



Exploring Extended Reality as a Simulation Training Tool Through Naturalistic Interactions and Enhanced Immersion

Daniel Duggan¹, Caroline Kingsley^{1(✉)}, Mark Mazzeo²,
and Michael Jenkins¹

¹ Charles River Analytics Inc., Cambridge, MA, USA
{dduggan, ckingsley, mjenkins}@cra.com

² Combat Capabilities Development Command Soldier Center, Simulation
and Training Technology Center (CCDC SC STTC), Orlando, FL, USA
mark.v.mazzeo.civ@mail.mil

Abstract. Augmented reality (AR), mixed reality (MR), and virtual reality (VR) technologies – collectively extended reality (XR) – have been discussed in the context of training applications since their introduction to the market. While this discussion was initially founded on potential benefit, the rapid pace and advancements within the XR industry allow the technology to fulfill initial expectations in terms of training utility. As XR becomes an effective training tool, two open areas of research, among many, include integrating experiences between VR and AR/MR simulations, as XR training often obscures the trainer from the trainee’s environment, and determining the optimal level of fidelity and overall usability of interaction within XR simulations for the training of fine- and gross-motor control tasks, as new peripherals and technologies enter the market supporting a variety of interaction fidelities, such as data glove controllers, skeletal motion capture suits, tracking systems, and haptic devices. XR may show more promise as a training platform the closer it can replicate real world and naturalistic training interactions and immersion. This paper discusses efforts in these two research areas, including a planned usability study on the impact of the fidelity of interactions on training of fine- and gross-motor control tasks in virtual environments.

Keywords: Extended reality · Naturalistic interaction · Training transfer · Virtual reality · Augmented reality · Mixed reality · Interaction fidelity

1 Introduction

This paper presents an overview of our ongoing work across the Extended Reality (XR) Head-Mounted Displays (HMDs) and Simulation domain, with a focus on two areas; integrating experiences between Augmented and Mixed Reality HMDs and Virtual Reality HMDs; and a planned usability study on the impact of the fidelity of interactions on training of fine- and gross-motor control tasks in virtual environments. While seemingly disparate, this work is being done under a larger, existing research and development program exploring ways to increase users’ naturalistic interactions

(through the support of a variety of devices including AR/MR HMDs, VR HMDs, data glove controllers, skeletal motion capture suits, tracking systems, and haptic devices) and immersion (through time- and event-based scripting for dynamic environments, voice-based interactions, etc.) in extended reality environments, with the goal of improving XR's utility as a training tool [1, 2]. The projected outcome of this larger effort is an open-source XR software development kit, designed to enable more rapid and consistent development of XR-based simulations with a focus on supporting natural human interactions to foster more realistic virtual simulations and training.

2 Networked Training

While the augmented reality (AR), mixed reality (MR), and virtual reality (VR) industries rapidly progress—especially in the healthcare training domain [3]—there remains a gap in how these extended realities can interface with one another. In an effort to advance the greater XR field in the spirit of its core principles, we are exploring how users wearing AR/MR HMDs and VR HMDs can observe each other's virtual environments and interactions. Our motivation to investigate this area is driven by the limitations of XR in the training domain [4]; specifically, while existing, non-XR computer-based training fits well into the traditional trainer-trainee model, whereas the nature of HMDs makes instruction more difficult as the trainer cannot directly see what actions the trainee is performing in their environment. Currently, the primary approach for observing a trainee's actions in XR is to watch from a computer screen that is mirroring the trainee's HMD screen. The mirrored screen approach is insufficient because it does not convey the spatial nature of manipulation in XR, particularly VR. Furthermore numerous VR and AR HMDs, especially the lower-cost models appealing to large scale training needs, are trending towards standalone models that do not interface with desktop computers. For example, the Oculus Go used in Walmart's experimental VR training is a standalone model. Microsoft's HoloLens and the Magic Leap One are the current front-runners in the AR domain and operate without a desktop computer.

We recently explored a solution to this problem in the training domain, allowing a “trainer” wearing an MR HMD (Magic Leap One) to observe a “trainee” in a virtual environment wearing a VR HMD (HTC Vive). In our prototype interface, focused on a medic training scenario based on Tactical Combat Casualty Care (TCCC) [5], we synchronized key data between the VR simulation and the trainer's MR HMD over a local network connection: the key data including the position and rotation of critical objects and the medical state of simulated patients. By initializing the origin point for both the MR and VR HMDs to the same point, we enabled the trainer to observe the trainee's environment overlaid on the same locations. The trainer was additionally allowed to view simulated patient data unavailable to the trainee such as heart rate, blood volume, and blood pressure. We believe that this asymmetrical multi-user approach will be valuable for creating training systems in the future; it lends itself to a classic trainer-trainee and/or teacher-student relationship where the trainer/teacher withholds information and the trainee/student is provided the tools or framework to

perform an action or produce an answer in relation to that information, for the trainer/teacher to then evaluate.

