84,51.65)=25.45$, $p<0.001$, $\eta_p^2=0.48$, correct compressions, $F(3,84)=10.63$, $p<0.001$, $\eta_p^2=0.28$, too deep compressions, $F(3,84)=6.81$, $p<0.001$, $\eta_p^2=0.20$, and compressions with incomplete release, $F(2.24,62.72)=3.28$, $p<0.05$, $\eta_p^2=0.11$. Table 3 reports the p-value of each pairwise comparison. The main effect of repetition on the other measures was not statistically significant ($p>0.05$ for all).

Repetition	Repetition	Cpm	Error in cpm	Correct compressions	Too deep compressions	Compressions with

						incomplete release
1	2	< 0.001	< 0.001	NS	NS	NS
	3	< 0.001	< 0.001	< 0.001	< 0.05	NS
	4	< 0.001	< 0.001	< 0.001	< 0.005	NS
2	3	NS	NS	NS	NS	NS
	4	< 0.001	NS	NS	NS	NS
3	4	NS	NS	NS	NS	NS

Table 3. Statistical significance (p-value) of pairwise comparisons between repetitions for cpm, error in cpm, correct compressions, too deep compressions, and compressions with incomplete release, during training. NS indicates statistical significance was not reached (p>0.05).

The analyses showed an interaction between group and repetition on compression depth, $F(3,84)=3.87$, $p<0.05$, $\eta_p^2=0.12$, too deep compressions, $F(3,84)=6.50$, $p<0.001$, $\eta_p^2=0.19$, and too shallow compressions, $F(3,84)=3.23$, $p<0.05$, $\eta_p^2=0.10$. For each group (respectively repetition), Table 4 (Table 5) reports the p-value of each pairwise comparison between repetitions (respectively groups). For the other measures, interaction was not statistically significant (p>0.05 for all).

Group	Repetition	Repetition	Compression depth	Too deep compressions	Too shallow compressions
VRON	1	2	NS	< 0.05	NS
		3	NS	< 0.001	NS
		4	NS	< 0.001	NS
	2	3	NS	NS	NS
		4	NS	< 0.01	NS
	3	4	NS	NS	NS
VRMA	1	2	NS	NS	NS
		3	NS	NS	< 0.05
		4	< 0.05	NS	< 0.05

	2	3	NS	NS	< 0.05
		4	< 0.05	NS	< 0.005
	3	4	NS	NS	NS

Table 4. Statistical significance (p-value) of pairwise comparisons between repetitions on average depth of chest compression, too deep compressions, and too shallow compressions for each group during training. NS indicates statistical significance was not reached ($p>0.05$).

Repetition	Compression depth	Too deep compressions	Too shallow compressions
1	< 0.005	< 0.001	< 0.001
2	< 0.005	< 0.005	< 0.001
3	< 0.05	< 0.05	< 0.05
4	NS	NS	NS

Table 5. Statistical significance (p-value) of pairwise comparisons between groups on average depth of chest compression, too deep compressions, and too shallow compressions for each repetition, during training. NS indicates statistical significance was not reached ($p>0.05$).

4.2. Transfer of manual skills and procedural knowledge in final assessment

During final assessment, all participants performed the first repetition of chest compressions, but one participant in the VRMA group forgot to perform the second one. Therefore, we were able to analyze data from 30 participants for measures concerning manual skills in the first repetition (Figure 5), and 29 participants for measures in the second repetition (Figure 6).

