The Fidelity Matrix: Mapping System Fidelity to Training Outcome

Peter Muller
Potomac Training Corporation
Lansdowne, Virginia
muller@potomactrainingcorp.com

LCDR Joseph Cohn
Naval Research Laboratory
Washington, DC
cohn@itd.nrl.navy.mil

CDR Dylan Schmorrow
Office of Naval Research
Arlington, Virginia
schmord@onr.navy.mil

Roy Stripling
Naval Research Laboratory
Washington, DC
Roy.stripling@nrl.navy.mil

Kay Stanney, Laura Milham,
and David Jones
Design Interactive, Oviedo, FL
Kay, Laura@designinteractive.net

Mary C. Whitton
University of North Carolina
Chapel Hill, NC
whitton@cs.unc.edu

Jennifer E. Fowlkes
CHI Systems
Orlando, Florida
JFowlkes@chisystems.com

ABSTRACT

One of the biggest challenges in designing Virtual Environment (VE) training systems is identifying the fidelity requirements for the component technologies. Initial fidelity-related design decisions are often motivated by the belief that the more accurately the VE stimulates individual components of the human sensory system, the more likely the system will provide effective training. Given that stimuli in the real world are not presented in a simple, scripted manner, it is quite probable that this is an unrealistic goal. Consequently, the development of effective VE training systems requires a more holistic approach and must focus on how these sensory systems converge to support performance at the task level within the VE. To evaluate the success of this approach, this process also requires the development of performance metrics that enable the assessment of how a component's fidelity relates to training outcomes, in terms of different types of sensory information. The current work discusses an initial application of this method to investigate the relationship between system design and performance in the context of a basic Military Operations in Urban Terrain (MOUT) task. While these results provide specific design recommendations for MOUT training, they also suggest a broader application for designing, testing, and evaluating training systems.

ABOUT THE AUTHORS

Peter Muller is president of Potomac Training Corporation. He serves as the Systems Engineer for ONR's VIRTE program and also supports the ONR Human Performance, Training, and Survivability program. He has a BS degree from Syracuse University, and a MS in Systems Management from USC.

LCDR Joseph Cohn is an Aerospace Experimental Psychologist in the U.S. Navy Medical Service Corps. Dr. Cohn received his commission in 1999 as a Lieutenant and completed naval flight training in February 2000. He has overseen the validation and transition of three combat simulators to the Fleet and Force and was the Founder and past Director of the Naval Research Lab Warfighter Human Systems Integration Laboratory.

CDR Dylan Schmorrow, MSC, USN, Ph.D., is a U.S. Naval Officer and an Aerospace Experimental Psychologist in the Navy's Medical Service Corps. He is serving as the Executive Assistant to the Chief of Naval Research (CNR) and as the Director of the Warfighting Enhancement Program Office managing Expeditionary Warfare and Warfighter Performance programs that are transforming promising technologies into operational capabilities.

Roy Stripling, Ph.D. is a Research Biologist at the U.S. Naval Research Laboratory in Washington, D.C. He received his Ph.D. in Neuroscience from the University of Illinois at Urbana-Champaign. Dr. Stripling is the Director of the Warfighter Human System Integration Laboratory at the Naval Research Laboratory, where he seeks

to leverage neuroscientific and technological advances to enhance human performance and improve training technologies.

Kay Stanney, Ph.D. is president of Design Interactive, Inc. and Professor and Trustee Chair with UCF's Industrial Engineering & Management Systems Department, where she joined in 1992. She conducts research in the areas of multimodal interaction, human-systems integration, and augmented cognition. She is currently conducting user-centered design research for ONR's Virtual Technologies and Environments Program.

Laura Milham, Ph.D. is the Director of Training Systems at Design Interactive, Inc. She has served as a Lead Scientist within the Navy's VIRTE program, guiding the design of human computer interaction aspects of virtual environment training systems, designed transfer of training experiments, and conducted training evaluation experiments.

David L. Jones is a Research Associate at Design Interactive, Inc. His primary research focus is on designing multimodal interactive systems. He has a Masters in Industrial Engineering from the University of Central Florida, with a focus on human-computer interaction and usability. He has a Bachelors in Human Factors Psychology from Embry-Riddle Aeronautical University, with a focus on human factors and computer science.

