Usability Evaluation of Augmented Reality Training for Battlefield Care Tasks

Larry A. Moralez, Claire Hughes, JoAnn Archer, Cameron Miller, and Kay Stanney

Design Interactive, Inc.

Orlando, FL

Larry.Moralez@designinterctive.net,
Claire.Hughes@designinteractive.net,
Joann.Archer@designinteractive.net,
Cameron.Miller@designinteractive.net,
Kay.Stanney@designinteractive.net

Darrin Frye, Frank Karluk, and Raymond Vazquez

Joint Program Committee 1-Medical Simulation Portfolio

Ft. Detrick, MD

darrin.l.frye.civ@mail.mil,
frank.j.karluk.ctr@mail.mil,
raymond.vazquez.ctr@mail.mil

ABSTRACT

The ability of Combat Lifesavers and Combat Medics to render efficient and effective battlefield care is often the deciding factor on whether injuries lead to death or survival. Over 80% of American fatalities with survivable wounds from Iraq and Afghanistan wars died from hemorrhage wounds that are potentially survivable and treatable when the first responder has been suitably trained. Combat Medics receive advanced training to include simulated scenarios that provide the necessary realism of battlefield stress, complexity, and dynamics. Yet, due to the prohibitive cost of medical manikins and the time and money spent conducting live training events, in addition to limited opportunities to practice under contextually representative conditions, Combat Lifesavers are often ill prepared for providing appropriate, timely battlefield trauma care for Soldiers with survivable injuries. However, Augmented Reality (AR) can couple with instructional design strategies to support tactical combat care at the Combat Lifesaver level. This paper presents a usability evaluation of a tablet-based AR training system which uses constructs associated with contextualization and embodied cognition to enhance training effectiveness. The key aim of the system is to systematically utilize these features within an AR environment such that it leads to contextualization of declarative knowledge faster than standard training methods, thereby increasing trainees' understanding of material and decreasing the overall training time required to reach a competent level. However, a crucial step in developing AR training solutions that leverage the principles of contextualization and embodied cognition to enhance training is to evaluate the usability of novel contextually-based design elements and embodied interactions afforded by AR. Careful examination of these features can highlight positive and negative experiences trainees experience in AR, possible improvements to usability, and future directions for evaluating the application of contextualization and embodied cognition principles into the design of AR training solutions for Combat Lifesavers.

Usability Evaluation of Augmented Reality Training for Battlefield Care Tasks

Larry A. Moralez, Claire Hughes, JoAnn Archer, Cameron Miller, and Kay Stanney

Design Interactive, Inc.

Orlando, FL

Larry.Moralez@designinterctive.net,
Claire.Hughes@designinteractive.net,
Joann.Archer@designinteractive.net,
Cameron.Miller@designinteractive.net,
Kay.Stanney@designinteractive.net

Darrin Frye, Frank Karluk, and Raymond Vazquez

Joint Program Committee 1-Medical Simulation Portfolio

Ft. Detrick, MD

darrin.l.frye.civ@mail.mil,
frank,j.karluk.ctr@mail.mil,
raymond.vazquez.ctr@mail.mil

INTRODUCTION

The ability of battlefield care providers to render efficient and effective battlefield care is often the deciding factor on whether injuries lead to death or survival. Studies using data from the Joint Theater Trauma Registry suggest that there are at least three potentially survivable battlefield wounds: extremity hemorrhage, tension pneumothorax, and airway obstruction. Of these, of American fatalities with survivable wounds from Iraq and Afghanistan wars died from hemorrhage alone (Lawton, Granville-Chapman, and Parker, 2009). More than 17% of combat casualties, or deaths related to combat, could have been survivable if the first responders in the field had sufficient training (Eastridge et al, 2011).

Extremity hemorrhage, tension pneumothorax, and airway obstruction are potentially survivable and treatable injuries when battlefield care providers have been suitably trained to administer care in the field. Typical battlefield care providers include Combat Medics and Combat Life Savers (CLS). Combat Medics are military personnel that have completed at minimum 16-weeks of training to provide first aid and frontline trauma care. They receive advanced training in the form of simulations to provide the necessary realism of battlefield stress, complexity, and dynamics (Fitzhugh, 2015). Continuous training and advancement of their skills is the norm, making them the definitive lifesaving resource in combat. On the other hand, CLS are intended to provide an intermediate step between the basic life support skills taught to every Soldier and Combat Medics. While Priority 1 units (i.e., actively deploying brigade combat teams) must retake the course every year, CLS typically receive only 40 hours of instruction prior to certification. Thus, they can often be ill prepared to provide appropriate and timely battlefield trauma care for Soldiers with survivable injuries due to limited opportunities to practice under the same contextually representative conditions as Combat Medics. The extent of their tactical trauma care training is inhibited by the prohibitive cost of live training, medical manikins, and a lack of anthropically diverse manikins or casualty simulators that can present with physiological realism of the types of injuries experienced on the battlefield. This training gap limits access to repeatable and representative training examples when military personnel are not deployed. Accessible, goal-specific, and guided just-in-time training for trauma resuscitation has the potential to improve survivability statistics in complex and time-sensitive trauma situations experienced by non-medic (bystander) military personnel and civilians. These first responders have few opportunities to practice under conditions representative of battlefield stress, complexity, and dynamics.

Augmented Reality (AR) and supporting technologies such as haptic feedback, 3D audio, and personalized performance assessment may offer a solution for providing high fidelity training in core battlefield life-saving competencies. Augmented reality training, in which the physical world is augmented by superimposing computer graphics and relevant data points onto trainees' natural world view, has the potential to provide contextually rich and immersive medical training solutions that substantially enhance learning (Chan et al., 2010; Lee, 2012). Notably, effective AR training solutions have been developed for medical (Barsom et al., 2016), industrial maintenance and assembly (Gavish et al., 2015), aerospace (Macchiarella et al., 2005), and education (Tekedere & Göke, 2016) applications. It follows that AR is emerging as a realistic, low-cost immersive solution with increasing potential in

improving and supporting performance of lifesaving resuscitation skills. The technology enables the embedding of high-fidelity augmentations within real-world contexts such that the training environment better matches the diversity of battlefield constraints. Contextual realism concomitant with real-time, guided instruction that better matches psychomotor responses and in field decision making may increase transfer of training of life-saving skills while under battlefield conditions.