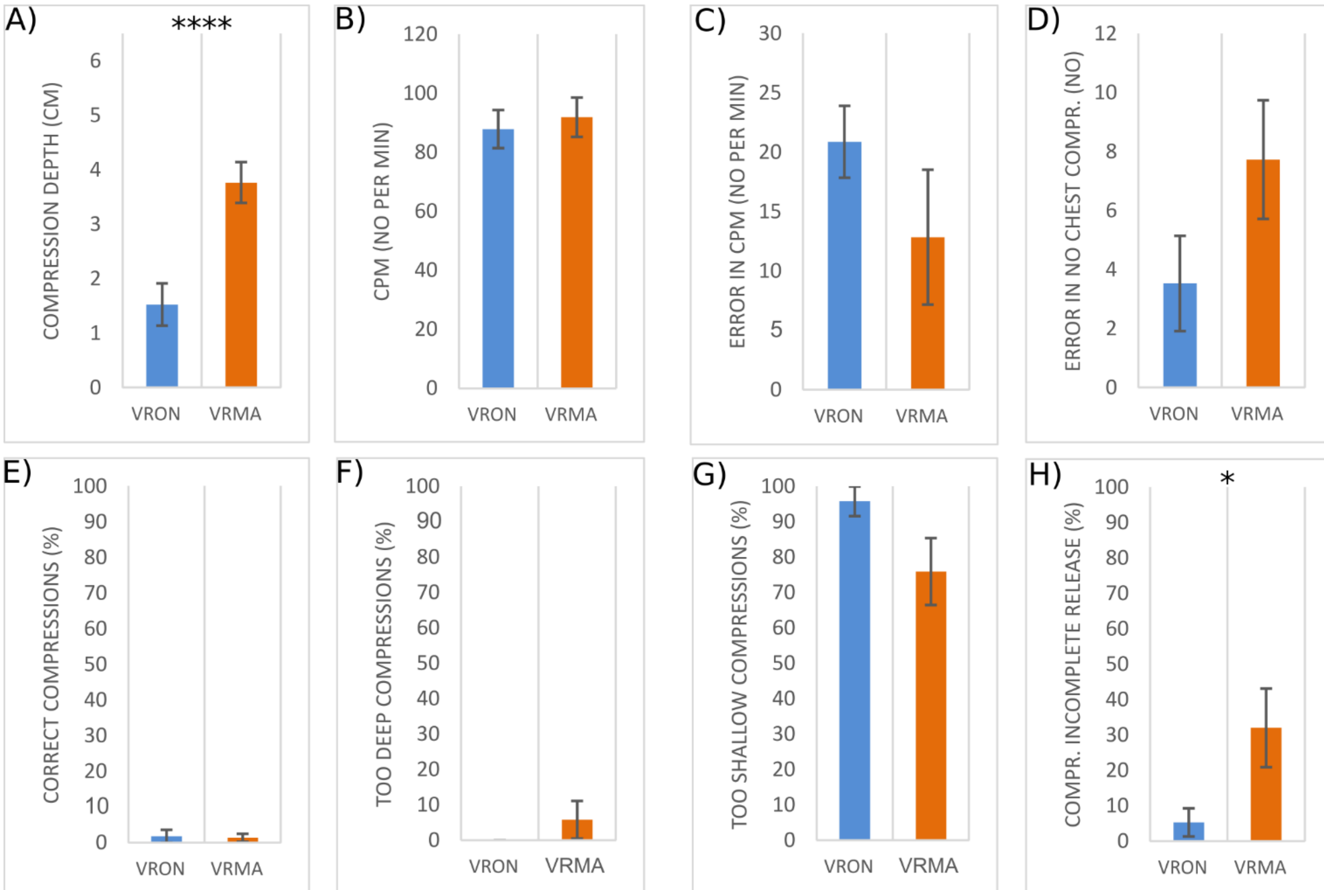


Figure 5. Mean of all measures concerning manual skills during the first repetition of chest compressions in final assessment. Capped vertical bars indicate ± SE. The *, **** signs indicate statistically significant differences between the two groups with p-values respectively <0.05, <0.001.

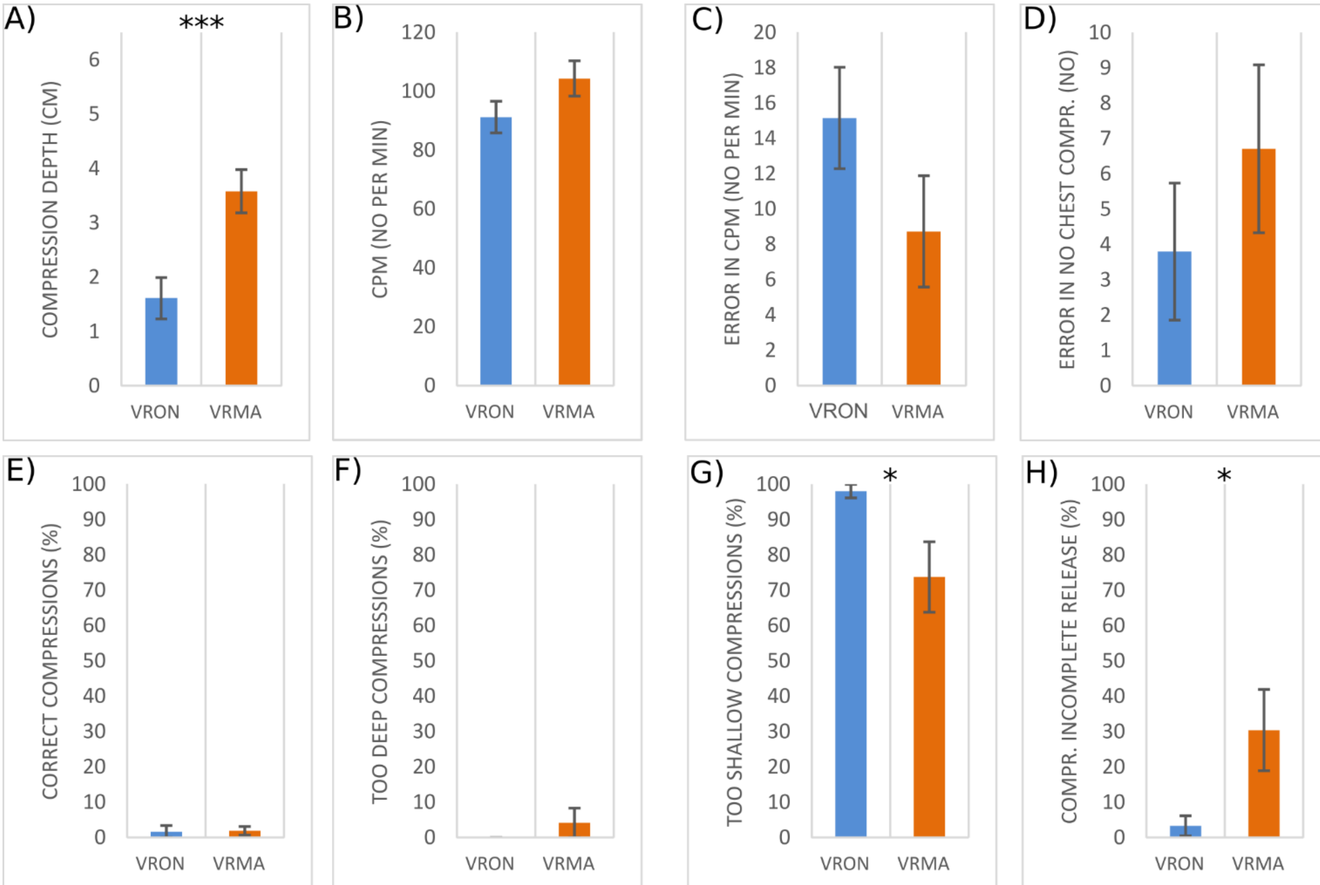


Figure 6. Mean of all measures concerning manual skills during the second repetition of chest compressions in final assessment. Capped vertical bars indicate ± SE. The *, *** signs indicate statistically significant differences between the two groups with p-values respectively <0.05, <0.005.