Mary C. Whitton has been involved in the development and evaluation of high performance graphics, visualization, and virtual environment systems for over 25 years. She came to Carolina after 16 years in industry and is now a Research Associate Professor. Whitton's research focuses on understanding what makes virtual environment systems effective and on developing technology and techniques to make them more effective. She has an M.S. degree in Electrical Engineering from North Carolina State University.

Jennifer E. Fowlkes, Ph.D. has twenty years of experience in areas of human factors and training which includes team training and performance, training effectiveness evaluations, and simulator sickness research. Most recently, her research has focused on measuring team performance in distributed training environments and assessment of haptics technologies for training. Dr. Fowlkes holds a Ph.D. in Experimental Psychology from the University of Georgia.

The Fidelity Matrix: Mapping System Fidelity to Training Outcome

Peter Muller
Potomac Training Corporation
Lansdowne, Virginia
muller@potomactrainingcorp.com

LCDR Joseph Cohn
Naval Research Laboratory
Washington, DC
cohn@itd.nrl.navy.mil

CDR Dylan Schmorrow
Office of Naval Research
Arlington, Virginia
schmord@onr.navy.mil

Roy Stripling
Naval Research Laboratory
Washington, DC
Roy.stripling@nrl.navy.mil

Kay Stanney, Laura Milham,
and David Jones
Design Interactive, Oviedo, FL
Kay, Laura@designinteractive.net

Mary C. Whitton
University of North Carolina
Chapel Hill, NC
whitton@cs.unc.edu

Jennifer E. Fowlkes
CHI Systems
Orlando, Florida
jfowlkes@chisystems.com

INTRODUCTION

Early behavioral theorists suggested that the transfer of learning is dependent upon the presence of identical elements (Osgood, 1949; Thorndike, 1906) and thus necessitates high-fidelity learning environments. As cognitive learning theory evolved, greater emphasis was placed on the learner's role and the argument over fidelity evolved into a need to replicate only certain elements of the real situation. Hays and Singer (1988) suggest that, "The real issue is how to replicate those parts of the task situation which are necessary for learning to perform the task." The challenge is to determine the fidelity requirements that constitute the most effective training environment. Milham, Hale et al (2004) suggest there is a range of fidelity factors which must be taken into consideration when analyzing fidelity requirements. These include functional fidelity, the ability of any system to support the appropriate stimulus response set; psychological fidelity, the degree to which the system affords the appropriate performance cues; and physical fidelity, the extent to which the system provides multi-modal sensory stimulation. Not surprisingly, these three factors will be weighted differently depending on the types of skills that a training system is meant to train. Ultimately, a mixed or "blended" fidelity training solution can be designed, which identifies how best to meet training objectives through an optimal mix of classroom instruction, training technologies, and live events to ensure a desired level of readiness is achieved (Carter & Trollip, 1980). This paper seeks to identify a blended training solution for a dismounted infantry MOUT trainer.

Motivation

As one walks the exhibition halls of conferences such as I/ITSEC, HCII, and others that have a significant focus on learning science research and development, one will encounter dozens of Modeling and Simulation based training tools using a host of various technologies. These tools are often advertised as providing effective training, enhanced situational awareness, improved cognitive capabilities, and other performance enhancement related panaceas. A closer examination of these performance claims typically reveals that surprisingly little actual assessment has been performed. This does not indicate malfeasance; more likely it speaks to the relatively high degree of complexity associated with assessing these systems *in vivo*. In an ideal world, every system component would be subjected to human factors assessments, such as Usability Analyses (Nielsen, 1993), and following integration into the complete system would be assessed using a Transfer of Training paradigm (ToT; Cohn, Schmorrow et al, 2005; Milham, Hale et al, 2004; Murdock, 1957). In this way, problems with the system could be identified early in the development stages, and success could be assessed at the later ones.

However, limitations of both time and money often preclude this complete approach from being taken. This has led to a negative cycle in the training world, in which, on the one hand, the need for training tools such as virtual environments (VEs) are becoming increasingly clear (c.f. Davies, 2002), while on the other hand, the need for increased clarification of the utility of these systems is becoming more pronounced. While there is certainly an increase in the frequency of

ToTs being conducted on the final, developed training system, it is often at the cost of not doing earlier evaluations on the system components. This will necessarily lead to any development program shouldering additional risk, since the later in the development cycle these assessments are done, the more difficult it is to incorporate any changes indicated by the ToT results. This paper describes one large-scale research effort, the Virtual Environments and Technologies (VIRTE) program that is emphasizing the role of component-wide analysis prior to the final ToT evaluation.