The Department of Defense is addressing this training gap by supporting the development of a medical AR classroom tool, AUGMED (Augmented Reality for Adaptive, Effective Learning and Execution of Medical Care), that superimposes synthetic imagery onto live patients and/or medical manikins in training of realistically modeled combat casualty states. AUGMED is an adaptive, modular, immersive, and extensible AR training system geared toward CLS providers who are learning the Tactical Combat Casualty Care (TCCC) curriculum. The system leverages AR technology to bridge the gap between traditional classroom-based learning and hands-on skills training. It is designed to effectively deliver spatialized instructional content and contextualized, realistic training cues to learners. Because the system was developed to be modular and extensible, it can support integration with multiple current and future hardware platforms. Currently, it can be experienced with the Microsoft HoloLens I AR head-mounted display and AR-enabled tablet platforms. Both hardware platforms present unique technical constraints impacting overall usability and must be considered in the system design process; with respect to the HoloLens I implementation, the device has a limited Field of View (FoV), fixed depth mapping, discoloration, and burdens trainees with eyestrain and the weight of the device. The AR tablet platform has FoV restrictions that are a function of the device camera and screen size. Moreover, the tablets must often be held by two hands, limiting interactivity and increasing the likelihood that tracking is lost when trainees place the device on any surface that occludes the room view.

The key aim of this system was to implement contextualization and embodied learning within an AR environment such that it leads to contextualization of declarative knowledge faster than standard training methods, thereby increasing trainees' understanding of material and decreasing the overall training time required to reach a competent level. However, a crucial first step to developing AR training solutions that leverage principles of contextualization and embodied cognition to enhance training is to evaluate usability of novel contextualized design elements and embodied interactions afforded by AR. If the embodied interaction techniques featured in the system are too cumbersome for trainees to intuitively use, then transfer of training is likely to be diminished. Careful examination of these features can highlight positive and negative experiences in AR, possible improvements to overall usability, and future directions for evaluating the application of contextualization and embodied cognition principles into the design of AR training solutions for Combat Lifesavers.

BACKGROUND

Contextualization

Contextualization is one of the key motivating empirical and theoretical principles underlying development of the AUGMED system. Combat lifesaving is a complex, high tempo process situated in and demonstrably interwoven within the battlefield context. The battlefield is highly fluid, filled with dynamic noises (e.g., explosions, weapon fire, vehicle engine, etc.), taxing environmental factors (e.g., heat, cold, humidity, etc.), visual confusion (e.g., due to fog, smoke, sandstorms, etc.), and other such factors that present stressors that directly impact operational performance (Merlo, 2018). These contextual factors may hinder a CLS's ability to rapidly scan a scene, focus on relevant cues, develop a casualty response plan, and administer care. It is critically important to prepare CLS to perform potentially lifesaving care in a variety of challenging battlefield environments. Providing such training without considering the impact of contextual factors would be expected to impede training transfer. AR technology can provide CLS with training that is contextually relevant scenarios that require trainees to adapt to the demands of the battlefield.

Contextualized learning suggests that if trainees are embedded into relevant context, the knowledge they gain will be more likely to transfer (Bransford, & Brown, & Cocking, 2000). However, successful transfer is largely determined by the degree of mastery of initial declarative and procedural knowledge acquisition (including the "what," "when," "why," "where," and "how"), as without such learning, transfer cannot be expected (Papert, 1980). This suggests that the appropriate level of contextualization depends on proficiency level, with low levels appropriate for novices and advanced beginners, while high levels appropriate for competent, proficient, and expert learners (Dreyfus, 2004).

Novices require low levels of contextualization because they have no relevant schemas via which to reason from, thus cognitive load is intrinsically high (e.g., trainees often do not know where to look, what to do, or how to make sense of new knowledge) and there can be a lot of extraneous processing going on (Mayer, 2010). If high levels of contextualization are added to task relevant information for novice and advanced beginner AR training scenarios, it would thus be expected to hinder initial learning as it would predictably add to the burden of extraneous processing. Thus, contextualization for novices and advanced beginners may best be focused on feedback regarding the degree to which they know when, where, and how to use the information they are learning so that they appropriately conditionalize their knowledge (Bransford, 1979). This can be enhanced through use of "contrasting scenarios" (Schwartz, Lin, Brophy, & Bransford, 1999) that place focus on critical cues that trainees have yet to recognize, the latter of which can be evidenced by AR headset eye tracking. As proficiency grows and trainees reach competent level and beyond, more contextual factors can be integrated into AR scenarios. Varying contextual factors is also important, as research has demonstrated that transfer across contexts, such as dynamic battlefield contexts, is especially difficult when skills, such as CLS skills, are only taught in a single context (Bjork & Richardson-Klavhen, 1989). Thus, if trainees who have reached competent level and beyond engage in AR training solutions that require wide and varied demonstration of applicable skills (e.g., with time stress, sounds of screaming and explosions, dense smoke, etc.), they would be expected to abstract relevant features of concepts and to develop a flexible representation of knowledge that can be transferred flexibly across a multitude of battlefield contexts (Gick & Holyoak, 1983).

Embodied Cognition

Embodiment during training is another one of the key motivating principles underlying development of the AUGMED system. It draws inspiration from decades of research in Embodied Cognition (EC), the view that there is a deep continuity between the mind and the body. EC is distinct from traditional perspectives in the sciences of the mind that view the body as merely peripheral to the brain and therefore establishes a novel means of conceptualizing and investigating the nature of cognitive processing, including learning.