For the first repetition of chest compressions, ANOVAs revealed statistically significant differences between the two groups for compression depth, $F(1,28)=17.22$, $p<0.001$, $\eta_p^2=0.38$, and compressions with incomplete release, $F(1,28)=5.13$, $p<0.05$, $\eta_p^2=0.16$. For the second repetition of chest compressions, ANOVAs revealed statistically significant differences between the two groups for compression depth, $F(1,27)=12.79$, $p<0.005$, $\eta_p^2=0.32$, too shallow compressions, $F(1,27)=6.15$, $p<0.05$, $\eta_p^2=0.19$, and compressions with incomplete release, $F(1,27)=5.55$, $p<0.05$, $\eta_p^2=0.17$. In both repetitions, compression depth (Figure 5A and 6A) was higher with VRMA (1st: M=3.77, SD=1.45; 2nd: M=3.58, SD=1.49) than VRON (1st: M=1.52, SD=1.51; 2nd: M=1.61, SD=1.47). Compressions with incomplete release (Figure 5H and 6H) were higher with VRMA (1st: M=32%, SD=43; 2nd: M=30%, SD=43) than VRON (1st: M=5%, SD=15; 2nd:

M=3%, SD=11). Too shallow compressions in the second repetition (Figure 6G) were higher with VRO (M=98%, SD=7) than VRM (M=74%, SD=37).

No statistically significant difference between the groups was found on too shallow compressions during the first repetition, and on cpm, error in cpm, correct compressions, too deep compressions, and error in the number of chest compressions during both repetitions ($p>0.05$ for all).

Among the measures concerning procedural knowledge (Figure 7), ANOVAs revealed a statistically significant difference between the two groups only on the number of wrong steps, $F(1,28)=7.39$, $p<0.05$, $\eta_p^2=0.21$, which was higher with VRMA (M=1.07, SD=1.16) than VRON (M=0.20, SD=0.41). No statistically significant difference was found on the other measures concerning procedural knowledge ($p>0.05$ for all).