VIRTE

The Virtual Technologies and Environments (VIRTE) program is an Office of Naval Research sponsored interdisciplinary research program that seeks to examine technologies that fully immerse marines and sailors in a Virtual Environment, so they can train on tasks that are too dangerous, too expensive, or impossible to do in the real world. The first stage, Virtual Environment Expeditionary Warfare (VE EW) tried to demonstrate the degree to which a Human Centric Simulation Design Model could be developed, validated, and used to rapidly prototype vehicle based VE training systems. The second stage, Virtual Environment Human Interface Technology (VE HIT) capitalized on this success and focused on non-vehicle team-based VE training systems that directly interface to individual users, rather than through the metaphor of a vehicular interface. The third stage, Multi-platform Operational Team Training Immersive VE (MOT2IVE) synthesized the results of the earlier efforts into a unique, cross-platform multi-spectrum training environment that will include advanced training enhancement tools and strategies, and which will lay the foundation for addressing larger training challenges.

Focus

This paper will focus on efforts devoted to developing dismounted infantry simulations under the VE HIT program. These technologies raise a unique set of complex Human Computer/Human Systems Integration questions by their very nature. The most fundamental challenge in designing these tools is that the interaction between the human and the simulation is not easily mapped to the real world. Turning a steering wheel in a simulation is very much like turning a steering wheel in a vehicle. Using a joystick to walk, however, is not at all like walking. While children learn this mapping early with computer games, the training transfer is not well understood. A second unique feature is that in vehicle simulations, most of the required information is presented to the trainee through the vehicle systems. In a dismounted setting, however, the trainee must directly experience and interact with their virtual world. This presents a formidable challenge, since the real world is filled with multiple sets of informational cues, impinging on all five senses, yet capturing the physics of these sensory cues, and replicating them virtually is beyond the state of the art. A further complication is that modeling the underlying dynamics of human interactions with objects (including other humans) is still a formidable task, from both a simulation and technology perspective.

COMPONENT TECHNOLOGY RESEARCH

The first step in developing any VE system is to identify the training requirements and objectives and to identify the component technologies that will comprise the actual system. For example, within the MOUT domain, there is a range of training goals (targets), from exposing students to the basic facts surrounding MOUT through providing students with realistic scenarios to support learning consolidation (Table 1). The challenge is to map the desired training to the type of training technology and then identify the types of technologies that need to be integrated to support this system. Ideally, the specifications for each of these pieces will be based on rigorous human performance based test and evaluation efforts, but it is often based on the lowest cost COTS technology. The following sections outline VIRTE’s research efforts at the component level to evaluate their expected performance impact prior to integration into the VIRTE MOUT system.

Table 1. Theorized use of Training Options to Optimize Training Effectiveness

Type of Training	Target of Training	Issues
School House Functional Fidelity: Low Physical Fidelity: Low Psychological Fidelity: Low	Declarative knowledge, facts	Difficult to practice skills and consolidate knowledge
Low Fidelity (partially immersive) Functional Fidelity: High Physical Fidelity: Medium Psychological Fidelity: Low	Consolidate declarative knowledge and acquire procedural knowledge	Difficult to acquire higher-order skills and strategic knowledge
High Fidelity (fully immersive) Functional Fidelity: High Physical Fidelity: High Psychological Fidelity: Medium	Higher-order skills and strategic knowledge (e.g., SA, team coordination)	Can overwhelm and distract early declarative and procedural learning; Cost; Limited availability; may require support staff to run
Live Functional Fidelity: High Physical Fidelity: High Psychological Fidelity: High	Higher-order skills and strategic knowledge (e.g., SA, team coordination)	Can overwhelm and distract early declarative and procedural learning; Cost; Limited availability; may require expert trainers to run

Framework For Multi-Modal Sensory Integration for Effective Training

As Table 1 suggests, the types of systems likely to be most effective for VE HIT's purposes involve some level of immersive technology, and require the robust delivery of sensory information (Visual, Haptic, and Aural). Since the majority of human performance research in this area has focused on the visual domain, little is known about how to blend other modalities to 'recreate' a sensorial experience similar to that experienced in the real world. Thus, one must first consider what the multimodal information requirements are and how they could be integrated into a single training package.