The main thesis of EC is that cognition is highly dependent on the features of the agent's physical body. In addition to learning, recent research in embodiment of cognition spans a rich range of phenomena, including action and perception, language, communication, memory, phenomenology, and even robotics. Some of the earliest EC work came in the field of linguistics, where Lakoff and Johnson (1980) illustrated how an extensive range of complex cognitive concepts are mapped to spatial relationships that directly reflect the nature of human anatomy. Language is thus endowed with meaning by way of experience, an interactive process that is constrained by the neural and anatomical features of our bodies, as well as properties of affordances in the distal environment. That is, meaning emerges from the interaction between an organism and its environment and not simply from the internal cognitive structures found in its central nervous system. Further complimenting this view is research investigating the nature of gesture. It is not uncommon for people to move their bodies while communicating. Though typically thought to be a means of communicating contextually relevant information to the listener, emerging research suggests such gesturing occurs at least partially to enhance cognitive performance. The Gesture as Simulated Action framework lays out the claim that gestures are the consequence of mental simulations of action and perception states that enhance cognitive mechanisms in service of reasoning skills, such as mental imagery, memory consolidation, and problem solving (Hostetter & Alibali, 2019). The research indicates that gesturing activates sensorimotor networks in the brain and in doing so recruits them to enhance thinking (Cole et al., 2002). This highlights the idea of neural reuse, whereby high-level abstract reasoning recruits the brain's sensorimotor regions responsible for perception and action rather than relying on distinct cortical regions dedicated to specific categories of reasoning (Anderson, 2014). Additional examples of EC range from the intuitive (i.e., distinct color-sensing cones in the eye constraining what colors can be detected) to the radical (i.e., explaining intelligent behavior only in terms of agent-environment dynamics), serving as a foundation in support of emerging interest in embodied learning. The implication is that, to promote learning and transfer, CLS AR training scenarios should incorporate contextually relevant gestures and other associated embodied behaviors (e.g., kneeling to apply a tourniquet) that simulate appropriate actions and perceptions associated with the application of critical CLS skills.

Embodied Learning

Embodied learning (EL) is an educational method that promotes bodily interaction during the learning process to increase learning outcomes. It contrasts starkly with traditional instruction principles that follow mentalistic theories separating body and mind. Instead of sitting, watching, listening, and writing, learners are encouraged to interact with their learning environments, whether it be through interactive exploration or enacting gestures.

Early attempts to integrate the body in the learning process come from research in 2nd language acquisition. What began as an instructional method involving learners acting out the commands given to them in the 2nd language, has since evolved into the use of representative gestures to clarify difficult concepts and increase vocabulary memorization efficacy (Asher & Price, 1967; Larsen-Freeman and Anderson, 2013; Quinn-Allen, 1995). Later investigations into the influence of gesture on learning mathematics showed that learners better comprehended problem-solving strategies if they performed a gesture that matched the strategy compared to a no gesturing condition. What's more, performing a mismatching gesture caused students to comprehend the strategies worse than a no gesture condition (Goldin-Meadows et al., 1999). Thus, should AR training solutions require non-contextually valid gestures, such as pinching and dragging menu items, transfer performance could be hindered. Continuing down this line of inquiry, Crollen and Noël (2015) asked learners to carry out hand or feet movements during a counting task. Both hand and feet movements during counting led to disruption in counting compared to the control condition, but the hand movement condition had the largest impact. This suggests that the process of counting and hand motion may rely on a shared limited neural resource. In a follow-up fMRI study, researchers found increased hemodynamic activity in the contralateral motor cortex when numbers were presented, indicating a deep continuity between sensorimotor substrates and those recruited during higher-order cognitive processes, such as those associated with combat casualty care (Sato & Lalain, 2008).

Embodiment theories have since developed into an emerging paradigm in virtual training and simulation studies. In these studies, researchers leverage the underlying principles of EC and EL to build instructional content in virtual space to augment the learning process. In one case, researchers placed learners in a room where virtual objects were projected onto surfaces (Embodied Mixed Reality Learning Environment, or EMRELE) and could be interacted with. They found that a high embodied learning condition that encouraged students to interact with and explore their environment increased long-term retention of generative knowledge related to centripetal force (Johnson-Glenberg et al., 2016). Additional EMRELE studies found that learning gains for both chemistry titration and disease transmission were greater when compared to traditional classroom training (Johnson-Glenberg, 2014). Similarly, when middle school students were given the opportunity to learn about gravity and planetary motion in an immersive, whole-body interactive simulation, they demonstrated significant learning gains, increased attention, and more positive attitudes towards science (Lindgren et al., 2016). The implication is that CLS AR training scenarios that incorporate requisite embodiment of critical skills should lead to greater learning gains than those associated with traditional classroom training alone.

AR TRAINING SYSTEM

Augmented Reality Tablet

An AR tablet application was designed for use at U.S. Army Medical Simulation Training Centers (MSTCs). The initial AUGMED Tablet design facilitated a desktop-based learning experience that could be carried out within the physical classrooms at the MSTCs. This was made possible by presenting a miniaturized wounded Soldier on a desk in front of a trainee, while walking the trainee through an instructional scenario. Although the experience was viscerally interesting to end-users, early observational data indicated that it didn't provide any specific learning value beyond that provided by traditional computer-based training (CBT) systems, as the design did not incorporate significant levels of contextualization or embodiment. After initial, informal usability observations, the team took a step back and revisited the very important question of "Why AR?".

Initial observations indicated that the AR 3D casualty model used in AUGMED Tablet was viscerally engaging, but only to an extent. It became clear that the miniaturized casualty models did not afford any significant gross motor movement from trainees. In fact, participants could easily just sit at a desk and complete the AR training experience.