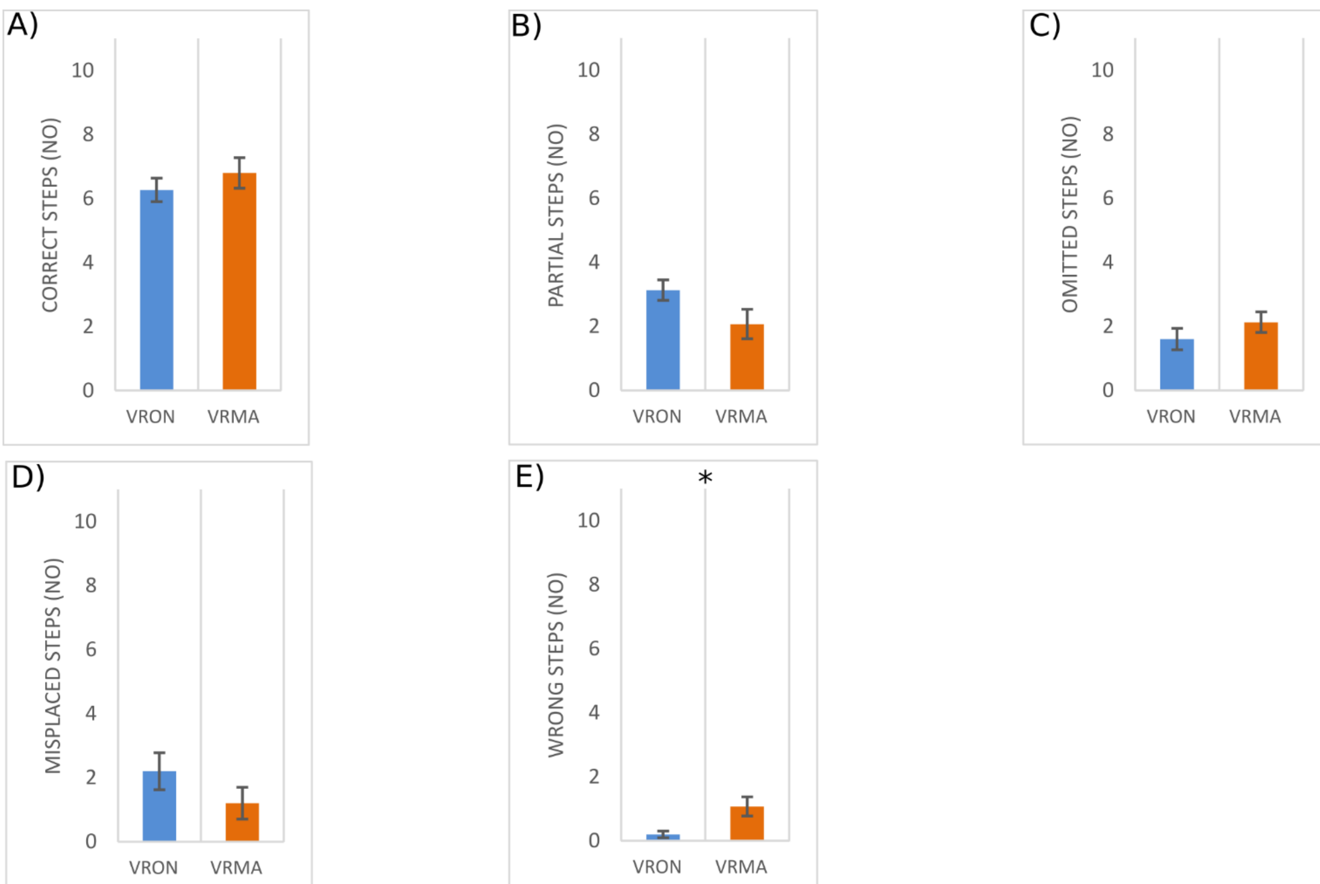


Figure 7. Mean of measures concerning procedural knowledge in final assessment. Capped vertical bars indicate ± SE. The * sign indicates a statistically significant difference between the two groups with p-value <0.05.

4.3. Self-efficacy

Mixed-design ANOVAs showed no main effect of group, $p>0.05$, a main effect of time of measurement, $F(1.36,38.08)=25.45$, $p<0.001$, $\eta_p^2=0.59$, and no interaction between group and time of measurement, $p>0.05$. Pairwise comparisons showed that pre-training self-efficacy ($M=3.24$, $SD=1.47$) was lower than post-training ($M=4.73$, $SD=1.03$) as well as post-assessment ($M=4.88$, $SD=1.15$) self-efficacy, $p<0.001$ for both, while no difference was found between post-training and post-assessment self-efficacy, $p>0.05$. Figure 8 shows evolution of self-efficacy over the three times of measurement for the two groups.

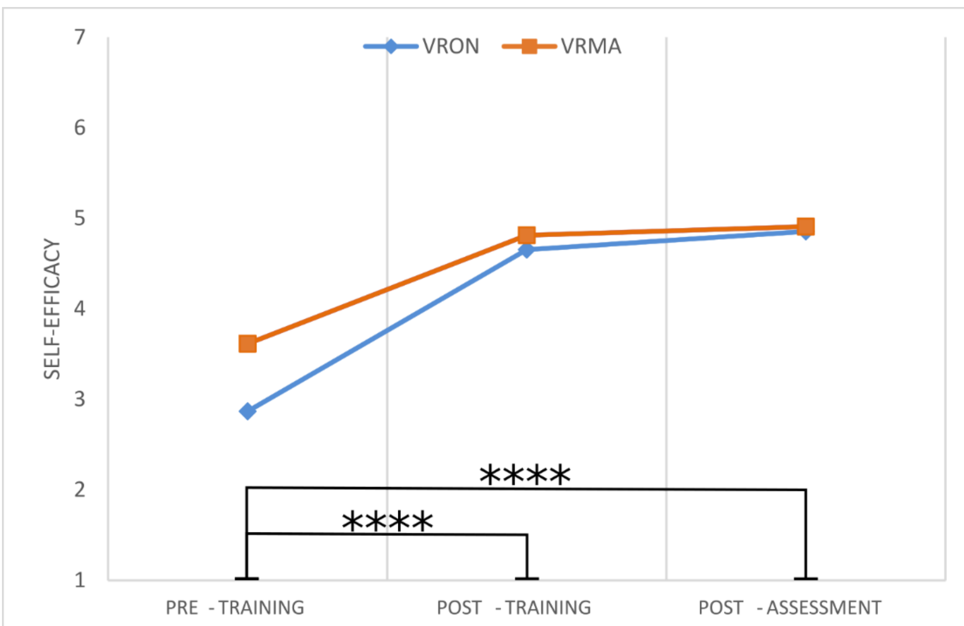


Figure 8. Self-efficacy over the three times of measurement for the two groups. The **** signs indicate statistically significant pairwise comparisons between repetitions with p-values <0.001.