2.1 VR Training Simulation

Our VR training simulation was built in Unity using VSDK, an XR development kit we developed under an Army-funded project called VIRTUOSO, for rapid prototyping of XR scenarios. VSDK will be released as Open Source in 2019 and enables developers to create immersive simulations by providing interoperability between COTS XR hardware devices—including HMDs, controllers, tracking systems, hand trackers, skeletal trackers, and haptic devices—as well as systems supporting naturalistic XR interactions, including a high-level gesture recognition system, and an XR-focused event handling system. Our training simulation was focused on combat medic training and consisted of a narrative component followed by an active component where the trainee must treat several casualties following a blast.

For the VR training simulation, we constructed a virtual environment consisting of a city street with an adjacent alley and park in a fictional megacity. The virtual environment had three distinct zones: (1) a military checkpoint on the side of the street; (2) an alleyway adjacent to the roadblock; and (3) an open field or park where a medical helicopter could land. The trainee was able to freely teleport between the three areas. For this scenario, we created four simulated patients—three squad members and a squad leader. During the narrative component, the trainee is instructed to guard the alleyway. Once in the alleyway, they are instructed to participate in a “radio test” which is used to calibrate a voice recognition system for later portions of the medical training scenario. Once calibration is completed, an enemy vehicle arrives at the roadblock location and detonates, injuring the three squad members (the squad leader is uninjured and issues orders to the trainee). At this point, the trainee can proceed with treatment as they see fit. Each patient has unique injuries backed by a real-time patient simulation. The patient simulator is driven by a set of finite state machines with associated physiological variables. States correspond to different wounds, and trainee interventions transition state machines between states. For example, a wound might have the states bleeding and not bleeding. By applying a bandage to the wound, the state can transition to not bleeding which would affect the blood loss per tick.

Unscripted Training. Although combat medic training [5] specifies how treatment should be prioritized based on established triage principles (e.g. treat massive hemorrhage first, treat airway obstruction next, treat respiratory distress next, treat circulation-related issues next, treat head injury/hypothermia next, etc.) [5, 6], the trainee is free to treat the patients in any order in the simulation just as they are in the real world (applying trained triage principles accordingly). In the simulation, the trainee can use both voice commands and hand-based interactions to treat the patients, to mimic a real world situation as much as possible. For voice commands, the trainee has the option of asking the squad leader to perform a security sweep, asking patients if they can treat themselves, or calling for an evacuation helicopter on the radio. In terms of physical interaction fidelity, we mounted a Leap Motion hand tracker on the VR HMD to allow the user to interact with the tools using their hands rather than a

controller. This naturalistic interaction is supported with the goal of improving immersion and training transfer. Further enabling freedom of trainee decision and hopefully promoting prior triage training and retention, the virtual environment has several tools that the trainee can choose from and use to perform treatments, including bandages, scissors, a needle chest decompression, a tourniquet, and a Nasopharyngeal airway (NPA) tube, just as they may in their pack. In our simulation scenario, the first patient's injuries are an amputated leg and a small shrapnel wound on the chest. The trainee must apply a tourniquet to the patient's leg, remove his shirt with the scissors, and apply a bandage to the chest wound. The second patient has a significant chest wound that is causing blood to enter the lungs. Again the user must use the scissors to remove the shirt and apply a bandage to the wound, but a needle chest decompression is also required to prevent the patient's lungs from filling with blood. The third patient is unconscious and having trouble breathing, so an NPA tube is used to open the Nasopharyngeal airway. However, the only way a user can diagnose these injuries and decide how to treat them is through visual cues (i.e. visible wounds, casualty vitals, casualty skin pallor), existing knowledge, and the equipment they have available, as it would be in the field. Once all of the patients have been treated and the helicopter has been called, the trainee must load the patients (using a menu item) and then board the helicopter to finish the scenario.