The act of sitting at a desk while completing training is something that can be more easily carried out through CBT training scenarios, without the potential for added frustrations and quirks associated with immersive AR technology. After analyzing early prototype use, the team quickly realized the differentiating factors between traditional CBT and AR-based training: contextualization and embodiment. An AR training experience that incorporates a spatially aware device can afford, via contextually relevant cues, appropriate gross motor function from the trainee. This powerful trait quickly became the crux of a rapid design iteration.

During redesign, the team reflected on the initial desktop-based design along with the context of use. The team assessed the general context of use of the MSTC site and determined that the desktop-based constraint was an unnecessary limiting factor in the design of the AUGMED Tablet experience. At this point, the team rapidly iterated to a life-size scale virtual casualty manikin and redesigned instructional activities as to better facilitate contextualized and embodied learning. This was the answer to the question of "Why AR?".

Observations showed that by introducing a life-size casualty in a contextually rich AR training environment, trainees were much more likely to move around the casualty with the tablet to explore symptoms and crouch down to assess injuries, learn more about training cues, and perform virtual interventions with the tablet. These are contextually anchored gross motor functions that are performed in the field and can now be facilitated through a training scenario that only requires an AR-enabled tablet and AUGMED software. Although the tablet presents limitations regarding training of granular motor function due its handed form factor, the iterative design process for the AUGMED Tablet training experience has resulted in a much more engaging, contextually relevant, and embodied training experience than what a seated, desktop-based solution could provide. In this more immersive design, novice and advanced beginners are physically engaged in knowing when (proper sequence), where (proper location), and how (proper physical embodiment) to apply critical care so that they appropriately conditionalize the CLS knowledge they are learning. For competent trainees, additional contextualization is incorporated, including time stress and contextually relevant sounds (e.g., explosions, weapons fire), and more intricate embodiment is required (e.g., prompting the user to physically respond to spatially varied dynamic visual and auditory cues).

Head Worn Display

The Head Worn Display (HWD)-based training compliments the AUGMED Tablet training experience. The Microsoft HoloLens I was the HWD selected for use due to availability and synergy with the U.S. Army's Integrated Visual Augmentation System (IVAS), which is based on the HoloLens platform. Unlike the tablet-based training experience, the HWD-based training was initially designed as a room-scale, life-size experience where trainees could view critical training cues overlaid onto a physical manikin. Early observations of HWD-based training uncovered usability issues relating to contextualized and embodied interaction design. These observations, paired with data gathered from the tablet-based experience, resulted in formalization of interaction design patterns for embodied learning, or Spatial Learning Activity Patterns (SLAPs). A total of seven SLAPs were developed. Each SLAP has been carefully crafted to afford a variety of contextually rich and embodied formative assessments within a room-scale 3D training experience. The SLAPs include the following:

1. **Discovery**: The trainee is presented with a virtual casualty that features various points of interest are highlighted through spatial iconography. Trainees are prompted to physically move close to areas of interest to learn more information. As a trainee approaches an area of interest, contextually relevant visual and auditory information is revealed.
2. **Locate and Select**: Spatialized iconography is appended to the casualty and trainee is prompted to select one or more areas of interest. This requires the trainee to physically look at a specified area and pay it visual attention before interacting with it either gesturally or verbally.
3. **Object Matching**: The trainee is presented with a 3D shelf of various objects. Objects may take the form of 3D models, images, or labels. The trainee is prompted to "pick up" an object from the shelf and physically match its associated location in space per the training content.
4. **Question and Answer**: Audiovisual questions are presented to the trainee for a quick assessment. The questions prompt the trainee to respond, either verbally or gesturally, based on what they are perceiving in 3D space.
5. **Action Plan** The action plan includes a list of items that need to be sorted into a correct sequence. The action plan can be presented to a trainee in an incomplete or partially complete state depending on proficiency level. After the action plan is completed, the trainee can carry out the plan of action verbally or using gestures.

6. **Timed Activity**: Time to task completion is captured through the "timed activity" SLAP. This is a simple interaction design that introduces contextually relevant visual and auditory time stress while capturing trainee performance.

The team was been able to increase the levels of contextualization and embodiment in the HWD-based training experience by formalizing a specified and deliberate set of contextualized and embodied interaction design patterns. The usability study reported here focused on further improving the usability of each SLAP.

METHODOLOGY

Participants and Materials

Fourteen (14) individuals (n = 14; 7 = Male and 7 = Female; Age Range=26-35) were recruited to participate in a usability evaluation of the AUGMED HWD and Tablet systems. All had normal or corrected-to-normal vision. In the HWD group, participants were equipped with a Microsoft HoloLens I AR headset. In the Tablet group, participants held an AR Samsung Galaxy Tab S5E.

The System-Usability Scale (SUS) survey was administered following training sessions for each AR device. The SUS is an industry-standard research instrument comprised of 10 questions, each with five (5) response options ranging from Strongly Disagree to Strongly Agree, which taken together provide a high-level view of overall system usability (Sauro, 2011). The Multidimensional User Experience Questionnaire (MUEQ) was used to measure additional features of the AUGMED training system that expand beyond basic SUS assessment (Bevan et al., 2016). The MUEQ provides data spanning multiple dimensions of user experience, such as Instruction & Support, Visual Design, Navigation, Interactivity, and Motivation to learn. The Single Ease Questionnaire (SEQ) was also used, which is designed to gauge ease-of-use of specific interactions within a system. The goal of the SEQ rating is to identify specific difficulties at key interaction points in a system so they may be remedied in future design iterations.

