Technical Report 1307

How Simulator Interfaces Affect Transfer of Training: Comparing Wearable and Desktop Systems

John S. Barnett
U.S. Army Research Institute

Grant S. Taylor
University of Central Florida
Consortium Research Fellows Program

June 2012

United States Army Research Institute for the Behavioral and Social Sciences

U.S. Army Research Institute
for the Behavioral and Social Sciences

Department of the Army
Deputy Chief of Staff, G1

Authorized and approved for distribution:

BARBARA A. BLACK, Ph.D.	**MICHELLE SAMS, Ph.D.**
Research Program Manager	**Director**
Training and Leader Development Division	

Technical Review by

Robert J. Pleban, U.S. Army Research Institute
Cheryl Johnson, U.S. Army Research Institute

NOTICES

DISTRIBUTION: Primary distribution of this Technical Report has been made by ARI. Please address correspondence concerning distribution of reports to: U.S. Army Research Institute for the Behavioral and Social Sciences, ATTN: DAPE-ARI-ZXM, 6000 6th Street (Bldg 1464 / Mail Stop: 5610), Ft. Belvoir, Virginia 22060.

FINAL DISPOSITION: This Technical Report may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The findings in this Technical Report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

REPORT DOCUMENTATION PAGE

1. REPORT DATE (dd-mm-yy) June 2012	2. REPORT TYPE Final	3. DATES COVERED (from. . . to) March 2010 - September 2011

4. TITLE AND SUBTITLE How Simulator Interfaces Affect Transfer of Training: Comparing Wearable and Desktop Systems	5a. CONTRACT OR GRANT NUMBER 5b. PROGRAM ELEMENT NUMBER 655803
6. AUTHOR(S) John S. Barnett (U.S. Army Research Institute), and Grant S. Taylor (University of Central Florida)	5c. PROJECT NUMBER D 730 5d. TASK NUMBER 5e. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences 6000 6th Street (Bldg. 1464 / Mail Stop: 5610) Fort Belvoir, VA 22060	8. PERFORMING ORGANIZATION REPORT NUMBER Technical Report 1307
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences 6000 6th Street (Bldg. 1464 / Mail Stop: 5610) Fort Belvoir, VA 22060	10. MONITOR ACRONYM ARI 11. MONITOR REPORT NUMBER Technical Report 1307

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

Subject Matter POC: John Barnett

14. ABSTRACT *(Maximum 200 words)*:

The present research compared the training effectiveness of a wearable computer interface versus a desktop interface. Teams of two participants were trained in military hostage rescue procedures, one using the wearable interface, a second group using a desktop interface, while a control group was trained in a live physical room using replica weapons and military equipment. Teams completed four training sessions, followed by four test sessions in the live room. Teams were scored on number of correct actions and time to complete the session. The control condition performed significantly better on the test scenarios than either of the simulator conditions. However, there were no significant differences in performance between the wearable or desktop simulation conditions. The number of correct actions for both simulator conditions were statistically equivalent, and slightly lower than the control condition. There was also no significant difference between the two simulator training groups in terms of time to complete the test scenarios. Both simulator conditions were significantly slower to complete the scenarios than the control condition, with the magnitude of this difference diminishing over time.

15. SUBJECT TERMS

transfer of training, wearable computer, simulator training, virtual environments

SECURITY CLASSIFICATION OF			19. LIMITATION OF ABSTRACT	20. NUMBER OF PAGES	21. RESPONSIBLE PERSON (Name and Telephone Number)
16. REPORT Unclassified	17. ABSTRACT Unclassified	18. THIS PAGE Unclassified	Unlimited	26	Peggy Maloney 703-545-2325

Technical Report 1307

How Simulator Interfaces Affect Transfer of Training: Comparing Wearable and Desktop Systems

John S. Barnett
U.S. Army Research Institute

Grant S. Taylor
University of Central Florida
Consortium Research Fellows Program

Orlando Research Unit
Joan H. Johnston, Chief

U.S. Army Research Institute for the Behavioral and Social Sciences
6000 6th Street, Bldg. 1464
Fort Belvoir, VA 22060

June 2012

Army Project Number	**Personnel Performance**
655803D730	**and Training**

Approved for public release: distribution is unlimited

HOW SIMULATOR INTERFACES AFFECT TRANSFER OF TRAINING: COMPARING WEARABLE AND DESKTOP SYSTEMS

EXECUTIVE SUMMARY

Research Requirement:

The U.S. Army is considering the use of wearable computers as interfaces for dismounted infantry simulation training. The wearable interfaces are expected to give Soldiers a greater sense of immersion (being there) and lead to better training. However, there is no empirical evidence to verify that the wearable interface provides better training than other methods. The present research is the latest in a series of experiments testing the training effectiveness of the wearable computer interface.

Procedure:

Teams of two participants with no prior military training were trained in military hostage rescue procedures. One group trained using the wearable interface and the Game-based Distributed Interactive Simulation (GDIS) dismounted infantry simulation, a second group got the same training using a desktop interface with GDIS, while the third, a control group, was trained in a live physical room in an office building using replica weapons and military equipment. Teams completed four training sessions in their respective training conditions, followed by four test sessions in the live room to assess their ability to transfer their training to a realistic environment.

