Comparing Training Transfer of Simulators: Desktop versus Wearable Interfaces

John S. Barnett
U.S. Army Research Institute for the Behavioral and Social Sciences
Orlando, FL
john.barnett1@us.army.mil

Grant S. Taylor
University of Central Florida: Institute for Simulation and Training
Orlando, FL
gtaylor@ist.ucf.edu

ABSTRACT

The use of simulators for training provides advantages over training in the field, but often at the cost of reduced realism. New ideas in interface design promise to reduce this potential disadvantage of simulation-based training while maintaining its benefits. One such design is the use of a wearable computer in which the simulator interface is embedded into a Soldier's load-bearing equipment, allowing the user's natural body movements to become inputs into the simulator environment. This type of interface may be more immersive and have advantages over traditional desktop interfaces. This research seeks to identify training benefits of this wearable interface relative to a more traditional desktop computer.

To evaluate the system, participants with no prior military experience were trained in hostage rescue procedures in a game-based simulation environment using either a wearable or desktop interface. A control group was trained in the same procedures in a live action condition. Following training, each group completed a series of missions in the live condition, with their performance video and audio recorded for scoring purposes. Participants were scored on the number of correct actions and the time to complete each mission. Results indicated that participants trained in the live condition performed better and were faster than those trained in either simulator condition. However, there were no significant differences between the simulator conditions for performance or speed. This indicates that although the wearable interface was expected to provide better training than the desktop interface, each interface provided equivalent levels of training transfer. These results underscore the importance of determining the training effectiveness of novel training methods before fielding. Although a novel training method may appear to be superior to more traditional methods, the new method should still be evaluated empirically to determine its training effectiveness.

ABOUT THE AUTHORS

Dr. John S. Barnett is a research psychologist with the U.S. Army Research Institute for the Behavioral and Social Sciences, Technology-Based Training Research Unit, in Orlando, Florida. He conducts research in the use of simulators and virtual environments to train U.S. Army Soldiers. He holds a Ph.D. in Applied Experimental and Human Factors Psychology from the University of Central Florida and is a Certified Professional Ergonomist (CPE).

Dr. Grant S. Taylor is a researcher with the Applied Cognition & Training in Immersive Virtual Environments (ACTIVE) lab at the Institute for Simulation and Training, where his research focusses on the application of various psychophysiological measurement techniques. Grant obtained a Ph.D. in Applied Experimental and Human Factors Psychology from the University of Central Florida in 2012.

Comparing Training Transfer of Simulators: Desktop versus Wearable Interfaces

John S. Barnett
U.S. Army Research Institute for the Behavioral and Social Sciences
Orlando, Florida
john.barnett1@us.army.mil

Grant S. Taylor
University of Central Florida: Institute for Simulation and Training
Orlando, Florida
gtaylor@ist.ucf.edu

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 advances 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. Simulation developers often assume that improving realism automatically improves training. Features found in the field environment may be added to the simulation to make the user experience more like the field. However, there are many elements in the field environment that have no practical effect on training. Therefore, adding those elements to the simulator environment may enhance realism but does not necessarily 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.

This research is the second of two experiments 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.

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.

To test transfer of training, participants were trained to conduct a hostage rescue task using either 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.

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 study.

BACKGROUND

Simulation Training

The essence of training is to introduce learners to declarative and procedural knowledge related to certain skills and then give them the opportunity to correctly practice and improve those skills. Simulators give learners the opportunity to practice skills in situations that otherwise would be difficult or dangerous. It is much safer for a student pilot to practice landings in a simulator where the penalty for failure is much less catastrophic than with a real aircraft. Training Soldiers in a simulated environment allows them to use weapons and tactics that would be dangerous to practice in the real environment.

A key question in simulator training and the use of virtual environments is how realistic must practice be to improve performance. Ideally, the procedures practiced in the simulator should be exactly the same as those for the real environment. However, for practical reasons some actions in a simulated environment cannot be exactly the same as in real life. For example, if a game-based simulator uses a desktop computer as an interface, the avatar in the simulated environment is moving through the environment controlled by the learner's mouse and keyboard, but the learner is normally seated in a chair and not moving. Is the student in this simulator still learning, even if they aren't moving? The answer to this question often depends on the type of skills to be learned.