Procedures

Two researchers, a moderator and a scribe, were present for each usability testing session. Once a participant arrived at the testing site, he/she was greeted by the researchers. The moderator then narrated a welcome script before handing the participant an AR device. A within-subjects design was employed, with counterbalanced HWD and tablet experiences. The HWD display streamed a first person's view of the training experience to an external laptop display, which was monitored by the researchers during the evaluation. Table 1 below lists each of the training experiences featured in the tablet and HWD versions of the system. During the experience, participants followed a think-aloud protocol by verbalizing their thoughts about the system, while the scribe transcribed their feedback. Throughout each participant experience, the moderator asked the Single Ease Question (i.e., overall, this task was [1] very difficult to [7] very easy) after key interaction events to support granular documentation of specific pain points within the system. After the participant completed the AUGMED experience, they completed the SUS questionnaires and provided additional verbal insight about their immersive experience.

Table 1: AUGMED Training Experiences

Tablet	Head-worn Display
Virtual Patient Placement	Massive Hemorrhage Cue Matching
Chest Wound Symptom Exploration	Individual First-Aid Kit Item Selection
Open Chest Wound Symptom Identification	Virtual Tourniquet Placement
Open Chest Wound Identification	
Tension Pneumothorax Symptom Exploration	
Intercostal Space Exploration	
CLS Bag Item Selection	

Virtual Needle-Catheter Insertion	

RESULTS

System Usability Scale

Figure 1 shows SUS results for the AR HWD evaluation. The first five questions are positively coded such that agreement with them indicates ease of system use. The right skewness of responses to these questions suggests that participants would like to use the AUGMED HWD training system, found the system well integrated, and felt comfortable using the system. The last five questions are negatively coded, such that their agreement would indicate difficulty using the system. The left skewness of these questions further suggest that participants had a favorable view of the AUGMED HWD system's usability, as they did not find the system unnecessarily complex or cumbersome to use, did not perceive inconsistencies in the system design, and did not feel the system required them to learn a lot before they started training. The AUGMED HWD training system's total SUS was 62.5. The average SUS is 68 (Sauro, 2011), which suggests the HWD system is scoring slightly below average, as compared to desktop and mobile software, the latter of which features nearly universal standardized interaction (WIMP) techniques that do not yet exist in AR. Usability for AR HWD will likely increase when a standardized interaction paradigm has been realized, as the SUS results indicate that 50% of participants were neutral or in agreement with the statement "I think that I would need assistance to be able to use this system." This indicates that learnability of the system should be a focus of future design efforts.
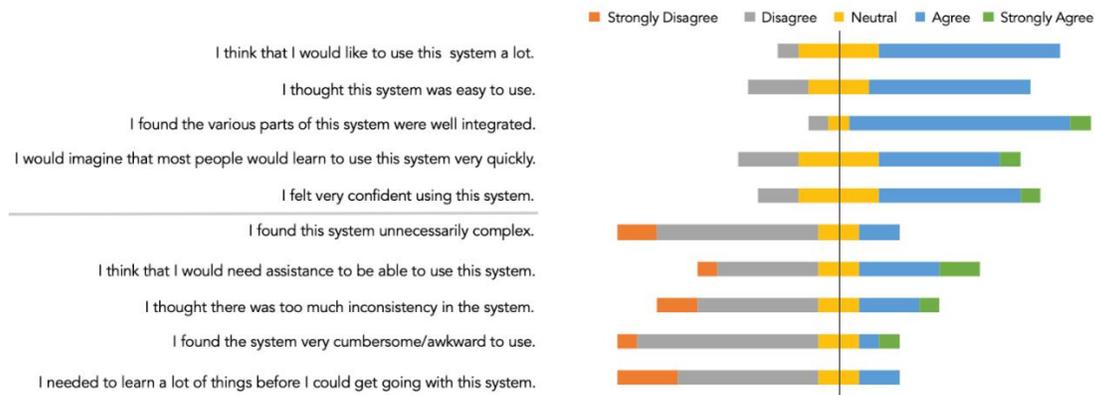
Figure 1. AUGMED HWD System Usability Scale Results

The AUGMED Tablet training system's total SUS was 76.4, scoring well above average for more traditional desktop and mobile applications. However, approximately 36% of participants were neutral or within agreeance with the statement "I think that I would need assistance to be able to use this system" (see Figure 2), indicating that learnability may be a worthwhile aspect to focus on in future design efforts.
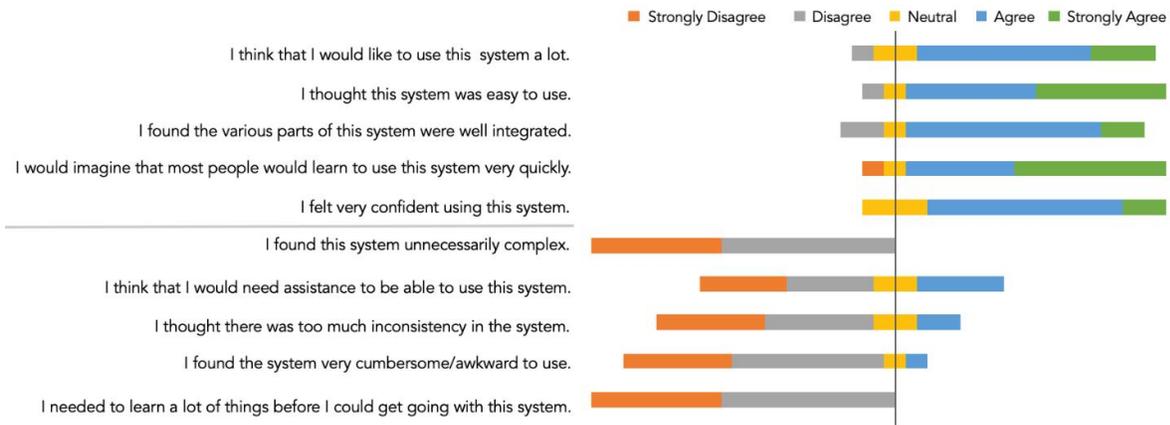
Figure 2. AUGMED Tablet System Usability Scale Results