Findings:

Not surprisingly, the control condition, which trained using live rooms, performed significantly better on the test scenarios than either of the simulator conditions. However, there were no significant differences in performance between the wearable or desktop simulation conditions. The number of correct actions taken in the test sessions by both simulator conditions was statistically equivalent, and slightly lower than the control condition. There was also no significant difference between the two simulator training groups in terms of time to complete the test scenarios. Both simulator conditions were significantly slower to complete the scenarios than the control condition, with the magnitude of this difference diminishing over time. The only other significant difference was that the wearable was rated higher in symptoms of simulator sickness than the desktop or live conditions, with no difference between the desktop and live conditions.

Utilization and Dissemination of Findings:

The findings of this research can be used to guide decisions about which interfaces are most effective for training dismounted infantry skills using game-based simulators. Discussion of how best to incorporate simulation into the training process is also provided.

These findings were presented to TRADOC Capabilities Manager, Virtual on 28 February 2012.

HOW SIMULATOR INTERFACES AFFECT TRANSFER OF TRAINING: COMPARING WEARABLE AND DESKTOP SYSTEMS

CONTENTS

LIST OF FIGURES

Page

HOW SIMULATOR INTERFACES AFFECT TRANSFER OF TRAINING: COMPARING WEARABLE AND DESKTOP SYSTEMS

Introduction

The use of simulators for training can, in certain domains, improve performance substantially, but at considerably lower costs than similar training in the field. The effectiveness of simulator training depends on a number of factors, including the types of skills to be trained, the capabilities of the simulator to facilitate correct practice of skills, and the usability and appropriateness of the simulator interface.

Often, new advances in technology will promise improvements in the training effectiveness of simulators. Faster computers allow more elements to be introduced into the simulation, and novel interfaces that improve the realism of the simulation are expected to provide better training. However, although improvements in simulator technology are often expected to provide better training, in practice they may provide no better training, or even worse training, than less complex systems. When training in the field, there are many elements in the field environment that have no practical effect on training. Therefore, adding those elements to the simulator environment does not improve training effectiveness.

The best way to determine the effectiveness of novel elements in a simulation is to test those elements through research and user testing. By evaluating how well the novel simulator elements train novices to learn skills compared to other simulators and field training, researchers can determine the training effectiveness of the novel simulator elements.

The current project was designed to test the training effectiveness of a novel simulator interface based on a wearable computer. The ExpeditionDI™ wearable interface was expected to provide a better training experience in a simulation environment than using a desktop interface, since it was theorized to provide a more immersive environment than a desktop simulation. To test this theory, participants were trained to conduct a hostage rescue task using the wearable interface, a desktop interface, or in a live room as a control. After training, each participant was evaluated on their ability to transfer the training to the performance of realistic hostage rescue missions in the live room. Group scores were then compared to determine the relative training transfer of the wearable and desktop simulators compared to training in the live room.

This research is the second in a series of experiments investigating the training transfer of wearable and desktop simulators. The first experiment (Taylor & Barnett, 2011) measured participant's recall of procedures trained in wearable and desktop simulators. Following the experiment, questions arose as to whether cognitive recall of what is essentially procedural knowledge could be considered training transfer. The current research was conducted to clarify the previous results. In this second experiment, the dependent measure was a physical demonstration of learned procedures, requiring participants to transfer procedural knowledge learned in the training scenarios to a live environment.

The following sections will include a brief review of relevant research, a description of the experimental methodology, results of the statistical analyses conducted, and a discussion of the conclusions found from the research.

Background

Virtual environments can be effective training tools for military tasks. For example, Pleban, Eakin, Salter, and Matthews (2001) used a virtual environment to train decision-making skills. In addition, virtual environments have been shown to provide effective dismounted infantry training (Pleban & Salvetti, 2003).

Knerr (2007) conducted a review that analyzed the need for, and expected benefits of, dismounted Soldier training in virtual environments. One of the recommendations of this review was to evaluate the cost effectiveness of fully immersive simulators compared to desktop simulators for dismounted infantry training. Based on the recommendations in this review, the TRADOC Capability Manager, Virtual Training Environment (TCM Virtual) requested that ARI conduct research to determine the relative advantages and cost effectiveness of wearable and desktop interface simulators for dismounted infantry training.

Therefore, an experiment was conducted which compared how well military tasks were trained using a wearable simulator interface and a common desktop computer interface, with the U.S. Army's Interactive Multimedia Instructional videos currently in use as a control (Taylor & Barnett, 2011). This experiment did not find any significant difference in the participant's ability to recall correct and incorrect steps for the military tasks, regardless of the training condition.

The dependent measure for the previous experiment was having the participants view videos of avatars performing military tasks and asking participants to describe both the correct and incorrect actions demonstrated in the videos. It was reasoned that this method of measurement assessed participants' memory learning, but possibly not the type of procedural knowledge necessary to perform the tasks in the field. Therefore, despite these findings, it was possible that the use of wearable and desktop interfaces may prove beneficial for the training of procedural skills.

For this reason, a second experiment was conducted which trained procedural skills and evaluated the transfer of this training to a realistic performance environment. The goal of this experiment was to determine whether any differences in learning occurred based on using a wearable interface, desktop interface, or traditional live training. This report presents the results of this research effort.

A discussion of previous research relevant to this series of experiments can be found in Taylor and Barnett (2011).

Method

Participants

A total of 62 participants completed the research, with 20 in each of the Desktop and Wearable training conditions, and 22 in the Live condition. To match the Army's restrictions for Soldiers conducting hostage rescue missions (the task to be trained), all participants were males between 18-30 years old (M = 20.27, SD = 2.128) and in good health. All participants were verified to have no prior military or ROTC experience to ensure they had no previous training on hostage rescue tasks.