2.2 MR Training Simulation

We built the MR trainer's application using VSDK as well, starting with the same environment as the trainee application. We removed the city environment because MR works best when only key elements of the scene are shown. We then built a networking component to enable us to send data to the MR application from the VR training simulation. In order to minimize network traffic and latency, we narrowed the shared data down to position and rotation of the trainee and patients as well as the state data of the patients (see Fig. 1). When configured correctly, the trainer would be able to see a helmet hovering over the trainee's HMD and would be able to see the trainee perform treatments. However, tool position was not synchronized so the current version only allowed the trainer to see the result of a treatment after it occurred. For example, after a tourniquet is applied to the first patient's leg, the trainer would see a tourniquet appear on the first patient's representation.

The trainer also receives additional information hidden from the trainee from the simulated patients. This information, such as blood pressure, heart rate, and hydration, would not be immediately measureable by the trainee without the proper tools but would be valuable to a trainer instructing a trainee on which patient or wound to address next. We displayed this information to the trainer via information cards floating above each patient (Fig. 1). The trainer can grab and move the information cards using the 6-degrees-of-freedom (6dof) controller provided with the Magic Leap One MR HMD in order to rearrange the information to see it better.

Another novel networked feature we implemented between our VR and AR simulation, is a mini-map feature. While the two environments are networked, if the AR user flips over the controller in their hand (in this case, the 6dof controller provided with the Magic Leap One), a mini-map spawns above and locks to the backside of the

controller, allowing the user in AR to see a map of the full simulation environment. Within that mini-map, the user in VR is represented by a colored icon within the map, as well as the user in AR. We believe this functionality can greatly enhance situational awareness between the user in AR and the user in VR, as it provides a snapshot of the full VR environment as well as the two users' respective locations in the context of that environment.

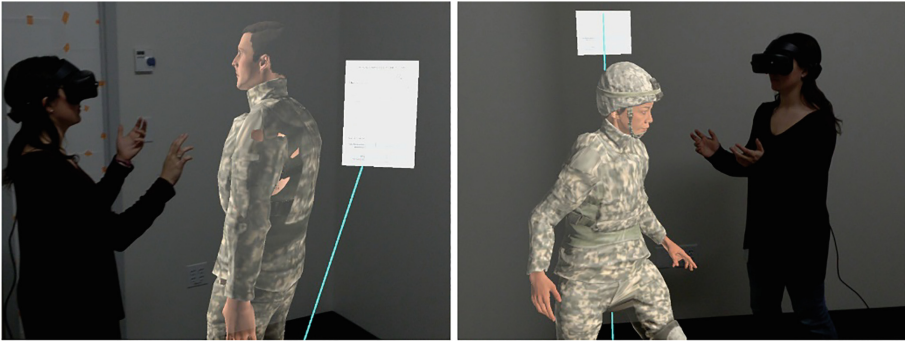


Fig. 1. These images demonstrate networking between our VR environment (worn by user pictured) and AR environment (worn by photographer), from the AR point of view (Magic Leap One). This simulation shows the presence of a non-player character (NPC) which is present in both VR and AR environments, as well as dynamic TCCC card [5] depicting NPC health state metrics, only viewable by the user in AR.

2.3 Recommendations

We believe this approach helps to address one of the key limitations in XR training, which is the reduced ability of trainers to observe trainee behavior and provide recommendations. However, our current implementation has a few drawbacks that could be improved on. The main drawback is that communication is only in one direction: from the trainee's VR simulation to the trainer's MR HMD. While this gives the trainer the ability to observe the trainee, it does not provide them any opportunities for providing additional information or assistance to the trainee (unless the trainee removes their headphones, reducing immersion). We are currently exploring approaches for meaningful bidirectional collaboration, including reversed visualizations, remote observation, and trainer tools for manipulation of the environment and the simulation.

Reversed Visualizations. Reversed visualizations are an important next step, not only for educational benefits but also for trainee safety. When operating in a VR environment, a trainee cannot see the real environment and is at risk of colliding with a trainer. While most trainers can stay out of the way, any collisions are likely to reduce user trust in virtual reality technology. We are exploring two avenues for supplying trainer positional data to the trainee: skeletal motion capture suits and networked transmission of AR/MR HMD position. Skeletal motion capture suits (such as Xsens or Synertial), frequently used by animation studios to capture actor performance, are seeing increased

interest in the XR domain because they enable applications to track the user's entire body [7, 8]. We have explored using real time motion capture to pose a 3D model of a human in real time, allowing one user in VR to see another who is wearing the motion capture suit. Full body capture requires a desktop computer, which limits the usage for trainers with standalone HMDs. The alternative option is to transmit the position and rotation of the trainer's HMD to the trainee's application. This seems like an option that is more likely to gain traction due to the high cost of tracking suits. Both options enable the trainer to direct the trainee; even without additional features, the trainer can point and gesture to patients and injuries for non-verbal instruction.