Skills are often divided into motor skills and cognitive skills. Motor skills involve bodily movement and fine muscle coordination, such as those that are used in sports such as golf or tennis. Hitting a golf or tennis ball and having it go where you want involves training groups of muscles to make very fine movements. Typically, this level of skill requires a considerable amount of practice. Cognitive skills involve remembering procedures required to perform a task and sometimes problem solving. Cognitive skills involve memory more than musculature. For example, remembering how to change a flat tire on a car is more of a cognitive skill than motor skill. It is more important, for safety and practical reasons, to remember how to perform the steps in the correct order than how to physically operate the tools, such as the jack. The tire wrench, jack, and other tools do not require fine motor skills; rather, almost anyone who has the strength to operate the tools has the ability to use them correctly. Soldier tasks often require the performance of both motor and cognitive skills. For example, a Soldier operating a checkpoint must have a thorough understanding of the Rules of Engagement and Escalation of Force procedures in order to know when it is necessary to fire on a suspected threat or enemy, and they must also be well trained on the physical operation of their weapon in order to effectively eliminate a threat.

Thus, learning motor skills through simulation requires the simulation to be an accurate representation of the physical operation of the real world system. On the other hand, learning cognitive skills requires the learner to remember and think through the correct procedures, while the exact physical movements are less important (Wickens, 1992).

Training Transfer

How well skills learned via a simulator or other training device improve performance of the same skills in the real world is termed transfer of training. Generally, if the behaviors practiced in the simulator are similar to the behaviors required to perform the real world skill, the transfer of training is high (Wickens, 1992). Often, though, practical considerations prevent the simulator behaviors from being exactly like those in the real world. For example, a driving simulator may have a steering wheel and pedals similar to a real automobile, but "driving" the simulator will not include the acceleration and deceleration forces of driving a real car. If those forces are important for learning how to drive, then the training transfer may be poor, however, if experiencing acceleration and deceleration forces is not important to learning to drive, then training transfer should not suffer.

Ideally, training in a simulator or simulation should improve the performance of skills in the real world. This is known as positive transfer of training. If the training did not improve or worsen those skills, it would be considered zero transfer. However, sometimes poor or inappropriate training can interfere with the performance of real world skills. This is known as negative transfer. For example, learning to type using an atypical keyboard layout, such as the Dvorak keyboard, would make it more difficult to learn to type on another keyboard layout, such as the common QWERTY layout. Obviously, any training, especially simulator training, must have positive training transfer to be useful. Zero transfer would be a waste of valuable training time, and negative transfer would make the student worse at performing the skill. Transfer of training is an important consideration when developing training and selecting training systems.

With many training systems it is not as simple as positive, zero, or negative overall training. Often, a training system will produce different levels of positive (or sometimes negative) transfer for different skills. Therefore, in most practical situations, training developers are looking for a system that can produce some level of positive transfer of training for the majority of skills, with little or no negative transfer.

Immersive Simulators

A virtual environment that has a greater sense of immersion should produce higher levels of presence, that is, the subjective feeling of being in one environment when actually being in another (Knerr, et al., 1998). While immersion is primarily a mental state, the physical analog is fidelity. A training system has high fidelity if it matches the real world system very closely. Although it is logical to believe that a simulator with high fidelity will train better than a lower fidelity system, research has shown that this is not always true (Wickens, 1992). In some cases, the added realism of high fidelity simulators may not provide enough training improvement to justify the increased costs. In other cases, simulators with high fidelity, but which are not exact matches to the simulated system, can force users to learn simulator-unique actions that are incompatible with the real system. These simulator-unique behaviors can actually interfere with the learning of skills needed for the real system. Wickens (1992) suggests it is important to know which components of training have to be similar to the target task and which are less important to learning.

The use of wearable simulators for dismounted Soldier training is a relatively recent development. Initial studies investigating their effectiveness found that although early systems did allow Soldiers to perform basic Infantry tasks, they were too bulky and lacked the fidelity in their visual and weapons systems necessary to be truly useful (Lockheed Martin, 1997; Pleban, Dyer, Salter, & Brown, 1998). Over the past decade, simulation technology has continued to advance, and researchers have continued to investigate their usefulness for the training of dismounted Soldiers (see Knerr, 2007). However, this research has been limited (due in no small part to the costs associated with wearable simulators), and the research that has been done has primarily revolved around subjective questionnaires to assess how effective users felt the systems were, rather than objective measures of their training effectiveness. Of the few studies to objectively measure training, only one (Loftin et al., 2004) compared the immersive system to a standard desktop simulation, but they used a CAVE (CAVE Automatic Virtual Environment; a simulator that projects images onto large fixed screens) rather than a wearable system. Their results showed a minor improvement in training from the CAVE over the desktop, but not enough to justify the tremendous increase in cost.

Current Research

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.