The VIRTE effort has concentrated on the following major sub categories:

1. Visual Displays, Navigation, and Locomotion
2. Haptic Interactions,
3. Aural/Auditory Interactions

The method used for determining the multi-modal requirements of a task was a Sensory Task Analysis (STA; Milham et al., 2006). This framework describes a methodology for identifying critical training design needs by decomposing task components into the multi-modal sensory elements (e.g. visual, auditory, haptic) necessary to support successful task accomplishment. Using this information, it was then possible to determine how to represent these operational cues in the virtual environment, either as operationally realistic visual, auditory, or haptic cues or as metaphoric cues. Because of limitations of the VE, information may not always be provided through the same mode. For example, haptic information may need to be provided visually or aurally. From this, scenarios can be designed that fold training needs into the virtual environment scenario design.

Visual Displays, Navigation, and Locomotion Interfaces for Dismounted Users in Virtual Environments

Dismounted Infantry simulations require a deep understanding of how humans plan and execute movements, in this case, locomotion. A range of studies has demonstrated the primacy of vision in controlling locomotion (Lee, 1976), while others have explored the role of locomotion interface characteristics on performance (Grant & Magee, 1998; Chance, Gaunet, Beall, & Loomis, 1998). To investigate the intersection of these two areas, and to understand the effect of interface on the motion paths followed by users, a set of studies was undertaken to characterize movement trajectories and task performance under different visual and locomotion interface conditions for users who are (virtually) moving on foot in VEs. The goal was to be able to rank locomotion interfaces used in virtual environments by the similarity of results, comparing them to results using natural vision and locomotion.

In these studies, participants moved between target positions and performed tasks in both a computer-generated environment and the corresponding real environment, to allow for a comparison of a range of visual and interactive interfaces across multiple conditions. The five experimental conditions used throughout were a combination of one of three locomotion interfaces (walking, walking-in-place, and joystick walking), and one of three visual conditions (head-mounted display, unrestricted natural vision, or field-of-view-restricted natural vision (Table 2).

Table 2. Visual and locomotion interface pairs studied. Table cells contain short name for each condition.

Locomotion \ Visual	Normal	Restricted FOV	HMD
Really Walk	Real	Cowl	VRWalk
Gamepad-Joystick			JS
Walk-in-Place			WIP

In the first study, the goal was to understand how technology choice impacted the execution of simple motions in a simple task: users walked to targets on walls and stopped as close to them as they could without making contact (Cohn, et al. 2004, Whitton, et al. 2005). The second study added a layer of complexity, requiring users to avoid aurally presented gunfire while moving between barriers behind which they were told to hide.

A critical result from this first study was that the correlations of critical motion path values (e.g. peak velocity and peak deceleration) for the conditions taken pair-wise suggest a coarse ordering of locomotion interfaces by “naturalness” of both the visual and the locomotion condition: high-bin correlations include only pairs that both use real walking; middle-bin correlations pairs all include the walking-in-place (WIP) condition; and low-bin correlation pairs all include the joystick condition. Of note is that, although the walking-in-place (WIP) implementation used in the studies was acknowledged to be difficult to use, the WIP results correlate more strongly with natural walking and seeing than does the joystick condition.

These pair-wise correlations of the values of the critical motion path points fall roughly into three bins (Table 3) with Bin A representing high correlations. These data support the notion of ranking visual and locomotion interfaces on *naturalness*: arguably, real vision is more natural than restricted FOV vision, which is, in turn, more natural than using an HMD, and really walking is more natural than walking-in-place, which is, in turn, more natural than using a joystick. As really walking in VR is impractical except in special circumstances, the message in this data is that VE system builders must continue to search for a truly usable walking-in-place technique.

Table 3. Pair-wise correlations fall into three bins. Results for the three VR (HMD) conditions suggest an ordering by the naturalness of their locomotion interface: really walking results in motions more like really walking with natural vision than walking-in-place, and walking-in-place results in motion more natural than a joystick interface.