Apparatus

Questionnaires. The Simulator Sickness Questionnaire (SSQ) (Kennedy, Lane, Berbaum, & Lilienthal, 1993) is a 16-item questionnaire that measures three separate dimensions of simulator sickness: Nausea (e.g., increased salivation, stomach awareness), Oculomotor (e.g., eyestrain, difficulty focusing), and Disorientation (e.g., dizziness, vertigo). The questionnaire has participants rate their current experience of symptoms on a four-point scale ranging from "None" to "Severe." Participants completed this questionnaire both before and after their exposure to a simulated environment, with the changing level of each of the three subscales used to determine the impact of the simulation on their physiological state.

Following the SSQ, participants completed the Gaming Experience Measure (GEM) (Taylor, Singer, & Jerome, 2009) and Game Performance Assessment Battery (GamePAB) (Taylor et al., 2009). The GEM is a questionnaire designed to measure the participant's experience with and knowledge of video games separately, consisting of 35 self-report items (e.g., average hours of play per week, experience with various genres, and experience with various types of controllers) and 21 multiple-choice test items (e.g., questioning what system a specific game was released for, or what is used to perform a specific function in a game). Experience is rated on a 1 – 5 scale, with larger values indicating greater experience, and knowledge is rated on a scale from 0 – 100 based on the percentage of correct responses, with higher values indicating greater knowledge. GamePAB is a measure of the participant's video game skill, specifically within the first-person shooter genre. Skill is measured through multiple performance metrics while the participant completes a series of tasks within a virtual environment, with metrics including reaction time and time-on-target.

The Interest/Enjoyment and Perceived Competence scales of the Intrinsic Motivation Inventory (IMI) (McAuley, Duncan, & Tammen, 1987) were collected following the training session. These two scales consist of a total of 13 statements (e.g., "I enjoyed doing this activity very much", "I was pretty skilled at this activity"), with participants rating their agreement with each on a 7-point Likert scale. These responses were averaged to form the independent measures of Interest/Enjoyment and Perceived Competence, each ranging from 1 – 7 with higher values indicating greater Interest/Enjoyment or Perceived Competence.

Desktop Simulation. Those participants assigned to the desktop simulation condition were trained on the procedures using a standard desktop computer system, using a keyboard and mouse to control a virtual Soldier. The desktop computers used were Dell XPS systems, with a 2.66 GHz Intel Core 2 Duo CPU, 4 GB of RAM, an NVIDIA GeForce 8800 GTX graphics card, and a 20" LCD monitor with a 16:10 aspect ratio. GDIS, the software used for these scenarios, was designed specifically for military training, and is similar to many popular first-person shooter video games (e.g., Modern Warfare, Half Life, Virtual Battlespace 2, etc.). The controls were typical of most first-person shooter video games, using the W, S, A, and D keys to move the virtual avatar forward, back, left, and right, and the mouse to look/aim and shoot. The GDIS display provided a first-person view of the environment (see Figure 1).

Figure 1. Participant's view within the GDIS virtual environment.

Wearable Simulation. Those participants assigned to the wearable simulation condition were trained on the procedures using an ExpeditionDI immersive wearable interface (by Quantum3D), using a combination of their own natural body movements and buttons on a simulated assault rifle to control their virtual avatar (Figure 2). The total system (including vest, helmet, and weapon) weighed approximately 25 pounds, with the majority of this weight distributed across the load-bearing vest. GDIS, the same virtual environment used with the desktop system, was used with the wearable system as well. The virtual environment was presented to the participant through a head-mounted display, with the participant's head movements controlling the orientation of their avatar within the environment, resulting in a natural control scheme. Similarly, the participant's posture was tracked through a gyroscopic sensor attached to their thigh, so that when the participant crouched their avatar crouched as well. The simulated weapon was the basis for the remaining controls. The weapon itself was tracked through space to control the position and aim of the virtual avatar's weapon, with the participant pulling the trigger on the simulated weapon to fire the virtual weapon. The front handgrip on the weapon contained a small joystick that the participant operated with their thumb to control the locomotion (walking, running) of their avatar. The front handgrip also contained four buttons, which served various functions such as cycling through different weapons and opening doors.

Figure 2. The ExpeditionDI wearable simulation system.

Live Environment. Those assigned to the live environment condition were trained on the hostage rescue tactics in real rooms, with life-size cardboard cutouts as enemies and hostages (Figure 3). They were provided with a replica M4 rifle, which was an unloaded Airsoft™ rifle designed to shoot plastic pellets, as well as replica frag grenades and flashbangs. Participants also wore an ammo vest to carry the grenades, and a helmet and goggles for safety (Figure 4).

Figure 3. Room used for all live scenarios (enemy/hostage targets and locations varied for each scenario). Pictured: hostage (left), and enemy targets (center and right).

Figure 4. A participant in the live condition, holding the Airsoft replica M4 rifle and wearing vest with frag grenade (left) and flashbang (right).

Procedure

Participants completed the research in groups of two. Upon arrival, both participants reviewed and signed an informed consent form and then completed a series of initial questionnaires on a desktop computer. These questionnaires began with a standard demographics form used to confirm that the participant's gender and age met the research requirements, and that they had no prior hostage rescue experience. Participants then completed a baseline measure of the SSQ (Kennedy et al., 1993). Following the SSQ, participants completed the GEM (Taylor et al., 2009) and GamePAB (Taylor et al., 2009) to measure their video game experience and skill.