Multi-dimensional System Usability Scale

ISO standards for establishing functional suitability of a prototype product following a user-centered design define learnability and operability as necessary evaluative criteria (Bevan et. al, 2016). The MUEQ was used to evaluate these criteria. No negatively coded questions are associated with the MUEQ, therefore right skewness of the answers always indicates favorable views of the evaluated components of the system. Apart from the statements, "The organization of information on the system display was clear," where 50% of participants were neutral or in disagreement, and "I always knew where my attention should be directed," where more than 50% of participants were in disagreement, participants' favorably viewed of the HWD system's learnability and operability (see Figure 3).
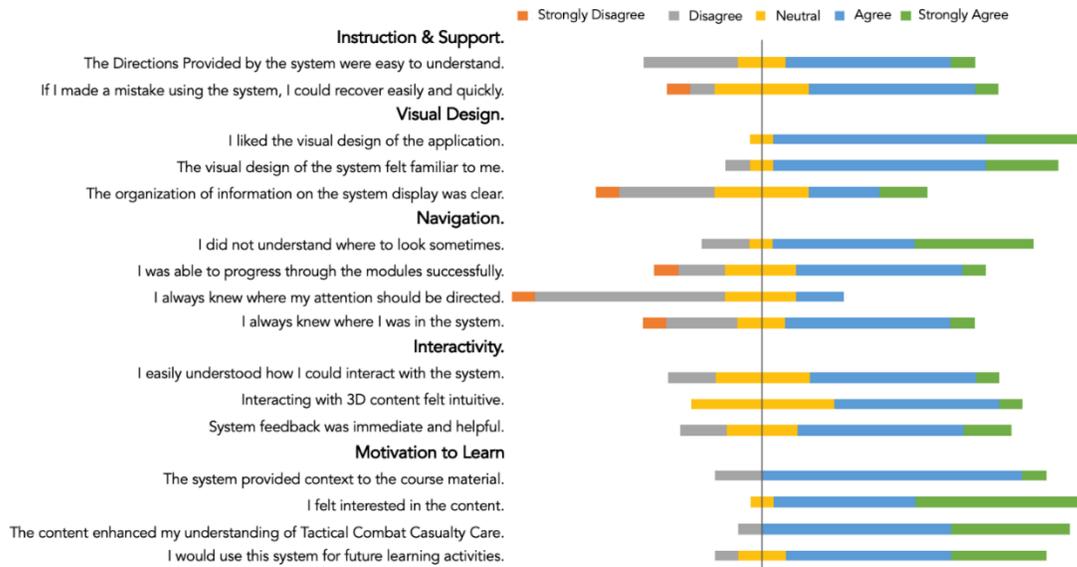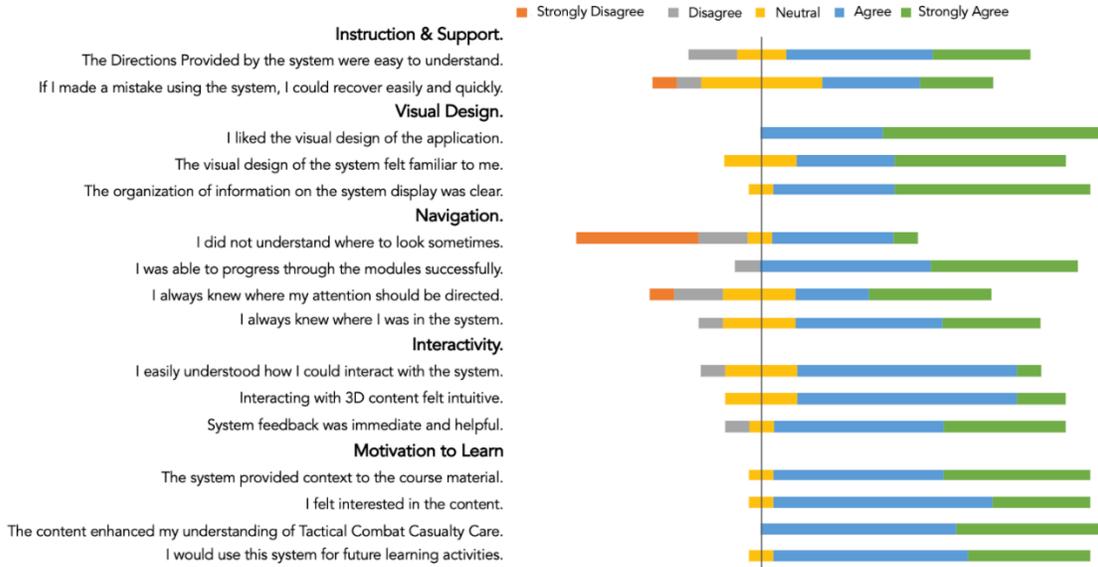

Figure 3. AUGMED HWD Multidimensional System Usability Scale Results

Figure 4 reveals that apart from the statements, "If I made a mistake using the system, I could recover easily and quickly," "I always knew where my attention should be directed," where at least one participant was within strong disagreement, and "I did not understand where to look sometimes," where 50% of participants were neutral or within disagreement, the participants' overall view of the AR Tablet system was favorable.
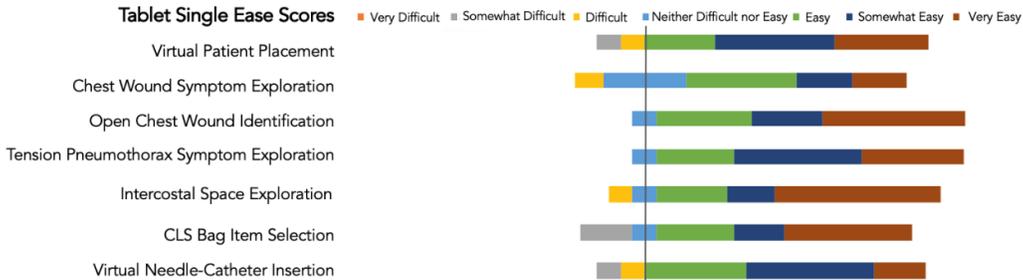
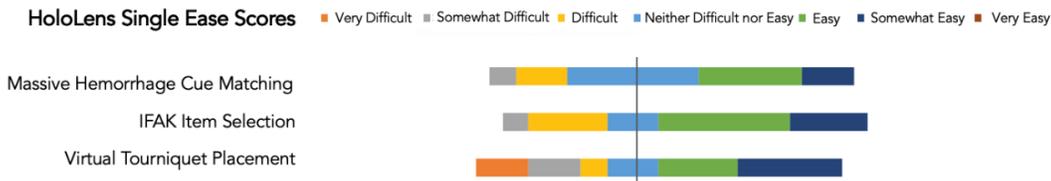**Figure 4. AUGMED Tablet Multidimensional System Usability Scale Results**