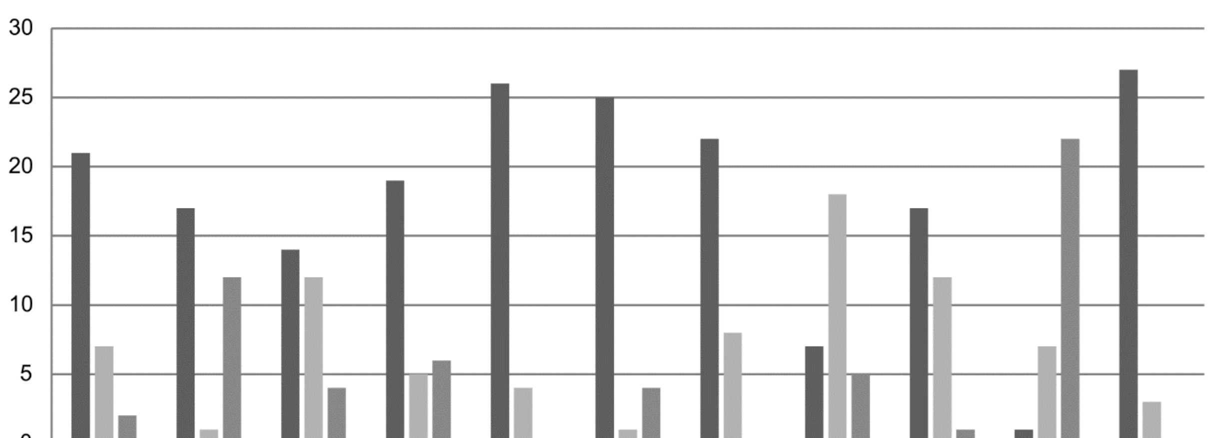
5. Discussion

The experimental evaluation showed the beneficial effects of the proposed VR methodology for CPR training on manual skills, procedural knowledge, and self-efficacy. Moreover, it shed light on the role of the physical feedback obtained through the mannequin chest.

More precisely, results about gain in manual skills during training showed that repeating chest compressions in VR could significantly reduce the error in cpm, with or without the addition of a mannequin (Figure 4C). In both groups, average cpm was more than 20 compressions below the minimum during the first repetition, while it was in the correct range during the fourth, with both groups above 100 cpm (Figure 4B). Correct compressions were only 27% during the first repetition, while they were 49% in the fourth (Figure 4D), indicating a positive effect of the proposed VR methodology for CPR training. There was an improvement also on compressions with incomplete release (Figure 4G), but pairwise comparisons of this measure did not reach significance, probably because it was already low at first repetition. For too deep compressions, too shallow compressions, and compression depth, results showed an effect of group and an interaction between group and repetition. Compression depth was expectedly lower in the VRMA group (Figure 4A), because more effort was required to perform compressions on the mannequin than moving the hands without the physical feedback of the mannequin chest. The VRON group performed thus more too deep compressions (Figure 4E) and less too shallow compressions (Figure 4F) than the VRMA group. Results showed that both groups benefited from repetitions in VR: compression depth decreased in the VRON group and increased in the VRMA group (Figure 4A). In this way, the averages for both groups in the fourth repetition were in the correct range. The result was statistically significant for the VRMA group as confirmed by the significant interaction and the significant difference between first and fourth repetition, while it was not significant for the VRON group, probably because this group was already in the correct range of the measure in the first repetition. The other two measures further clarify the effects of VR with and without mannequin, showing that the percentage of too shallow compressions significantly decreased in the VRMA group and was low in all repetitions in the VRON group (Figure 4F), while the percentage of too deep compressions significantly decreased in the VRON group and was low in all repetitions in the VRMA group (Figure 4E).

Overall, results concerning gain in manual skills assessed during training confirmed the hypothesized general improvement between the first and the last repetition, extending the results found in [33],[34] to CPR. Results also confirmed our hypothesis that improvement in the two groups would differ on measures concerning pressure, which indeed showed different improvement trends (Figure 4A, 4E, and 4F). Nevertheless, repeating chest compressions in VR had a beneficial effect on all measures, suggesting that periodical training in VR could be a convenient approach to counterbalance the inevitable decay in skill retention [40]. It is worth noting that the proposed VR methodology can deliver CPR training using consumer VR devices that have a much lower cost per user than traditional training with mannequins and instructors. This would allow to increase the frequency of CPR training sessions. From this perspective, longitudinal studies are needed to assess how much periodical repetitions of VR trials could be effective to promote retention over time.

Results of the final assessment confirmed our hypothesis that procedural knowledge acquired by participants during training transferred to performing a CPR procedure on the mannequin without the assistance of VR. Participants in both groups correctly remembered most of the steps of the procedure during final assessment (Figure 7A). Figure 9 shows for how many participants each step was correct, partial, or omitted, highlighting that the two steps with more errors were those that participants did not try directly, because it was the other rescuer who performed them in VR. This is in line with constructivist theories [28] and confirms the results of other experimental evaluations in which performing procedures in VR transfered to the real world [35][37]-[39]. It is worth noting that the presence or absence of the physical feedback provided by the mannequin had no statistically significant effect on most measures concerning procedural knowledge (Figure 7). We found a difference only on wrong steps, which were higher with the mannequin, but the average number of wrong steps was low in both groups (0.20 in VRON and 1.07 in VRMA).



	Safety check	Shoulder shaking	Head rotation	Airways check	Breath check	Phone call	Chest compres-sion	Insuffla-tion	Chest compres-sion	Insuffla-tion	Breath check
Correct	21	17	14	19	26	25	22	7	17	1	27
Partial	7	1	12	5	4	1	8	18	12	7	3
Omitted	2	12	4	6	0	4	0	5	1	22	0

Figure 9. Number of participants for whom the different steps were correct, partial, or omitted.