Following the questionnaires, the researcher trained the participants on the proper military hostage rescue techniques for roughly 20 minutes within one of three randomly assigned training conditions (desktop simulation, wearable simulation, or live environment; see descriptions below), with both participants working together as a team within the same training environment. The techniques were similar to those used by the U.S. Army and came from an Army field manual (Department of the Army, 2002), and Soldier Training Publications (Department of the Army 2004a, 2004b). These techniques described the proper way to enter a potentially hostile room, the paths to take once inside the room, and how to respond to enemy targets. Participants were scored on 22 individual task steps. The missions required the participants to work together as a team. Most task steps were consistent for both team members, but each team member did have some specific responsibilities. Each participant was randomly assigned to one team role (#1 or #2) before training began, and maintained this role throughout training and testing.

Regardless of condition, the training consisted of four practice missions. For the first mission, the researcher walked the participants through each step of the mission, explaining the important task components along the way. For the remaining three training missions, the researcher observed as the participant teams completed the missions on their own. Following each mission the researcher provided feedback describing the correct and incorrect steps taken by the members of the team. Participants typically reached near-perfect performance by the fourth training scenario.

After completing the training missions, the participants completed the SSQ again, as well as the Interest/Enjoyment and Perceived Competence scales of the IMI (McAuley, Duncan, & Tammen, 1987).

After the questionnaires, all participants completed a testing phase in which they conducted four missions in live rooms under the same conditions as described for the live practice scenarios. Their performance was videotaped to be scored later on their ability to correctly execute the procedures covered in the training, with no additional feedback provided from the researcher. Following this testing phase the research was complete, the entire experiment lasted two hours.

Results

The three training conditions were initially compared in terms of performance on the test scenarios. In addition to the percentage of actions performed correctly, scenario completion time was also used as a dependent variable due to the critical importance of speed in the hostage rescue missions. The analysis was conducted using a mixed-model ANOVA with training condition (between subjects: Desktop, Wearable, or Live) and scenario number (within subjects: 1^{st}, 2^{nd}, 3^{rd}, or 4^{th}) as independent variables.

Training condition was found to have a significant main effect on percent correct [$F(2, 69) = 4.399$, $p = .017$; Figure 5]. Pairwise comparisons determined that the Live training condition performed significantly better ($M = 96.7\%$, $SD = 2.00$) than both Desktop ($M = 93.1\%$, $SD = 3.80$, $p = .010$, $d = 1.23$) and Wearable ($M = 93.5\%$, $SD = 6.29$, $p = .018$, $d = 0.717$) training conditions, with no significant difference between the Desktop and Wearable conditions ($p = .826$). The main effect for scenario number, as well as the training condition x scenario number interaction, was not found to be statistically significant ($p > .05$ in each case).

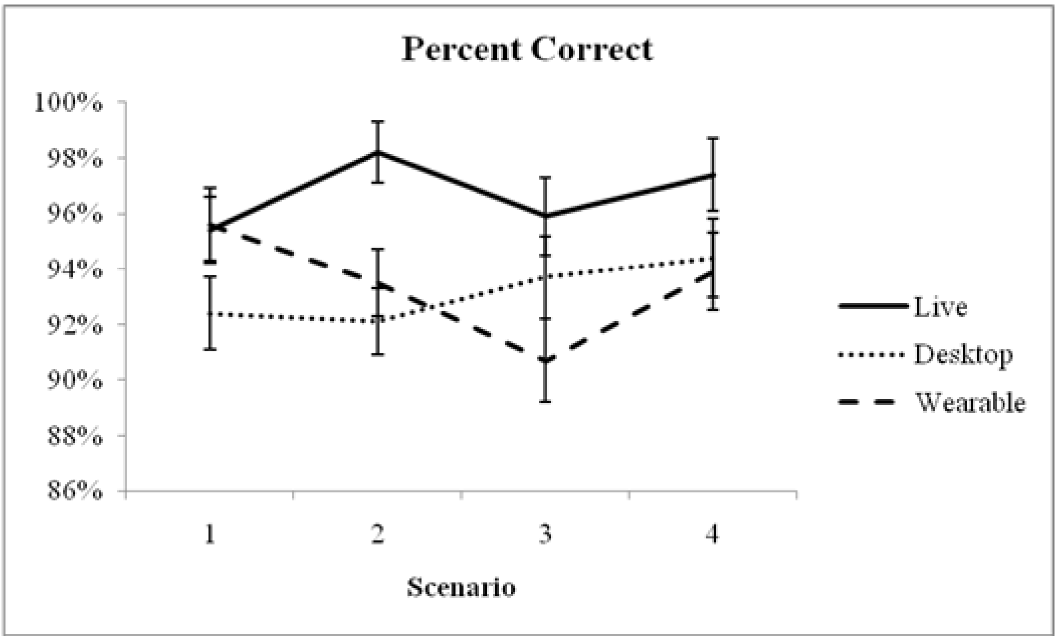

Figure 5. Percent of actions performed correctly on the four test scenarios.

Training condition also had a significant main effect on scenario completion time [$F(2, 69) = 25.056$, $p < .001$; Figure 6]. Pairwise comparisons found the Live training condition to perform the scenarios significantly faster ($M = 27.41$s, $SD = 3.48$) than both the Desktop ($M = 35.24$s, $SD = 4.74$, $p < .001$, $d = 1.94$) and Wearable ($M = 33.54$s, $SD = 2.96$, $p < .001$, $d = 1.94$) training conditions, with no significant difference between the Desktop and Wearable conditions ($p = .161$).