SEQ

Whereas the SUS and MUEQ measures give general overviews of ease-of-use and effectiveness of the AUGMED system, SEQ results provide a finer grained assessment of specific events in the training lesson that may have led to reported scores in the previous questionnaires. Figure 5 highlights that in the tablet version, the main tasks incorporated into participants experience were heavily skewed towards ease-of-use. None of the participants rated any of the tasks as Very Difficult, whereas at least two or more rated each task as Very Easy. Most responses fell in the Easy to Very Easy range. The responses suggest that AR related training tasks within the system were intuitive and easy-to-use. Nonetheless, a handful of Somewhat Difficult to Difficult responses indicate that there remains room for improvement.

Figure 5. AUGMED Tablet Multidimensional System Usability Scale Results

Figure 6 highlights that in the HWD build, the main tasks for the participant experience were only slightly skewed towards easiness. However, no participant reported their interactions to be Very Easy for any of the three tasks, with some participants indicating that it was Very Difficult to place a virtual tourniquet on the casualty's leg.

Figure 6. AUGMED HWD Multidimensional System Usability Scale Results

DISCUSSION AND CONCLUDING THOUGHTS

The purpose of this study was to evaluate the usability of an AR training system that was developed in alignment with the principles of contextualized and embodied learning. A usability study was planned prior to conducting a training effectiveness evaluation to ensure that the contextualized and embodied interaction techniques featured were designed in such a manner as to not impede usability, which could hinder training effectiveness. Overall SUS scores ranged from slightly below average (HWD) to above average (tablet), suggesting that improvements need to be made to increase the HWD's usability to levels higher than the average desktop and mobile software solutions. Fixing key issues is likely to significantly increase the HWD system's usability. However, it is possible that the increased number of training experiences in the Tablet compared to the HWD instantiation of the system led to the discrepancies in performance. Future studies can control for this variation in training content to allow for a more accurate comparison of usability.

AUGMED is a framework designed to be used with multiple technologies. For this study, it was implemented within both a HWD, which has a limited FoV, fixed depth mapping, and discoloration of visual assets, and a tablet device. As can be seen in the MUEQ results for the HWD (see Figure 3), the lowest rated issue for this implementation of AUGMED was that participants were unaware of where to look and direct their attention. This is a problem that does not present itself as pertinently when implemented on other devices that do not have as extensive FoV restrictions, such as a tablet. This indicates that this shortcoming is not inherent to the AUGMED system itself, but to the specific implementation of the system on the HWD device. Such technological limitations are likely to be diminished in future HWD technologies. Nonetheless, designing a system that properly guides trainees' attention to system elements outside of their FoV is a crucial consideration when developing embodied interaction techniques for such systems. Of noted value may be implementing spatial audio and lighting cues to direct attention such that trainees know where to look and how to direct their attention during task performance. Moreover, the HoloLens I gesture recognition capabilities are currently limited to two hand-based movements: the bloom gesture, whereby the user brings all of their fingers together and then opens their hand, and the air tap gesture, which the user can interact with virtual affordances by looking at them and bringing their pointer finger up and down in a tapping motion. Unlike the tablet version of the system that employs many standardized mobile interaction techniques, the user is not likely to be familiar with or skilled at performing the bloom and air tap gestures used in the HWD. Using novel and non-standardized interaction techniques likely made the system more difficult to use for users without previous experience using the HWD. Indeed, in many cases the user needed to perform the air trap gesture multiple times because they did not perform it in a manner that the HWD could recognize. Future iterations of the AUGMED system can focus on development of novel gesture recognition features that allow the user to not only move around and interact in a fully immersive environment, but to also perform contextually relevant gestures specific to the training scenario that may lead to increased training outcomes similar to those observed in research supporting embodied learning.

Overall, both system implementations performed admirably in most dimensions of the MUEQ (see Figures 3 and 4). As the system is an extensible AR training solution, it is crucial that the instructional content is delivered to trainees in a comprehensible manner. Responses to the "Instruction and Support" and "Motivation to Learn" dimensions indicate that the system is successful in this regard. This was true for both the Tablet and HWD versions of the system, further validating the system's overall efficacy as a training solution. Participants indicated that the system was easy to interact with. Thus, when they appropriately directed their attention and the system recognized their gestures, carrying out necessary tasks was not difficult.

Properly trained battlefield medical care providers are critical to the prevention of fatalities, motivating a strong push to develop more efficient and cost-effective training solutions. AR technology stands to add immersive medical training that could confer a significant learning advantage for CLS and beyond. What's more, AR is well-positioned to leverage the principles of contextualized and embodied learning to enhance its effectiveness. However, in order to properly assess the effectiveness of contextualized and embodied AR training, it is important to first ensure that such interaction techniques do not impede trainees' ability to advance in their training. Several such contextualized and embodied interaction techniques, across both HWD and Tablet devices, were evaluated and shown to be viewed favorably by participants, with the main caveat being the need to aid trainees in directing their attention. Future work, with aforementioned design issues remedied, will compare the training effectiveness of the AUGMED system to a traditional classroom setting to measure its impact on training knowledge retention and transfer.