Experimental Condition (Locomotion Condition : Visual Condition)	Real (Walk:Real)	Cowl (Walk: Restricted FOV)	VRWalk (Walk:HMD)	WIP (Walk-in-place: HMD)	JS (Gamepad-Joystick:HMD)
Real (Walk:Real)					
Cowl (Walk:Restricted FOV)	A				
VRWalk (Walk:HMD)	A	A			
WIP (Walk-in-place:HMD)	B	B	B		
JS (Gamepad-Joystick:HMD)	C	C	C	C	

Further validation of this ordering of systems was evident in the results from the second study, which looked at minimizing exposure of the trainee’s avatar to perceived gunfire. Here, the VR-Walk condition was the only condition that showed improved performance over successive trials; the joystick and walking-in-place conditions were consistently the worst. This suggests that learning is better facilitated through more natural types of interfaces.

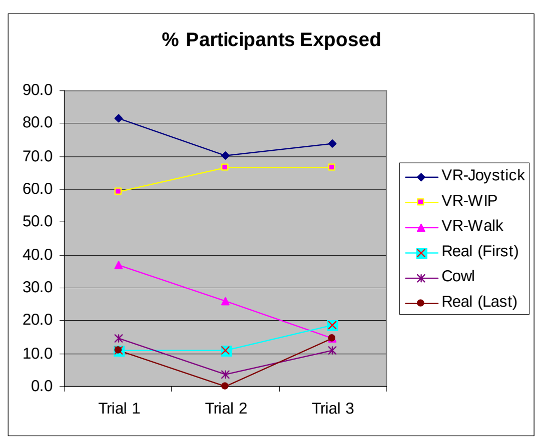

Figure 1. Over three training trails, participants in the VR Walk condition showed the best improvement of all five conditions

Kinesthetic and Proprioceptive Aspects of Locomotion Interfaces for Dismounted Infantry Training

A classic challenge in the design of GUIs and Web-based applications is ensuring that users do not get 'lost', and that, if they do, proper mitigation strategies are easily accessible (Sellen & Nicol, 1990) Similarly, in immersive VEs for dismounted infantry, trainees may get disoriented while moving and navigating through hallways, stairs, or rooms. In the real world, entering a dark hallway or stairwell can lead to similar confusion; however, years of experience walking through a range of real world environments trains our nervous system to maintain reasonable orientation through short stretches by relying on kinesthetic and proprioceptive information (Harris, Jenkin, & Zikovitz, 2000). In the virtual world, while it is not possible (yet) to faithfully provide this sort of feedback, it is possible to afford users an interface option that enables them to move in a manner similar to how they would in the real one by providing physical feedback. Consequently, it is expected that those interfaces that best provide equivalent information, and an equivalent way of interacting with this information, should be better at supporting locomotion and navigation in visually degraded virtual environments.

In addition to an evaluation of interfaces in terms of control and visual information, it is also possible to evaluate them in terms of the degree to which the self-reflective (kinesthetic and proprioceptive) information they afford maps onto the information requirements for executing specific behaviors (Grant & Magee, 1998). Therefore, a set of experiments was conducted that tested navigational precision and accuracy through three tasks, one of which provided visual cues, and two of which did not. The tasks included:

- Maze navigation with visual cues
  - Participants navigated a maze of two hallways and one turn, and at the end were asked to point back to their starting position. In this environment the first hallway was 9m long, the second hallway was 5m long, and the angle of the turn was 45°, 90°, or 135°.
- Rotating in place without visual cues
  - Participants were asked to turn 45°, 90°, or 135° on successive trials.
- Direct walking to a previously viewed target after visual cues were removed
  - Participants were shown an object at a distance of 3.33m, 6.66m, or 10m. After the object and other visual environmental cues were removed, the participant was asked to walk to where the object had been.

Each participant conducted these three tasks on one of three different locomotion systems:

- standard joystick (Thrustmaster Top Gun Fox 2 Pro), that controlled both locomotion and rotation
- optical tracking system that utilized an algorithm for translating "walking-in-place" movements of the legs into locomotion in the VE, and that translated real-world rotations and head movements into VE rotations and head movements
- "hybrid" system that used optical tracking to translate real-world rotations and head movements into VE rotations and head movements, but used a joystick to control all other aspects of locomotion.