The interaction between training condition and scenario number was also statistically significant [$F(6, 177) = 4.319$, $p < .001$]. Subsequent one-way ANOVAs evaluated the effect of training condition on completion time of each scenario individually. These analyses found the Live training condition performed significantly faster than both the Desktop and Wearable conditions across all four scenarios ($p < .05$ in each case). However, the strength of this effect diminished over time, as exhibited by repeated-measures ANOVAs conducted on each condition. The Live training condition's performance times remained consistent across all four scenarios ($p = .481$), while a significant main effect for scenario number was found for both the Desktop [$F(3, 57) = 11.344$, $p < .001$] and Wearable [$F(3, 57) = 10.382$, $p < .001$] training conditions' performance times. Post-hoc comparisons determined that scenario completion time improved from the first scenario (Desktop: $M = 40.10$s, $SD = 5.69$; Wearable: $M = 38.56$s, $SD = 7.28$) to the last for both groups (Desktop: *M* = 30.85s, *SD* = 6.44, *p* < .001, *d* = 1.56; Wearable: *M* = 30.36, *SD* = 3.42, *p* < .001, *d* = 1.48).

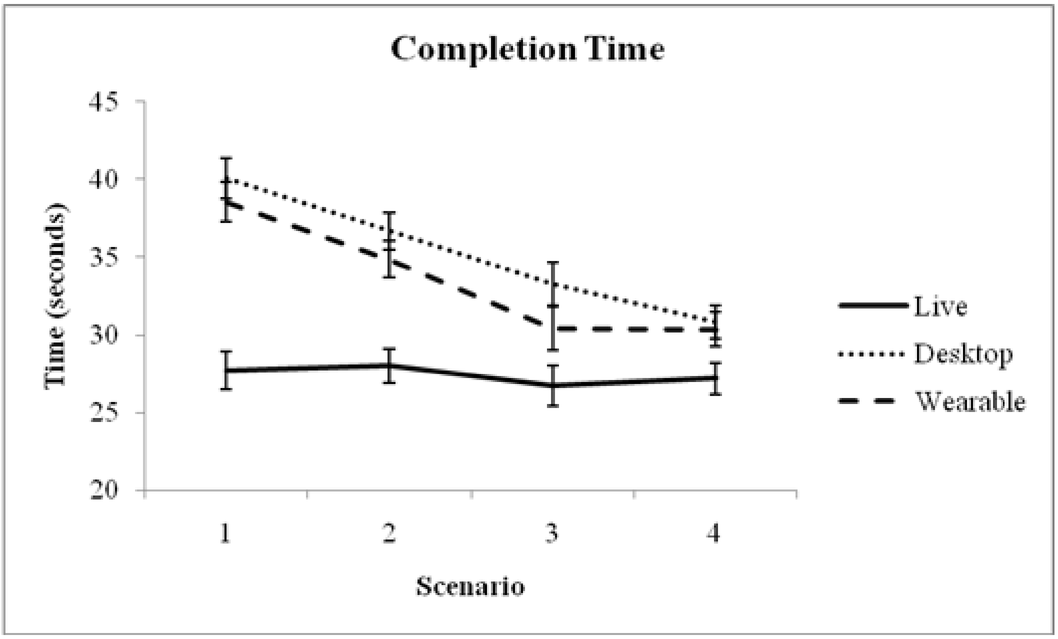

Figure 6. Scenario completion time for the four test scenarios by training condition.

The effect of training condition was also evaluated on the subjective ratings of simulator sickness and intrinsic motivation. For simulator sickness, each of the three subscales provided by the SSQ were obtained both before and after training, with the change scores for each subscale used as the dependent variable in a series of one-way ANOVAs with training condition as the independent variable (Figure 7). A significant main effect for training condition was found for the Nausea subscale [$F(2, 59) = 7.640$, $p = .001$], with the Wearable condition reporting significantly higher values (*M* = 18.60, *SD* = 30.71) than both the Desktop (*M* = -0.477, *SD* = 3.76, *p* = .001, *d* = 0.895) and Live training conditions (*M* = 0.000, *SD* = 4.16, *p* = .001, *d* = 0.891), with no significant difference between the Desktop and Live conditions (*p* = .931). The same trend was found for the Oculomotor subscale [$F(2, 59) = 13.192$, $p < .001$], with the Wearable condition reporting significantly higher values (*M* = 23.50, *SD* = 29.30) than both the Desktop (*M* = 0.379, *SD* = 1.69, *p* < .001, *d* = 1.14) and Live training conditions (*M* = 0.000, *SD* = 2.34, *p* < .001, *d* = 1.19) with no significant difference between the Desktop and Live conditions (*p* = .942). This trend was also present for the Disorientation subscale [$F(2, 59) = 4.144$, $p = .021$], with the Wearable condition reporting significantly higher values (*M* = 21.58, *SD* = 49.78) than both the Desktop (*M* = 0.000, *SD* = 0.000, *p* = .020, *d* = 0.629) and Live training conditions (*M* = -1.27, *SD* = 5.94, *p* = .012, *d* = 0.677) with no significant difference between the Desktop and Live conditions (*p* = .886).