Remote Observation. The current solution is only configured to work on a local network. Future versions could be configured to work over distance to enable remote observation. Remote observation could benefit trainees working in distant locations far from training centers without missing out on the experience of expert trainers. Adopting a more robust networking protocol could also enable multiple team members to train together. This would likely decrease existing barriers to training and improve overall training adherence.

Real Time Manipulation. The next step beyond allowing the trainer to see the trainee's environment is to enable the trainer to modify it in real time to create a more collaborative training environment. Depending on the task and training context, trainees are likely to be challenged by different types and aspects of scenarios. Trainers are ideally positioned to provide augmentations to training programs that account for specific trainee needs. Future developments of the extended reality integration should enable trainers to make meaningful augmentations. In the medic training example, a trainer may notice that the trainee has difficulty applying tourniquets in the default scenario; the trainer could augment the scenario by introducing additional casualties with injuries that require a tourniquet for treatment. Alternatively, the trainer may want to remove other casualties to encourage the trainee to focus or even step in and demonstrate the appropriate procedure. We can also see these first steps develop into full XR scenario customization systems to allow for even more specially tailored training scenarios unique to each trainer or trainee.

3 Training and Interaction Fidelity

While the coexistence of MR and VR HMDs strives to improve the trainee-trainer relationship within virtual simulation training, we are also exploring improved training transfer from virtual to real-world environments through fidelity of interactions. Specifically, we have designed and will be conducting a usability study designed to answer this research question: does the fidelity of fine- and gross-motor control interactions impact training of fine- and gross-motor control tasks in virtual environments? Our study will consist of three phases: a training phase, a usability phase, and finally an evaluation phase. The usability phase is dual purpose and strives to both distract the subjects from their initial training and to collect meaningful data on which levels of fidelity are appropriate for different tasks. During the training phase, the subjects will be trained in a particular task using both a real world device and a virtual

reproduction of that task, restricted to a level of fidelity. During the evaluation phase, the subjects will be instructed to perform the real world task a second time.

This planned study will elicit a mix of quantitative (i.e. task execution, reaction time, accuracy, errors) and qualitative data (i.e. usability questionnaire, simulator sickness questionnaire, NASA-TLX workload assessment). To balance this need with other constraints (i.e. time, recruiting), our goal is to run at least 30 participants for the planned study, 10 participants for each of our three study tasks.

3.1 Study Tasks

We have selected three tasks for study, based on difficulty, trainability, and ease of reproduction in virtual environments. We also selected these tasks to include a mixture of fine- and gross-motor control skills. Tasks must be difficult enough that some subjects could struggle with the task and yet trainable so that subjects can improve with virtual training. Our three tasks for the planned study are: (1) wire-tracing, fine-motor; (2) mirrored writing, fine-motor, cognitive; and (3) drum rhythms, gross-motor (Fig. 2). We will vary both the task and the level of fidelity on a per-subject basis.



Fig. 2. These images depict the real world materials that will be used for our three planned study tasks. From left to right; mirrored writing apparatus, electronic drum kit, and wire tracing game.

Fine Motor Wire Tracing Task. For the wire-tracing task (available off the shelf as an ‘electronic wire game’) subjects will need to move a small, wand-like object around a rigid wire. The wand is conductive and if it touches the wire a buzzer sounds. This seemingly simple task requires both hand-eye coordination and fine-motor skills to complete without error. We will be tracking both time to complete the task and the number of errors to determine a task completion score. Once the user has tried the real world version of the task, they will be able to train with an XR recreation of the real world object. The virtual version of the wire-tracing task will be an exact replica; our goal is to have the user training model the real world as closely as possible. Since this is a primarily hand-based task, we will be varying the fidelity of hand-based interactions by changing devices between users while trying to configure each device to work as close to the real task as possible. We will be using controllers, vision-based hand tracking, and data gloves. For controllers, such as the Vive “wand” controller, we will have the controller act as the wand rather than as the user’s hand. For vision-based hand tracking (e.g. the Leap Motion) and the data gloves (e.g. ManusVR gloves) we will have a virtual wand that users will be able to pick up and manipulate.