Analysis of the data provide some support for the general hypothesis that enabling more natural user interactions supports performance, but also revealed the importance of clearly defining the rules for using specific interfaces. All three interface systems were able to support equivalently accurate estimation of the point of origin in the maze task (where visual cues were provided), with both optically tracked systems (which each translated real-world turning into VE turning in a one-to-one manner) being statistically better than the joystick, and not different from one another. This lends credence to the notion that higher fidelity representations of movement (i.e. equivalent to anticipated kinesthetic and proprioceptive information) are valuable. Counter-intuitively, in the direct walking task, the joystick and "hybrid" system (both of which used a joystick for forward locomotion) were statistically better than the "walking-in-place" interface. Post-experiment interviews suggested that users were erroneously applying a "two for one" mapping when using the "walking-in-place" interface. That is, they assumed that to take one step forward in the VE would require them to take two steps-in-place in the real world. This interpretation of the user's behavior was supported by additional analysis of the data which revealed that users in the "walking-in-place" interface systematically overshot their intended target by 2-fold on average.

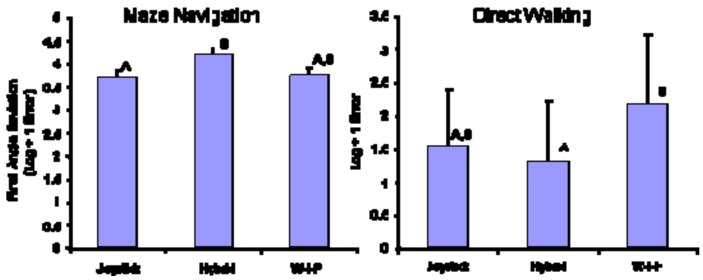

Figure 2. (Left) Results from the Maze task; (Right) Results for the direct walking task. Bars marked with different letters are statistically different from one another (within each task).

Haptics Interactions in Support of Dismounted Infantry Training

Haptics refers to the incorporation of tactile information into displays provided within operational systems and training environments (Hale & Stanney, 2004). Since normal interactions in the real world often involve a high degree of haptic information, it should be expected to support widespread training applications. Used this way, haptics can enhance the perception of immersion, expand the breadth of skills that can be trained, improve task performance, enhance learning and retention, improve spatial orientation or situation awareness, and facilitate teamwork.

However, despite progress, our understanding of how best to use haptics to support dismounted infantry training is in its very early stages. VIRTE's haptics effort focused on developing a taxonomy for using haptics to support MOUT training and using this taxonomy to develop mapping strategies between haptic tools (sensor placement, sensor stimulus properties) and desired performance.

The taxonomy identifies two broad areas in which haptics may impact training in **VEs**. *Haptic Simulation* refers to simulation of information **naturally** provided by the touch sensory modality. This information can support navigation, object detection and identification, and teamwork just as it does in the real world and potentially increase transfer. *Haptic metaphors* include haptic uses such as alerts and displays to provide spatial orientation. This information may enhance training efficiency by increasing the usability of VEs,

This taxonomy has been used to develop and test hypotheses regarding haptic effects. For example, tactors were placed as shown in Figure 3 to support the training of a team searching task in a VE (Fowlkes et al, 2005). Tactors were used to convey information about team member physical contact and to provide information

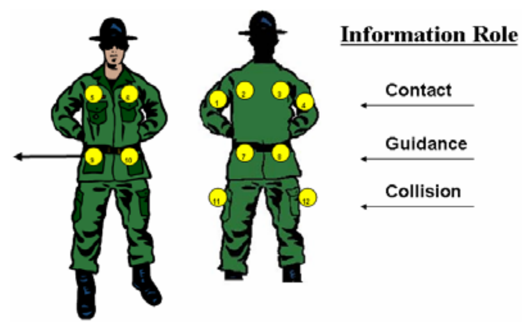

Figure 3. Potential mapping between sensor placement and desired information.

about team member physical contact and to provide metaphoric information to trainees (to provide steering guidance and to alert trainees that they had collided with objects in the VE). Haptics better enabled participants to maintain physical contact with a simulated team member. In addition, as shown in Figure 4, team communication skills showed better

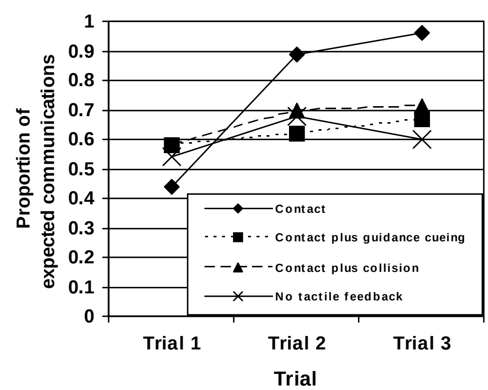

Figure 4. Communication scores for haptic groups across trials.

acquisition over the three training trials for the group provided with information about contact with the simulated team member compared to the other haptics groups (provided also with guidance or collision metaphoric haptic displays) and compared to the no haptics group.