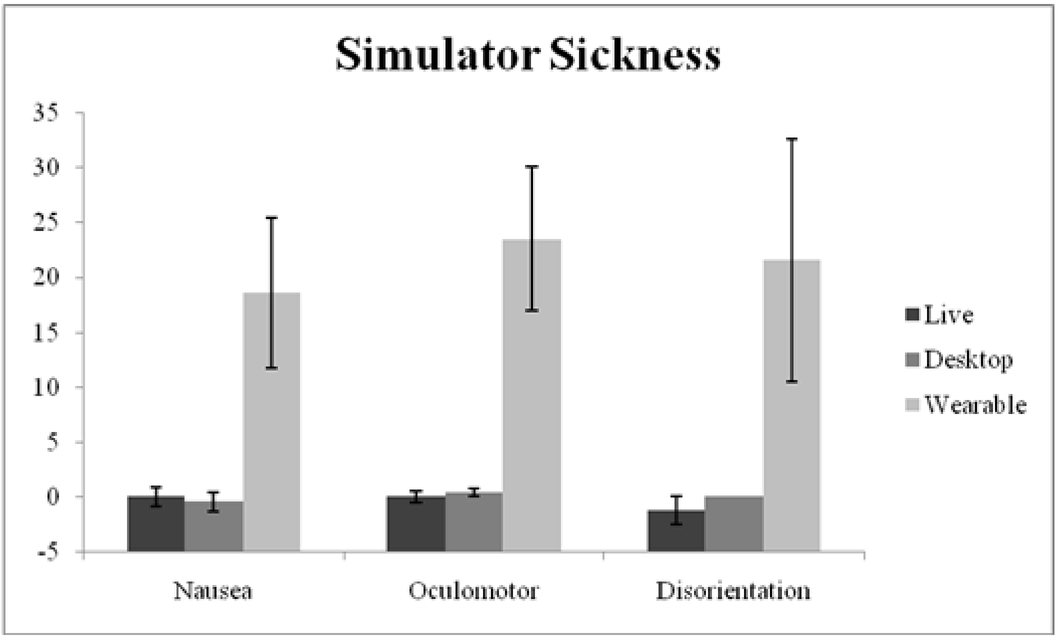

Figure 7. Simulator sickness values reported from each training condition.
Note: Values are reported as change from the baseline data collected prior to training, with positive values indicating an increase.

The effect of training condition was also evaluated on both the Interest/Enjoyment and Perceived Competence scales of the IMI (Figure 8). A significant main effect of training condition was found for Interest/Enjoyment [$F(2, 59) = 11.021$, $p < .001$]. Post-hoc comparisons determined that the Live condition reported significantly higher values ($M = 5.84$, $SD = 0.463$) than both the Desktop ($M = 5.10$, $SD = 0.632$, $p < .001$, $d = 1.38$) and Wearable conditions ($M = 5.26$, $SD = 0.510$, $p = .001$, $d = 1.22$), with no significant difference between the Desktop and Wearable conditions ($p = .338$). Training condition was also found to have a significant effect on the Perceived Competence scale [$F(2, 59) = 6.657$, $p = .002$]. Post-hoc comparisons again found that the Live condition reported significantly higher values ($M = 5.42$, $SD = 0.593$) than both the Desktop ($M = 4.81$, $SD = 0.831$, $p = .004$, $d = 0.873$) and Wearable conditions ($M = 4.77$, $SD = 0.517$, $p = .002$, $d = 1.19$), with no significant difference between the Desktop and Wearable conditions ($p = .842$).

The influence of video game experience and skill (measured by GEM and GamePAB, respectively) on mission performance was also evaluated using standard Pearson correlations. A significant relationship was found between video game experience and scenario completion time [$r(62) = -.332$, $p = .008$], with those higher in experience performing the missions faster. A regression determined that this relationship did not vary as a function of training condition ($p = .743$). No significant relationship was found between video game experience and percent correct, or the measures of video game skill with either percent correct or scenario completion time ($p > .05$ in each case).

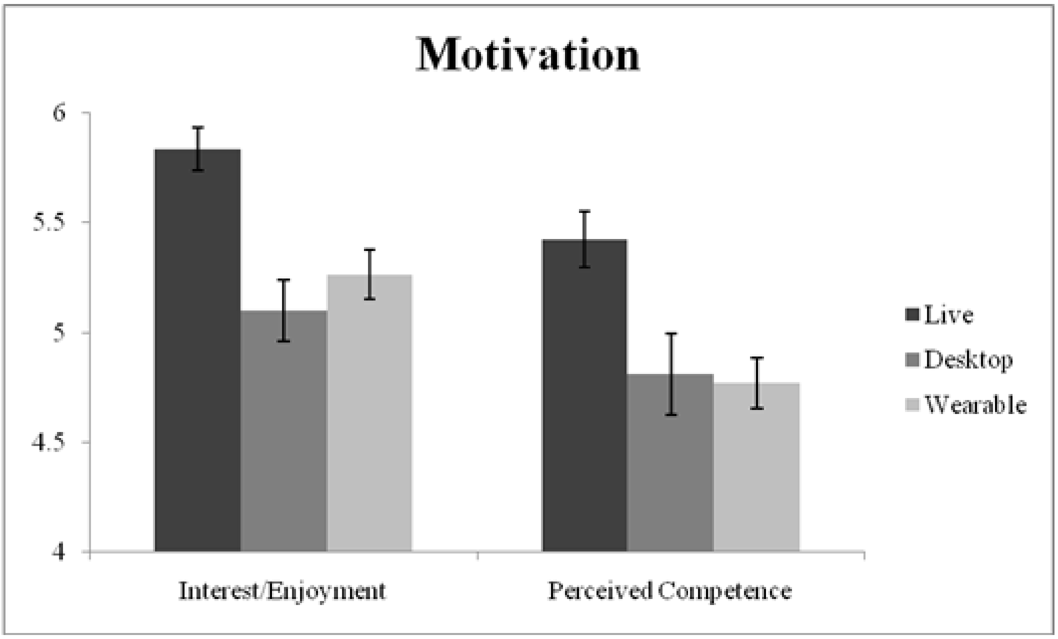

Figure 8. The Interest/Enjoyment and Perceived Competence subscales of the Intrinsic Motivation Inventory as reported from each training condition.