Considering transfer of manual skills, no statistically significant difference between groups was found on cpm and error in cpm. In the final assessment, the average cpm was in the correct range only for the VRMA group in the second repetition (Figure 6B), but all the other values of this measure were close to the range (Figure 5B and 6B), and better than the corresponding averages in the first repetition in VR (Figure 4B). On the contrary, the addition of the mannequin had a fundamental role on pressure-related measures: as hypothesized, compression depth was significantly higher in participants that tried the procedure with the physical feedback of the mannequin (Figure 5A and 6A). This measure was below the correct range for both groups and both repetitions in the final assessment, but the distance from the correct range was much higher in the VRON group. The other measures showed that too deep compressions (Figure 5F and 6F) were very rare for both groups (below 6%), while too shallow compressions (Figure 5G and 6G) were very common (above 73%), with a significant difference between groups in the second repetition (where the measure reached 98% for VRON), confirming that training in VR with the mannequin leads to less pressure-related errors. There were less incomplete releases (Figure 5H and 6H) without the mannequin because compression depth in the VRON group was very low (Figure 5A and 6A). Low percentages of correct compressions (Figure 5E and 6E) can be explained by the high percentages of too shallow compressions (especially in the VRON group) and compressions with incomplete release (in the VRMA group). Since compression depth increased by repeating chest compressions in VR with the mannequin, retraining with the same equipment might help further increasing compression depth, reducing too shallow compressions, and increasing correct ones.

Overall, results of the final assessment showed that the proposed VR methodology was effective to learn the CPR procedure and that the integration of the mannequin is needed only to train in manual skills concerning pressure. Since the VR application is able to monitor relevant measures using an inexpensive Arduino board and displays the full virtual body of the person to resuscitate, a basic version of a mannequin without legs, arms, and additional sensors is sufficient to support the integration, enabling low-cost, frequent training and refresher training sessions.

Results showed that the proposed VR methodology for CPR training had a positive effect also on self-efficacy (Figure 8), confirming our hypothesis and extending the results found in [5][31],[35]. Participants were more confident in their ability to perform CPR after using the VR application, regardless of the presence of a mannequin. We measured self-efficacy immediately after training as well as after final assessment, because a poor performance in the final assessment might have had a negative effect also on self-efficacy. On the contrary, no significant difference was found between these two times of measurements and both were higher than initial self-efficacy assessed before training. Self-efficacy could be particularly important in emergency medicine because it plays a role on performance outcomes, and people with similar skills may perform differently depending on variations in their self-efficacy, as suggested by Social Cognitive Theory [32][41].

A limitation of the experimental evaluation concerns the small sample of participants that were involved. While the size of the sample is consistent with typical evaluations in the field, especially those that required participants to wear a VR headset [16][17],[19],[21]-[24], and it was sufficient to show the statistically significant differences described above, a wider sample might have led to additional significant results. The evaluation compared the proposed VR methodology for CPR training with and without the physical mannequin, but did not include a control condition based on traditional, non-VR training with instructors. However, the proposed evaluation methodology could be followed to carry out future evaluations that could compare VR approaches with other, non-VR training approaches. The proposed evaluation methodology could also be extended to evaluate retention of knowledge and skills. Indeed, the experimental evaluation described in this paper assessed transfer of both manual skills and procedural knowledge after training, but knowledge and skills naturally decay over time [40]. Future evaluations are needed to assess how much periodical training in VR could improve and maintain CPR knowledge and skills, counteracting their decay. Moreover, assessment of knowledge and skills on a mannequin and in a laboratory setting is not the same as an assessment on a real patient suffering from a cardiac arrest in a real setting. Future work should exploit the potential of VR to make VR training more similar to real world settings, gradually adding distracting elements (e.g., loud environmental sounds or bystanders who distract and interfere with responders) and time pressure (e.g., virtual patients whose condition keep worsening and may die) as trainees progress with the trials.

Nevertheless, the results obtained in the described experimental evaluation already support the adoption of novel training methodologies in which trainees begin to learn CPR autonomously through VR using consumer, home VR devices as the ones we employed in the implementation of the proposed VR methodology for CPR training. This can reduce the amount of time human instructors are needed, and does not require to conduct all training in traditional classrooms or training centers. Moreover, trying the procedure on a mannequin would be necessary only to learn appropriate pressure. This reduces dependence on the continuous availability of a mannequin, because most aspects of CPR could be learned using VR only.

The described experimental evaluation showed the positive effects of the proposed VR methodology on CPR training, but the methodology can be applied to train in other BLS and ALS procedures that require both procedural knowledge and manual skills, such as the use of a defibrillator or more complex evaluations of patient symptoms. The proposed evaluation methodology could then be extended with new measures that could take into account the additional BLS and ALS skills and knowledge.

6. Conclusions

The experimental evaluation showed that the proposed VR methodology for CPR training benefits procedural knowledge and manual skills, also when a physical mannequin is not used. Trials on a mannequin are needed only to train in manual skills concerning the correct pressure for chest compression. The experimental evaluation suggests that the introduction of VR is promising to broaden CPR training at affordable costs. The next step in our research will concern the extension of the methodology to blend VR, mannequin, and instructor-based training in teaching and refreshing CPR over time. In particular, we will extend the VR methodology to gradually introduce complex challenges into the simulated scenarios making them more similar to the realistic settings where cardiac arrest could happen, and schedule a series of trials of increasing complexity to promote retention of procedural knowledge and manual skills over time. We are going to test the improved methodology with a sample of health professionals who belong to medical fields other than emergency medicine. This sample is ideal to assess retention because such medical participants are professionally required to be able to perform CPR, but very rarely (or never) have to perform it in their regular practice. This would be a further step that would evaluate the effectiveness of the proposed VR methodology to improve CPR refresher courses for health professionals.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was part of workpackage WP2 of the EasyNet project, funded by the Italian Ministry of Health and the Friuli Venezia Giulia region, project number NET-2016-02364191. The sponsors were not involved in the design, the writing, and the decision to submit the manuscript for publication.