Aural

Auditory (aural) information provides another critical source of sensory information. To understand how to optimize the presentation of this information within the MOUT setting, first, an auditory scene analysis was carried out to determine important auditory cues for a room clearing task as well as metaphoric cues used for guiding trainees. Then, research experiments were conducted to explore the effect of spatialization fidelity on the training effectiveness of the MOUT system.

The auditory scene analysis suggest the use of a number of cues from Greenwald's (2002) Critical Cue Inventory (CCI), which are naturally present in the real-world environment, as well as metaphoric cues that are expected to train participants to avoid dangerous situations. These sounds were targeted at assisting VE trainees in locating and differentiating between hostile and non-hostile units and included enemy voices, movement sounds, and hostile fire, as well as the movement and voices of non-hostile units present. The metaphoric cues were aimed at assisting trainees in developing effective strategies. They included a foghorn, to symbolize when participants were in a doorway, room entry area, or in front of a mouse hole; breaking glass, whenever participants were standing in front of a window; and a pan flute, whenever participants were too close to walls when walking down a hallway. By cueing trainees when they were in such areas, they could change their room clearing strategy in order to avoid such obstacles.

To determine the optimal approach for integrating these cues –and the best technology for doing so- an experiment was designed to examine the value of integrating the identified audio cue strategy at four levels of spatialization fidelity (i.e., no audio, non-spatialized audio, generalized HRTF, and best-fit HRTF) into the MOUT training system. Participants performed a room clearing task with two primary objectives: 1) clear rooms of hostiles as efficiently as possible and 2) confirm the detection of non-hostiles. Performance on this task was based on the time required to clear friendly units, time required to fire on enemy units, and the total number of hostile and non-hostile units engaged. While performing this task, participants were also required to avoid any areas in the environment that would pose a threat to them. The total time participants were in danger areas and number of times they entered them were recorded and used to compare performance under the various auditory spatialization fidelity levels.

The results suggested that:

- Trends were present suggesting that the average time required to clear rooms and engage hostiles decreased as audio spatialization levels increased.
- A decreasing trend in the time spent in entrance danger areas as fidelity increased except for under the best-fit HRTF condition. This condition resulted in performance approximately equal to the non-spatialized condition.
- A pattern was also present in the average time participants spent in front of mouse holes as a function of audio condition, with decreasing time spent in front of mouse holes as spatialization fidelity increased.
- In terms of perceived workload, participants felt that the non-spatialized and best-fit HRTF conditions were significantly more temporally demanding than the generalized HRTF condition ($p = .027$, $p = .003$, respectively). Interestingly, participants in the best-fit HRTF condition also rated the task as being more complex than all other audio conditions ($p < .05$).

Taken together, the results support the use of spatialized audio in VE training systems but are equivocal when it comes to the call between a generalized and best-fit HRTF.

Summary of Component Research

The above results suggested that two types of systems could be developed. The first, essentially a low fidelity training tool, could be comprised of:

- Visual display: flatscreen LCD Panel
- Locomotion: Joystick or low end walking in place
- Haptics: Built into the locomotion device, provided very general contact/collision cueing
- Audio: Minimal cueing, providing warnings and alerts

The second, essentially a high fidelity training tool, could be comprised of:

- Visual Display: wide field of view Head Mounted Display
- Locomotion: Full body optical tracking to support walking in place paradigm (Templeman, Denbrook & Sibert, 1999).
- Haptics: Series of arrays placed across the trainee's body, providing both real and metaphoric cues.
- Audio: Spatialized audio using general HRTFs.

In order to ensure that these systems serve as more than technology showpieces, a framework for integrating the range of studied sensory cues using a scenario based training approach was used.

SYSTEM WIDE RESEARCH: A PATH FORWARD

Muller, Cohn and Nicholson (2003) proposed a 5 stage process for Training Effectiveness Evaluation, starting with up front analyses, leading into component wise assessments –iteratively when possible- and ending with a transfer of training study. While much has been written about the initial (Chipman, Schraagen & Shalin, 2000) and final (Lathan, Tracey et al 2002) steps, there has been little guidance or utility demonstrations of the intermediate one. This paper endeavored to provide an example of how this component wise assessment could be conducted and how the results could be, used to define training system specifications. As suggested, the results from this type of effort may often indicate that there is more than one functional solution to a given training challenge. Resources permitting, the utilization of a training transfer paradigm (Murdock, 1957; Bessemer, Boldovici, & Bolton, 2002) could best provide decision support.