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 a wearable simulator interface could 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 occur based on using a wearable interface, desktop interface, or traditional live training. This report presents the results of this research effort.

METHOD

Participants

A total of 62 participants completed the study, 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. Participants completed the Gaming Experience Measure (GEM; Taylor, Singer, & Jerome, 2009) and Game Performance Assessment Battery (GamePAB; Taylor et al., 2009) to identify their level of gaming skill. 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.

In addition, participants completed the Interest/Enjoyment and Perceived Competence scales of the Intrinsic Motivation Inventory (IMI; McAuley, Duncan, & Tammen, 1987) to determine their motivation to learn the skill. 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.

Participants were also administered the Simulator Sickness Questionnaire (SSQ; Kennedy, Lane, Berbaum, & Lilienthal, 1993). The SSQ 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.

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 16:9 aspect ratio. The software used for these scenarios (see Figure 1), was the Game-based Distributed Interactive Simulation (GDIS). This software 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.

Wearable Simulation. Those participants assigned to the wearable simulation condition were trained on the procedures using an ExpeditionDI immersive wearable interface. This interface uses a combination of the user’s 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.

***Figure 1.* Participant’s view within the GDIS virtual environment.**

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.

Live Environment. Those assigned to the live environment condition were trained on the hostage rescue tactics in real rooms, with life-size cardboard cutouts of enemies and hostages (Figure 3). They were provided with a replica M4 rifle, which was an Airsoft™ rifle designed to shoot plastic pellets. Participants also wore replica fragmentation grenades and flashbangs (stun grenades), as well as an ammo vest to carry the grenades, and a helmet and goggles for safety (Figure 4). During the experiment, no plastic pellets were used with the rifle. Instead, the participants' shots were scored based on the sound of the rifle firing and the orientation of the rifle at the time of firing.

Figure 2. **The ExpeditionDI wearable simulation system.**

Procedure

Participants completed the study 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 study requirements, and that they had no prior hostage rescue experience. Participants then completed a baseline measure of the SSQ. Following the SSQ, participants completed the GEM and GamePAB 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), with both participants working together as a team within the same training environment. 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. 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.

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).**

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.

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

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 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 Intrinsic IMI.

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 study 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 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.

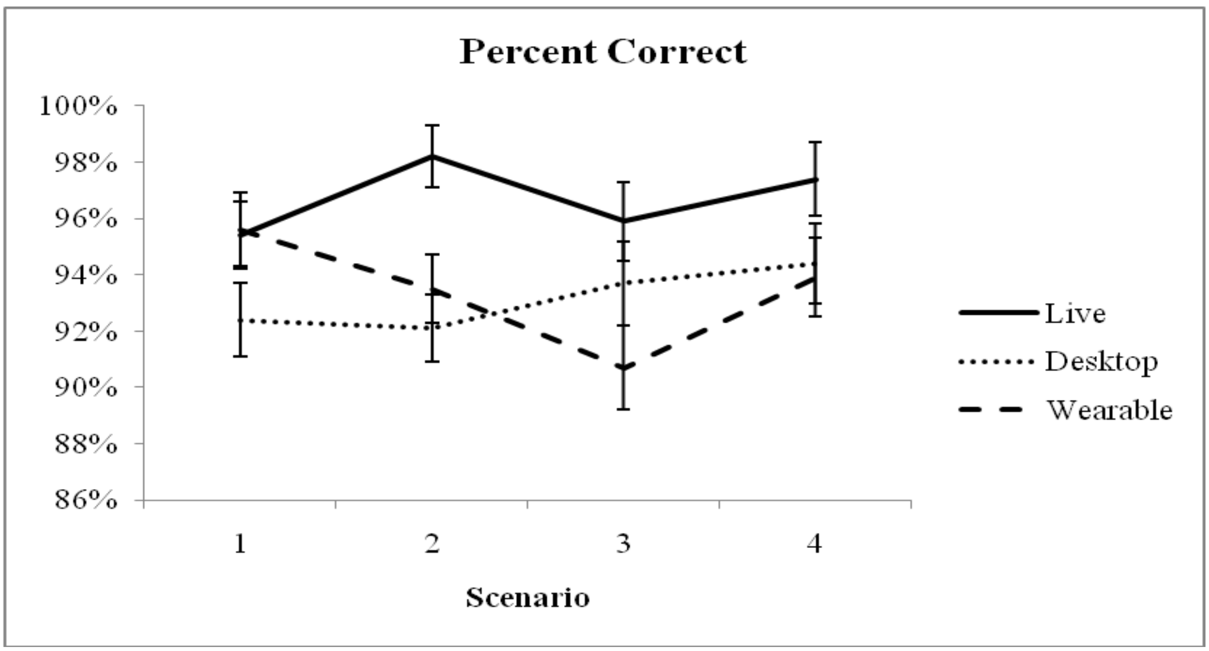

***Figure 5.* Percent of actions performed correctly on the four test scenarios.**

The effect of 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 ($M = 93.1\%$, $SD = 3.80$, $p = .010$) and Wearable ($M = 93.5\%$, $SD = 6.29$, $p = .018$) 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).