Cognitive Mirrored Writing Task. During the mirrored writing task, subjects will need to trace letters or shapes with a pen but will only be able to see a mirrored version of their writing hand. This task has been demonstrated to be difficult but trainable [9]. In the real version of the task, this will be accomplished through a mirrored writing apparatus designed specifically for research studies. Performance will be measured based on how many letters or shapes are backwards in the final version. For the virtual recreation, we will use a rendering effect to mirror the subject's hand and writing. For this task we will also be using controllers, vision-based hand tracking, and data gloves. Similarly to the first task, we will treat the controllers as pens to mirror the real world task as closely as possible. For the other input methods, the user will be able to pick up a virtual pen to attempt the task.

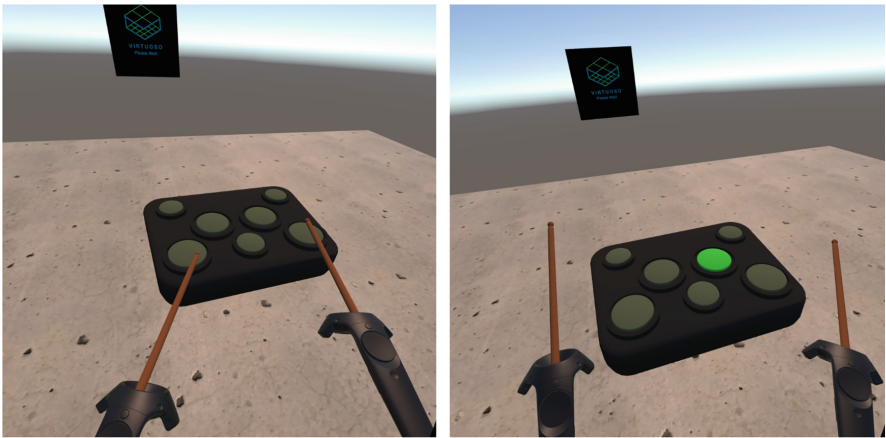


Fig. 3. Early iteration of VR drum model based on real world drum kit. These images depict a user holding the drum sticks via HTC VIVE controllers, which is one of the several interaction fidelities we plan to test. The image on the right shows a highlighted drum pad, indicating a pattern for the user to follow, learn, and rehearse.

Gross Motor Drumming Task. The third task will require the subject to memorize a drum rhythm. For the real version of the task, we will be using an electronic drum set with 7 drum pads to allow for complicated patterns. The user will be shown an animation indicating the order the pads must be activated in and will be able to listen to the pattern. The subject will be judged based on whether they hit the correct pads in the correct order and how long it took to complete the rhythmic pattern. For the virtual training session, we will recreate the drum pad and sounds, and the animation will play directly on the drum pads (Fig. 3). On top of the previous interaction methods mentioned, we will also secure trackers to both the drum sticks and the drum pad to allow for mixed reality training.

Usability Sub-tasks. In addition to the primary tasks, we will have a series of smaller XR usability tests to run each subject through; these usability tasks will function to both distract participants from the original task training and to collect data on the usability of XR interactions. This will include subjective questionnaires (i.e. likert scales) aimed at discovering whether the subject found an interaction easy to use. We have prototyped several interactions from medical and mechanical domains covering both gross and fine motor skills across a variety of XR devices. Table 1 lists some of the tasks we will use during the usability portion of the test.

Table 1. Example usability sub-tasks for interaction fidelity study.

Domain	Fine- or Gross-motor	Task	Devices
Medical	Fine	Scalpel incision	Controller, hand tracker, glove with force feedback, MR tool
Medical	Fine	Chest tube insertion	Controller, hand tracker, glove with force feedback
Medical	Gross	Check pulse	Controller, hand tracker, haptic feedback gauntlet, glove with force feedback
Medical	Gross	Tourniquet application	Controller, hand tracker, glove with force feedback
Mechanical	Fine	Bolt tightening/ loosening	Controller, hand tracker, glove with force feedback
Mechanical	Fine	Insert wire through small hole	Controller, hand tracker, glove with force feedback
Mechanical	Fine	Soldering component	Controller, hand tracker, glove with force feedback, MR device
Mechanical	Gross	Fitting large component (e.g. oil filter)	Controller, hand tracker, glove with force feedback
Mechanical	Gross	Welding	Controller, hand tracker, glove with force feedback
Mechanical	Gross	Hammering	Controller, hand tracker, glove with force feedback
Generic	Mixed	Firing a weapon	Controller, hand tracker, MR device

3.2 Impacts

We anticipate that our results will have several impacts that will influence the design of XR. First, we believe that both portions of our study will contribute insights into device selection for XR training simulations. Currently, devices are selected by either qualitative assessment by resident scientists and developers or out of convenience. Clearer evidence for which devices provide the most training impact will help demystify the process of device selection at the start of development. Second, we believe that both the insights into fidelity and the usability studies will highlight which approaches to XR interaction and scenario design create more effective training scenarios.