We wish to thank Nicola Zangrando, who carried out 3D modeling and programming activities for the VR application used in the experimental evaluation. We also thank Antonio Di Chiara (cardiologist) and Sabrina Licata (pharmacologist) for sharing their expert knowledge.

References

[1] Virani SS, Alonso A, Benjamin EJ, et al. Heart Disease and Stroke Statistics-2020 Update: A Report From the American Heart Association. 2020. doi:10.1161/CIR.0000000000000757

[2] Cheng A, Nadkarni VM, Mancini MB, et al. Resuscitation Education Science: Educational Strategies to Improve Outcomes From Cardiac Arrest: A Scientific Statement From the American Heart Association. 2018. doi:10.1161/CIR.0000000000000583

[3] Greif R, Lockey AS, Conaghan P, et al. European Resuscitation Council Guidelines for Resuscitation 2015. Section 10. Education and implementation of resuscitation. Resuscitation 2015;95:288–301. doi:10.1016/j.resuscitation.2015.07.032

[4] Rumsfeld JS, Brooks SC, Aufderheide TP, et al. Use of mobile devices, social media, and crowdsourcing as digital strategies to improve emergency cardiovascular care. Circulation 2016;134:e87–108. doi:10.1161/CIR.0000000000000428

[5] Creutzfeldt J, Hedman L, Heinrichs L, et al. Cardiopulmonary resuscitation training in high school using avatars in virtual worlds: an international feasibility study. J Med Internet Res 2013;15:e9. doi:10.2196/jmir.1715

[6] Wattanasoontorn V, Magdics M, Boada I, et al. A kinect-based system for cardiopulmonary resuscitation simulation: A pilot study. In: 4th International Conference on Serious Games Development and Applications. Springer, Berlin, Heidelberg 2013. 51–63. doi:10.1007/978-3-642-40790-1_5

[7] Latif MH, Ajmal M, Ahmad F, et al. LA-VIE: A serious game for cardiopulmonary resuscitation. In: 2017 IEEE 5th International Conference on Serious Games and Applications for Health. Institute of Electrical and Electronics Engineers Inc. 2017. doi:10.1109/SeGAH.2017.7939297

[8] Boada I, Rodriguez-Benitez A, Garcia-Gonzalez JM, et al. 30 : 2: A Game Designed to Promote the Cardiopulmonary Resuscitation Protocol. Int J Comput Games Technol 2016. doi:10.1155/2016/8251461

[9] Morrison-Smith S, Yao H, Wang I, et al. Staying alive with virtual humans. In: Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems. New York, New York, USA: Association for Computing Machinery 2018. 1–6. doi:10.1145/3170427.3188477

[10] Nas J, Thannhauser J, Vart P, et al. Rationale and design of the Lowlands Saves Lives trial: A randomised trial to compare CPR quality and long-term attitude towards CPR performance between face-to-face and virtual reality training with the Lifesaver VR app. BMJ Open 2019;9:e033648. doi:10.1136/bmjopen-2019-033648

[11] Nas J, Thannhauser J, Vart P, et al. Effect of Face-to-Face vs Virtual Reality Training on Cardiopulmonary Resuscitation Quality: A Randomized Clinical Trial. JAMA Cardiol 2020;5:328–35. doi:10.1001/jamacardio.2019.4992

[12] Leary M, McGovern SK, Chaudhary Z, et al. Comparing bystander response to a sudden cardiac arrest using a virtual reality CPR training mobile app versus a standard CPR training mobile app. Resuscitation 2019;139:167–73. doi:10.1016/j.resuscitation.2019.04.017

[13] Semeraro F, Frisoli A, Loconsole C, et al. Motion detection technology as a tool for cardiopulmonary resuscitation (CPR) quality training: A randomised crossover mannequin pilot study. Resuscitation 2013;84:501–7. doi:10.1016/j.resuscitation.2012.12.006

[14] Higashi E, Fukagawa K, Kasimura R, et al. Development and evaluation of a corrective feedback system using augmented reality for the high-quality cardiopulmonary resuscitation training. In: 2017 IEEE International Conference on Systems, Man, and Cybernetics. Institute of Electrical and Electronics Engineers Inc. 2017. 716–21. doi:10.1109/SMC.2017.8122692

[15] Balian S, McGovern SK, Abella BS, et al. Feasibility of an augmented reality cardiopulmonary resuscitation training system for health care providers. Heliyon 2019;5:e02205. doi:10.1016/j.heliyon.2019.e02205

[16] Varun Durai SI, Arjunan R, Manivannan M. Affordable hi-fidelity VR based CPR simulator with haptics feedback. In: 2017 IEEE International Symposium on Haptic, Audio-Visual Environments and Games. Institute of Electrical and Electronics Engineers Inc. 2017. doi:10.1109/HAVE.2017.8240349

[17] Bucher K, Blome T, Rudolph S, et al. VReanimate II: training first aid and reanimation in virtual reality. J Comput Educ 2019;6:53–78. doi:10.1007/s40692-018-0121-1

[18] Vaughan N, John N, Rees N. CPR virtual reality training simulator for schools. In: 2019 International Conference on Cyberworlds. Institute of Electrical and Electronics Engineers Inc. 2019. 25–8. doi:10.1109/CW.2019.00013

[19] Semeraro F, Ristagno G, Giulini G, et al. Virtual reality cardiopulmonary resuscitation (CPR): Comparison with a standard CPR training mannequin. Resuscitation 2019;135:234–5. doi:10.1016/j.resuscitation.2018.12.016

[20] Gent L, Sarno D, Coppock K, et al. Abstract 396: Successful Virtual Reality Cardiopulmonary Resuscitation Training in Schools: Digitally Linking a Physical Manikin to a Virtual Lifesaving Scenario | Circulation. Circulation 2019;140.https://www.ahajournals.org/doi/abs/10.1161/circ.140.suppl_2.396.