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, $SD = 4.74$, $p < .001$) and Wearable ($M = 33.54$s, $SD = 2.96$, $p < .001$) 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 to perform significantly faster than both the Desktop and Wearable conditions across all four scenarios, though the strength of this effect diminished over time, with the Live training condition's performance times remaining consistent as the Desktop and Wearable training conditions' performance times improved over time.

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 ($M = -0.477$, $SD = 3.76$, $p = .001$) and Live training conditions ($M = 0.000$, $SD = 4.16$, $p = .001$), 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 ($M = 0.379$, $SD = 1.69$, $p < .001$) and Live training conditions ($M = 0.000$, $SD = 2.34$, $p < .001$) 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 ($M = 0.000$, $SD = 0.000$, $p = .020$) and Live training conditions ($M = -1.27$, $SD = 5.94$, $p = .012$) with no significant difference between the Desktop and Live conditions ($p = .886$).

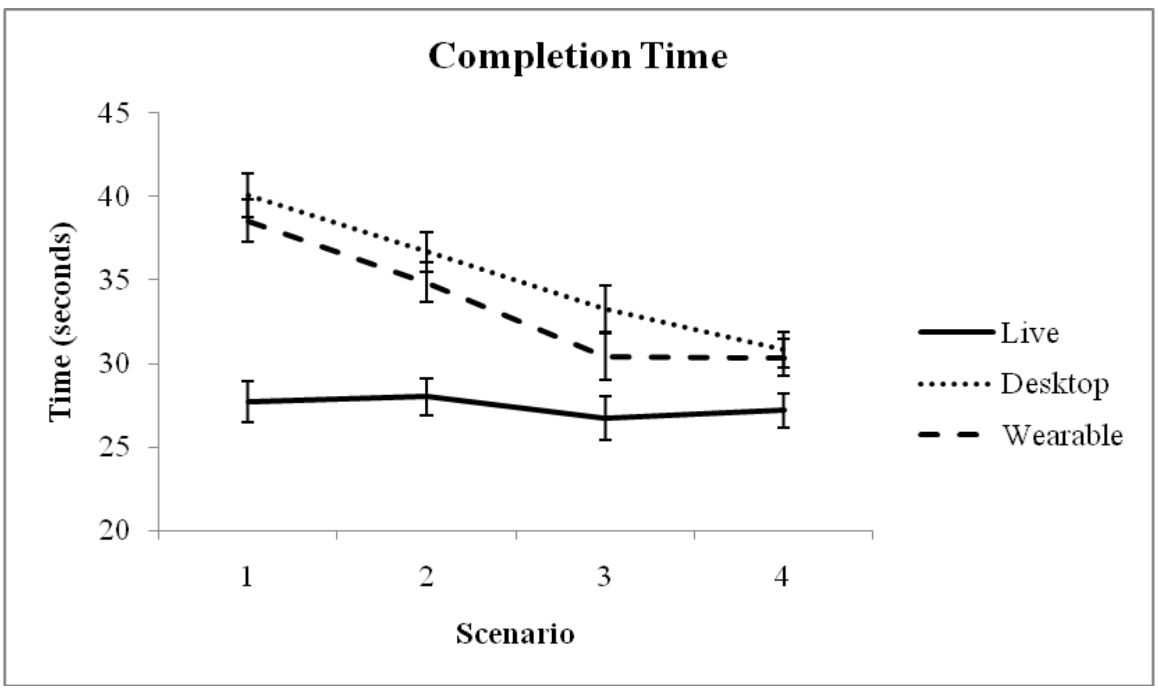

Figure 6. **Scenario completion time for the four test scenarios by training condition.**

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 ($M = 5.10$, $SD = 0.632$, $p < .001$) and Wearable conditions ($M = 5.26$, $SD = 0.510$, $p = .001$), 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) and Wearable conditions (M = 4.77, SD = 0.517, p = .002), 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).

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 the same 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. Determining whether improved time scores result from familiarity with the environment or improving psychomotor skills may be a productive topic for future research.