Discussion

Live Training

One not-particularly-surprising finding is that live training is superior to simulations for the learning of procedural skills. The results for both the percentage of actions performed correctly and the time to complete the scenarios showed live training to be superior to both simulation interfaces.

However, one possible confounding variable is that the live training condition trained in a similar environment (only slightly modified) in which their performance was tested. The live training group had the advantage of not having to transfer their knowledge to a new environment during the testing phase. Therefore, they were more familiar with the surroundings, which likely improved both their speed and performance accuracy. As participants trained in the desktop and wearable simulators completed the four test missions in the live environment, their time scores improved, whereas the live control group's time scores stayed about the same (see Figure 6). This suggests that, as they became familiar with the live testing environment, the simulator groups were able to perform more quickly, though performance accuracy remained consistent.

An alternate explanation is that the control group learned both procedural and psychomotor skills, whereas the simulation groups only learned procedural skills. The time improvement for the simulator groups may have been because they were learning the psychomotor skills required for performance. The control group's improved time scores may have resulted from training in the same room they were tested, rather than improved psychomotor skills. This question may be a productive topic for future research.

Simulator Training

The results also found there to be no significant differences between wearable and desktop interfaces, with the exception of simulator sickness symptoms. Participants who used the wearable interface rated it as inducing significantly stronger symptoms of simulator sickness than either the desktop or live training conditions. Although neither simulator condition trained as well as the live condition, both simulator conditions trained the procedural skills equally well.

The theory behind the wearable interface is that Soldiers would learn better if actions in the simulation were more natural and closer to those required in real life. The wearable allows Soldiers to turn to face different directions, look up and down, kneel, and aim and shoot their weapons using natural actions that are mirrored by their avatar in the simulated environment. However, the results of this experiment suggest that being able to perform these movements seems to have little influence on learning procedural skills. Results from the previous experiment (Taylor & Barnett, 2011) indicated that the features of the wearable interface also have little influence on learning cognitive skills, at least no more than the desktop interface or training videos. If the assumption that the improvement in mission completion time was due to the simulator groups learning psychomotor skills, then it is clear the wearable simulator's use of natural movements does not transfer to live performance.

Previous research on the usability of the wearable interface (Barnett & Taylor, 2010) indicated there were also elements of the wearable that were inconsistent with natural actions. Therefore, although the wearable simulator does allow for the use of some natural actions, other unnatural actions could negate the potential benefits of the natural actions, possibly even leading to negative training. However, the results of the present research suggest the non-natural actions do not influence training enough to provide negative training, either.

However, simulator training in general does seem to provide adequate training for procedural skills. The performance accuracy was high across all training conditions, averaging 93% to 97% depending on condition, indicating that all conditions provided acceptable training. The trend in time-to-complete for both simulator interfaces showed participants took less time to complete the live scenario each time it was performed. Although speculative, all groups would have had equivalent completion times on the fifth scenario if the trend had continued.

Gaming Skill

Although both the wearable and desktop interfaces, as well as the simulation environment, were based on game engines and used gaming conventions, prior experience playing computer games did not help or hinder learning via the game-based simulation environment. Participants seemed to learn the simulator interfaces fairly quickly, and were able to operate them well enough to learn the target skills, regardless of video game experience.

Simulations as Part of a Training Program

Previous research (Barnett, Singer & Taylor, 2010) showed that leaders saw simulations as good lead-ins for field training. Company and platoon leaders said they would like a week of simulation training just prior to engaging in field exercises. The simulation training would allow their units to practice procedural skills, which could then be mastered during field training. Simulators have much less "overhead" than field training. Simulators do not require transport to the training area, setting up equipment, maintaining vehicles, etc., thus almost all of the training time is spent practicing and improving skills (Barnett et al., 2010). Simulation may allow Soldiers to practice procedural skills to expertise, though probably not psychomotor skills. Soldiers can learn procedural skills in the simulator, and then transition to field exercises, where they can practice psychomotor skills to expertise. This is a more efficient use of training time, as valuable field time would not be needed to practice skills that can be mastered via simulation (i.e., cognitive and procedural skills). This would save time and effort in field exercises and allow all training time in the field to go towards practicing those skills that require field training to master.

Comparing Wearable and Desktop Simulation Interfaces

This experiment validates prior research (Taylor & Barnett, 2011) which found no training advantage from the use of a wearable simulator over a desktop computer when training basic field maneuvers. The current experiment tested the wearable and desktop interfaces for their ability to train more complex procedural skills and again found no appreciable difference between the wearable and desktop in their training capabilities.

However, two factors of wearable systems make them less suitable for training. First, the wearable interface is typically much more costly (by a factor of at least 10x) than even high-end desktop or laptop computers. Second, the experiment showed the wearable simulator to significantly increase symptoms of simulator sickness more than either desktop or live training conditions. Therefore, when training a large number of Soldiers, the costs and training time loss due to simulator sickness make the wearable a less attractive interface.

There were also some minor inconveniences associated with the wearable interface. There was some training time loss due to having to fit the wearable interface to the participant, taking approximately 10-15 minutes to get a proper and safe fit for each participant. Previous research (Barnett & Taylor, 2010) also determined that the wearable controls were more difficult to learn and use than the desktop interface. The wearable interface was also less stable than the desktop, consequently there was training time lost due to the need to reinitialize or reboot the system when it locked up or failed. However, as the technology matures, it may become more stable.