4 Conclusion

While our larger program addresses numerous aspects of XR in an effort to increase the realism and immersion of XR simulation and requisite training capabilities, this paper focused on two areas within this larger effort; real-time networking of different XR environments (VR and AR/MR), and a planned study on the impact of the fidelity of interactions on training of fine- and gross-motor control tasks in virtual environments as well as the usability of specific XR interactions.

We see promise in our concept for networking different XR environments, such as VR and MR (Magic Leap One), for a variety of training applications, as it can help solve the issue of VR training obscuring the trainer from the trainees' environment and interactions, and allows for specific objects and information to be kept separate between two environments. We believe we have barely scratched the surface in the space of networked multi-reality simulations and see numerous areas for future research and exploration, including reversed visualizations, remote observation, and trainer tools for manipulation of the environment and the simulation.

Additionally, we described the design of a planned study on interaction fidelity and usability of XR interactions. This study is designed to evaluate the effect of VR interaction fidelity (e.g. controller vs. data glove vs. camera-based hand tracking) on the training transfer of fine- and gross-motor tasks. This planned study will evaluate three distinct tasks; a fine-motor wire tracing task, a fine-motor and cognitive mirrored writing task, and a gross-motor drumming sequence task. Participants will complete the real world task, train on the task in VR (at varying levels of fidelity across participants), complete a series of tasks designed to evaluate the usability of XR tool-based interactions and act as distractors from the training task, then finally complete the real world task again. This study has the potential to provide interesting insights for XR developers and researchers, potentially on the optimal level of fidelity of XR interaction, selection of XR peripherals and devices, and the overall design of HCI/XR interactions.

The XR field will likely continue to grow and training applications will increase in both quantity and quality [2, 4]. Our intention is to aid the advancement of the XR field; thus our aforementioned VSDK will become open-source in 2019, supporting our networked XR capabilities and naturalistic interactions, and we will disseminate our findings from our usability study once we finalize data collection and analysis, later in 2019 as well.

Acknowledgements. Research was sponsored by the Combat Capabilities Development Command and was accomplished under Contract No. W911NF-16-C-0011. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Combat Capabilities Development Command or the US Government. The US Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the Combat Capabilities Development Command.

References

1. McMahan, R.P., Lai, C., Pal, S.K.: Interaction fidelity: the uncanny valley of virtual reality interactions. In: Lackey, S., Shumaker, R. (eds.) VAMR 2016. LNCS, vol. 9740, pp. 59–70. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-39907-2_6
2. Champney, R., Salcedo, J.N., Lackey, S.J., Serge, S., Sinagra, M.: Mixed reality training of military tasks: comparison of two approaches through reactions from subject matter experts. In: Lackey, S., Shumaker, R. (eds.) VAMR 2016. LNCS, vol. 9740, pp. 363–374. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-39907-2_35
3. de Ribaupierre, S., Kapralos, B., Haji, F., Stroulia, E., Dubrowski, A., Eagleson, R.: Healthcare training enhancement through virtual reality and serious games. In: Ma, M., Jain, L.C., Anderson, P. (eds.) Virtual, Augmented Reality and Serious Games for Healthcare 1. ISRL, vol. 68, pp. 9–27. Springer, Heidelberg (2014). https://doi.org/10.1007/978-3-642-54816-1_2
4. Cohn, J., et al.: Training evaluation of virtual environments. In: Baker, E., Dickieson, J., Wulfek, W., O’Neil, H. (eds.) Assessment of Problem Solving Using Simulations, pp. 81–106. Routledge, New York (2011)
5. Tactical Combat Casualty Care Handbook. 5th edn. Center for Army Lessons Learned, Fort Leavenworth (2017)
6. Montgomery, H.: Tactical Combat Casualty Care Quick Reference Guide. 1st edn (2017)
7. Xsens Customer Cases. <http://www.xsens.com/customer-cases>. Accessed 15 Mar 2019
8. Chan, J., Leung, H., Tang, K., Komura, T.: Immersive performance training tool using motion capture technology. In: Proceedings of the First International Conference on Immersive Telecommunications, p. 7. ICST, Brussels (2007)
9. Latash, M.: Mirror writing; learning, transfer, and implications for internal inverse models. *J. Mot. Behav.* **31**(2), 107–111 (1999)