[21] Almousa O, Prates J, Yeslam N, et al. Virtual Reality Simulation Technology for Cardiopulmonary Resuscitation Training: An Innovative Hybrid System With Haptic Feedback. Simul Gaming 2019;50:6–22. doi:10.1177/1046878118820905

[22] Aksoy E. Comparing the effects on learning outcomes of tablet-based and virtual reality-based serious gaming modules for basic life support training: Randomized trial. JMIR serious games 2019;7:e13442. doi:10.2196/13442

[23] Bench S, Winter C, Francis G. Use of a Virtual Reality Device for Basic Life Support Training. Simul Healthc J Soc Simul Healthc 2019;14:287–92. doi:10.1097/SIH.0000000000000387

[24] Semeraro F, Ristagno G, Giulini G, et al. Back to reality: A new blended pilot course of Basic Life Support with Virtual Reality. Resuscitation 2019;138:18–9. doi:10.1016/j.resuscitation.2019.02.034

[25] Buttussi F, Pellis T, Cabas Vidani A, et al. Evaluation of a 3D serious game for advanced life support retraining. Int J Med Inform 2013;82:798–809. http://www.sciencedirect.com/science/article/pii/S1386505613001147 (accessed 16 Sep 2013).

[26] Vankipuram A, Khanal P, Ashby A, et al. Design and development of a virtual reality simulator for advanced cardiac life support training. IEEE J Biomed Heal Informatics 2014;18:1478–84. doi:10.1109/JBHI.2013.2285102

[27] Khanal P, Vankipuram A, Ashby A, et al. Collaborative virtual reality based advanced cardiac life support training simulator using virtual reality principles. J Biomed Inform 2014;51:49–59. doi:10.1016/j.jbi.2014.04.005

[28] Chittaro L, Ranon R. Web3D technologies in learning, education and training: Motivations, issues, opportunities. Comput Educ 2007;49:3–18.

[29] Gallagher AG, Ritter EM, Champion H, et al. Virtual reality simulation for the operating room: Proficiency-based training as a paradigm shift in surgical skills training. Ann Surg 2005;241:364–72. doi:10.1097/01.sla.0000151982.85062.80

[30] Ragan ED, Bowman DA, Kopper R, et al. Effects of Field of View and Visual Complexity on Virtual Reality Training Effectiveness for a Visual Scanning Task. IEEE Trans Vis Comput Graph 2015;21:794–807.

[31] Buttussi F, Chittaro L. Effects of Different Types of Virtual Reality Display on Presence and Learning in a Safety Training Scenario. IEEE Trans Vis Comput Graph 2018;24:1063–76. doi:10.1109/TVCG.2017.2653117

[32] Bandura A. Self-efficacy: the exercise of control. New York, NY, USA: Freeman 1997.

[33] Vankipuram M, Kahol K, McLaren A, et al. A virtual reality simulator for orthopedic basic skills: A design and validation study. J Biomed Inform 2010;43:661–8. doi:10.1016/j.jbi.2010.05.016

[34] Lin Y, Wang X, Wu F, et al. Development and validation of a surgical training simulator with haptic feedback for learning bone-sawing skill. J Biomed Inform 2014;48:122–9. doi:10.1016/j.jbi.2013.12.010

[35] Chittaro L, Corbett CL, McLean GA, et al. Safety knowledge transfer through mobile virtual reality: A study of aviation life preserver donning. Saf Sci 2017;102:159–68. doi:https://doi.org/10.1016/j.ssci.2017.10.012

[36] Schwarzer R, Jerusalem M. Generalized Self-Efficacy scale. In: Weinman J, Wright S, Johnston M, eds. Measures in health psychology: A user's portfolio. Causal and control beliefs. Windsor, UK: Nfer-Nelson 1995. 35–7.

[37] Ragan ED, Sowndararajan A, Kopper R, et al. The Effects of Higher Levels of Immersion on Procedure Memorization Performance and Implications for Educational Virtual Environments. Presence Teleoperators Virtual Environ 2010;19:527–43. doi:10.1162/pres_a_00016

[38] Ganier F, Hoareau C, Tisseau J. Evaluation of procedural learning transfer from a virtual environment to a real situation: A case study on tank maintenance training. Ergonomics 2014;57:828–43. doi:10.1080/00140139.2014.899628

[39] Buttussi F, Chittaro L. Humor and Fear Appeals in Animated Pedagogical Agents: An Evaluation in Aviation Safety Education. IEEE Trans Learn Technol 2020;13:63–76. doi:10.1109/TLT.2019.2902401

[40] Vlasblom JID, Pennings HJM, Pal J Van Der, et al. Competence retention in safety-critical professions: A systematic literature review. Educ Res Rev 2020;30. doi:10.1016/j.edurev.2020.100330

[41] Bandura A. Social cognitive theory: an agentic perspective. Annu Rev Psychol 2001;52:1–26. doi:10.1146/annurev.psych.52.1.1