Although previous research (Taylor & Barnett, 2011) indicated that the wearable interface provided more motivation for the training than the desktop interface, the current experiment did not replicate this finding. This inconsistency may indicate that the motivation experienced as a result of using the wearable interface may vary based on the content of the training itself. Therefore, the motivational aspect of the wearable system is not reliable enough to use to justify its use in training.

Recommendations

Although training in a live environment identical to that found in the field results in superior performance to training in simulations, it has associated costs in terms of time, facilities, overhead, and safety. There are also some tasks that can be practiced in the simulator that cannot be practiced in the field, such as using demolitions, destroying buildings and facilities, and using rare or expensive equipment like aircraft or specialized vehicles. Also, kinetic scenarios in a simulation environment can have more realistic elements than field training, such as having enemies die when shot – something that clearly can’t be done in field training.

On the other hand, there are some skills that cannot be trained efficiently in a simulator, most notably psychomotor skills that require learning the "feel" of things. Therefore, a game-based simulator cannot teach a Soldier how to change a rifle magazine faster, or throw a grenade more accurately, or perform a correct parachute-landing fall. These skills require practicing the physical movements to master them.

Thus, both simulations and field training have advantages and disadvantages. Luckily, in many cases they are complementary, that is, the advantages of one make up for the disadvantages of the other. Therefore, each mode of training can be used to train different skills. Simulations are effective for training cognitive and procedural skills, like doctrine, tactics, and procedures, because they can be practiced in the simulation environment repeatedly at relatively low cost. Conversely, psychomotor skills requiring repeated practice of physical movements to master can be practiced during field exercises.

Most military training utilizes the "crawl, walk, run" model of skill acquisition. In this model, the "crawl" stage involves learning the basic terminology associated with the task and memorizing the procedural steps in the task. Next, in the "walk" stage, trainees practice the task steps until they can be performed in the correct order with no steps omitted. Finally, in the "run" stage, trainees further practice until the task becomes nearly automatic and can be performed quickly without errors. Following this model, the use of simulations are most appropriate for the "walk" phase, preparing Soldiers to take better advantage of field training for the "run" phase where skills can be polished and learned to mastery.

Based on the results of this, and previous research, it seems clear that desktop computer interfaces are ideal for this simulation-based training. Given their ubiquity, relatively low cost, and the degree of familiarity most Soldiers will likely have with them, desktop computers are the most accessible means of providing simulated training to Soldiers. The wearable interfaces are more expensive, more difficult to deploy (requiring additional technical expertise from training providers), require additional setup time as well as time for Soldiers to learn to use the system before they can use it to conduct training exercises, and carry an increased risk of simulator sickness. The disadvantages of the wearable interface might be acceptable if it was capable of providing training that is more effective than the desktop interface, as it is difficult to quantify the ability to transfer potentially life-saving information. However, repeated evaluation has found the wearable interface to provide training that is no better than the desktop interface.

References

Barnett, J. S., Singer, M. J., & Taylor, G. S. (2010). *Game-Based Simulators for Platoon-Level Training: What Coalition Soldiers Want.* Poster presented at the Army Science Conference, Orlando, Florida.

Barnett, J. S. & Taylor, G. S. (2010). *Usability of wearable and desktop game-based simulations: A heuristic evaluation* (ARI Study Note 2010-01). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Department of the Army (2002). *Combined arms operations in urban terrain*. Field Manual 3-06.11. Available from https://atiam.train.army.mil/soldierPortal/

Department of the Army (2004a). *Task: Perform movement techniques during an urban environment. Soldier training publication 071-326-0541*. Available from https://atiam.train.army.mil/soldierPortal/

Department of the Army (2004b). *Task: Select hasty firing positions during an urban environment. Soldier training publication 071-326-0557*. Available from https://atiam.train.army.mil/soldierPortal/

Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator Sickness Questionnaire: An enhanced method for quantifying simulator sickness. *International Journal of Aviation Psychology*, *3*(3), 203-220.

Knerr, B.W. (2007). *Immersive simulation training for the dismounted Soldier* (Study Report 2007-01). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

McAuley, E., Duncan, T., & Tammen, V. V. (1987). Psychometric properties of the Intrinsic Motivation Inventory in a competitive sport setting: A confirmatory factor analysis. *Research Quarterly for Exercise and Sport, 60,* 48-58.

Pleban, R. J., Eaken, D. E., Salter, M. S. & Matthews, M. D. (2001). *Training and assessment of decision-making skills in virtual environments* (ARI Research Report 1767). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Pleban, R.J. & Salvetti, J. (2003). *Using virtual environments for conducting small unit dismounted mission rehearsals* (ARI Research Report 1806). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Taylor, G. S. & Barnett, J. S. (2011). *Training capabilities of desktop and wearable simulator interfaces* (ARI Technical Report 1299). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Taylor, G., Singer, M., & Jerome, C. (2009). *The validity and reliability of the Game Experience Measure (GEM) and Game Performance Assessment Battery (GamePAB).* Human Factors and Ergonomics Society Annual Meeting, San Antonio, TX.

Acronyms

ANOVA	Analysis of Variance
GamePAB	Game Performance Assessment Battery
GDIS	Game-based Distributed Interactive Simulation
GEM	Game Experience Measure
IMI	Intrinsic Motivation Inventory
ROTC	Reserve Officer Training Corps
SSQ	Simulator Sickness Questionnaire
TRADOC	Training and Doctrine Command.
TCM Virtual	TRADOC Capability Manager, Virtual Training Environment
VBS2	Virtual Battle Space 2