References

Anderson, M. L. (2014). After phrenology: Neural reuse and the interactive brain. MIT Press.

Asher, J. J., and Price, B. S. (1967). The learning strategy of the total physical response: some age differences. Child Dev. 38, 1219–1227.

Barsom, E. Z., Graafland, M., & Schijven, M. P. (2016). Systematic review on the effectiveness of augmented reality applications in medical training. Surgical endoscopy, 30(10), 4174-4183.

Bevan, N., Carter, J., Earthy, J., Geis, T., & Harker, S. (2016, July). New ISO standards for usability, usability reports and usability measures. In International Conference on Human-Computer Interaction (pp. 268-278). Springer, Cham.

Bjork, R.A., & Richardson-Klavhen, A. (1989). On the puzzling relationship between environment context and human memory. In C. Izawa (Ed.), *Current issues in cognitive processes: The Tulane Flowerree Symposium on Cognition* (pp. 313-344). Hillsdale, NJ: Erlbaum.

Bransford, J.D. (1979). *Human cognition: learning, understanding, and remembering*. Belmont, CA: Wadsworth.

Bransford, J.D., Brown, A.L., & Cocking, R.R. (2000). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academy Press.

Chan, C. L., Ngai, E. K., Leung, P. K., & Wong, S. (2010). Effect of the adapted virtual reality cognitive training program among Chinese older adults with chronic schizophrenia: A pilot study. International Journal of Geriatric Psychiatry, 25(6), 643-649.

Cole, J., Gallagher, S., & McNeill, D. (2002). Gesture following deafferentation: A phenomenologically informed experimental study. Phenomenology and the Cognitive Sciences, 1(1), 49-67.

Dreyfus, S. E. (2004). The five-stage model of adult skill acquisition. *Bulletin of Science, Technology & Society*, *24*(3), 177-181.

Eastridge, B. J., Mabry, R. L., Seguin, P., Cantrell, J., Tops, T., Uribe, P., ... & Butler, F. K. (2012). Death on the battlefield (2001–2011): implications for the future of combat casualty care. Journal of trauma and acute care surgery, 73(6), S431-S437.

Fitzhugh, D. C. (2015). Technology Solutions for Combat Casualty Care. Newsletter HPT&B, 3, 3-5.

Gavish, N., Gutiérrez, T., Webel, S., Rodríguez, J., Peveri, M., Bockholt, U., & Tecchia, F. (2015). Evaluating virtual reality and augmented reality training for industrial maintenance and assembly tasks. Interactive Learning Environments, 23(6), 778-798.

Gick, M.L., & Holyoak, K.J. (1980). Analogical problem solving. *Cognitive Psychology*, 12, 306–355.

Johnson-Glenberg, M. C., Birchfield, D. A., Tolentino, L., & Koziupa, T. (2014). Collaborative embodied learning in mixed reality motion-capture environments: Two science studies. Journal of Educational Psychology, 106(1), 86.

Johnson-Glenberg, M. C., Megowan-Romanowicz, C., Birchfield, D. A., & Savio-Ramos, C. (2016). Effects of embodied learning and digital platform on the retention of physics content: Centripetal force. Frontiers in psychology, 7, 1819.

Goldin-Meadows, S., Kim, S., and Singer, M. (1999). What the teacher's hands tell the student's mind about math. J. Educ. Psychol. 91, 720–730.

Hostetter, A. B., & Alibali, M. W. (2019). Gesture as simulated action: Revisiting the framework. Psychonomic bulletin & review, 26(3), 721-752.

Lakoff, G., & Johnson, M. (1980). Metaphors we live by. Chicago, IL: University of Chicago.

Larsen-Freeman, D., and Anderson, M. (2013). Techniques and Principles in Language Teaching, 3rd Edn. Oxford: Oxford University Press.

Lawton, G., Granville-Chapman, J., & Parker, P. (2009). Novel hemostatic dressings. Journal of the Royal Army Medical Corps, 155(4), 309-314.

Lee, K. (2012). Augmented reality in education and training. TechTrends, 56(2), 13-21.

Lindgren, R., Tscholl, M., Wang, S., & Johnson, E. (2016). Enhancing learning and engagement through embodied interaction within a mixed reality simulation. Computers & Education, 95, 174-187.

Macchiarella, N. D., Liu, D., Gangadharan, S. N., Vincenzi, D. A., & Majoros, A. E. (2005). Augmented reality as a training medium for aviation/aerospace application. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 49, No. 25, pp. 2174-2178).

Mayer, R. E. (2010). Applying the science of learning to medical education. *Medical Education*, *44*, 543–549.

Merlo, J.L. (2018). Military operations: Humans not machines make the difference. In R. Steinberg, S. Kornguth, and M.D. Matthews (Eds.), *Neurocognitive and physiological factors during high-tempo operations* (pp. 195-207). London: Routledge.

Papert, S. (1980). *Mindstorms: Computers, children, and powerful ideas.* New York: Basic Books.

Quinn-Allen, L. (1995). The effects of emblematic gestures on the development and access of mental representations of French expressions. Modern Lang. J. 79, 521–529.

Sato, M., and Lalain, M. (2008). On the relationship between handedness and hand-digit mapping in finger counting. Cortex 44, 393–399.

Sauro, J. (2011). Measuring System Usability with the System Usability Scale (SUS). Retrieved from https://measuringu.com/sus/

Schwartz, D.L., Lin, X., Brophy, S., & Bransford, J.D. (1999). Toward the development of flexibly adaptive instructional designs. In C.M. Reigelut (Ed.), *Instructional Design Theories and Models: Volume II,* (pp. 183–213). Hillsdale, NJ: Erlbaum.

Tekedere, H., & Göke, H. (2016). Examining the Effectiveness of Augmented Reality Applications in Education: A Meta-Analysis. International Journal of Environmental and Science Education, 11(16), 9469-9481.