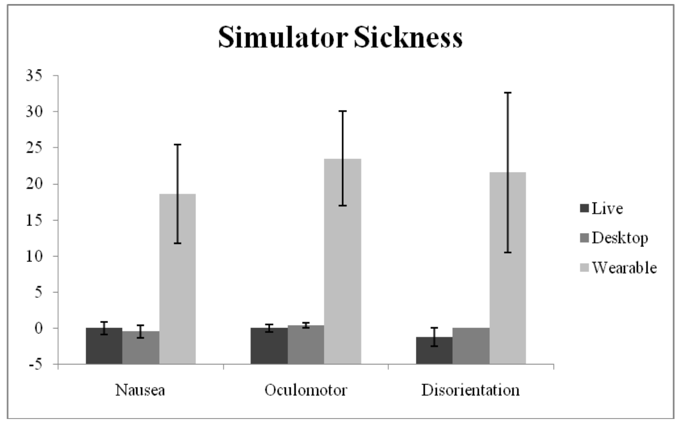

Figure 7. **Simulator sickness values reported from each training condition.**

Note: Values are reported as change from the baseline data collected prior to training, with positive values indicating an increase.

The 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.

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.

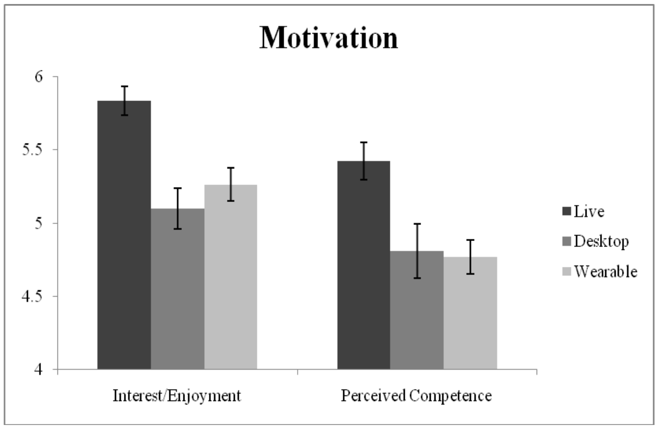

***Figure 8*. The Interest/Enjoyment and Perceived Competence subscales of the Intrinsic Motivation Inventory as reported from each training condition.**

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, an extrapolation of completion times suggests that all groups would have had equivalent completion times on about the fifth scenario if the trend had continued.

In conclusion, the results of this experiment, in conjunction with previous evaluations, suggests that using a wearable interface to train dismounted Soldier skills provides no discernible benefit over a more traditional desktop computer interface. In fact, the wearable interface has several disadvantages, the greatest of these being the expense of the system (a wearable interface currently costs roughly 10 times as much as a powerful desktop computer). The wearable interface's greater tendency to induce simulator sickness is another concern, as these symptoms are likely to interfere with the training process and limit the amount of time trainees can spend in the system. Most importantly, the results of this evaluation reiterate the critical need for empirical evaluation of novel technologies. Although early subjective ratings considered the wearable interface to be a promising new technology, empirical results are ultimately necessary to determine the true effectiveness of any system.

REFERENCES

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.

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.

Knerr, B. W. Lampton, D. R., Singer M. J., Witmer, B. G., Goldberg, S. L., Parsons, K. J., & Parsons, J. (1998). *Virtual Environments for Dismounted Soldier Training and Performance: Results, Recommendations, and Issues* (ARI Technical Report 1098). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Lockheed Martin Corporation. (1997). *Dismounted warrior network front end analysis experiments* (Advanced Distributed Simulation Technology II, Dismounted Warrior Network DO #0020, CDRL AB06, ADST-II-CDRL-DWN-9700392A). Orlando, FL: US Army Simulation, Training and Instrumentation Command. (ADA344365).

Loftin, R.B., Scerbo, M.W., McKenzie, R., Catanzaro, J.M., Bailey, N.R., Phillips, M.A., & Perry, G. (2004). *Training in peacekeeping operations using virtual environments.* Paper presented at the RTO HFM Symposium on "Advanced Technologies for Military Training." (ADA428142).

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., Dyer, J. L. Salter, M. S., & Brown, J. B. (1998). *Functional capabilities of four virtual individual combatant (VIC) simulator technologies: An independent assessment* (Technical Report No. 1078).

Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (ADA343575).

Taylor, G. S. & Barnett, J. S. (2011). *Training transfer 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.

Wickens, C. D. (1992). *Engineering psychology and human performance (2nd ed.).* New York: Harper Collins.