For the current MOUT training system, specific questions to be addressed in order to evaluate the final system could include the following:

1) How much live training can be saved via simulator training at both low and high fidelity?
2) How does low fidelity compare to high fidelity in its relative efficacy to save live training for technical and higher-order skills, both immediately upon training completion and some time later?
3) What is the relative amount of expert performance that can be expected from low versus high fidelity training? (informal comparison)

Studies to answer these questions should support the examination of training transfer, across low fidelity, high fidelity, and live training environments, of both technical (e.g., enemy engagement, room clearing, exposure, and survivability skills) and higher order (e.g., spatial relation knowledge, situation awareness, etc.) skill sets in ground-based operations that may rely on a combination of functional, physical, and psychological fidelity, paying attention to how persistent MOUT training is with a given system. Such a study would combine fidelity, skill type and retention interval into a structured matrix that would enable comparisons between trainees receiving instruction in a low fidelity desktop VE or a high fidelity fully immersive VE, in terms of performance of similar tasks in a real world indoor MOUT environment. The transfer effectiveness ratio (TER; Roscoe, 1971) could be used to specify the trials/time saved in the live environment as a function of prior trials/time in each training platform and the incremental transfer effectiveness ratio (ITER; Flexman et al., 1972) could be used to determine the transfer effectiveness of successive increments of training in each training platform; with successive increments of training predicted to decrease the average TER and ITER to a point where additional training is no longer effective.

CONCLUSION

The results from this effort should fill in a distinct gap in the training community, by providing trainers with a systematic methodology to identify the technology components necessary for developing a blended fidelity training solution. Importantly, this approach is best used in combination with other ones, such as up front analyses capitalizing on Task Analytic methods and back end, whole system assessment ones, like a transfer of training study. In this way, the Warfighter, the ultimate customer, will be assured of receiving effective training that improves performance in real-world situations.

ACKNOWLEDGEMENTS

The author's would like to thank the sponsorship of the Office of Naval Research's Virtual Technologies and Environments program.

REFERENCES

Bessemer, D., Boldovici, J, & Bolton, A. (2002) *The Elements Of Training Evaluation*, Army Research Institute for the Behavioral and Social Sciences, Alexandria, VA.

Carter, G., & Trollip, S.R. (1980). A constrained maximization extension to incremental transfer effectiveness, or, how to mix your training technologies. *Human Factors*, 22, 141-152.

Chance, S. S. Gaunet, F., Beall, A. C. & Loomis, J. M. (1998). Locomotion Mode Affects the Updating of Objects Encountered During Travel: The contribution of vestiular and proprioceptive inputs to path integration, *Presence*, vol. 7, pp. 168-178.

Chipman, S.F., Schraagen, J.M., & Shalin, V.L. (2000). Introduction to cognitive task analysis. In J. Maarten Schraagen, S. F. Chipman, & V.L. Shalin (Eds.), Cognitive task analysis (pp. 3-23). Mahwah, NJ: Lawrence Erlbaum Associates.

Cohn, J., Schmorrow, D., Nicholson, D., Muller, P., Stanney, K., & Milham, L. (2005). Trainee-Centric Design for Developing Military-Based Virtual Environment Training Systems. *Proceedings of Human Computer Interaction International 2005, Las Vegas, Nevada.*

Davies, R.C. (2002). Applications of system design using virtual environments. In K.M. Stanney (Ed.), *Handbook of Virtual Environments: Design, Implementation, and Applications* (pp. 1079-1100). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Flexman, R.E., Roscoe, S.N., Williams, A.C., Jr., & Williges, B.E. (1972). Studies in pilot training: The anatomy of transfer. *Aviation Research Monographs*, 2(1), 1-87.

Fowlkes, J.,E., Durlach, P.J., Drexler, J. Daly, J., Alberdeston, R., & Metevier, C. (2002). *Optimizing Haptics Perceptions for Advanced Army Training Systems: Impacts On Performance*. Paper presented at the 23rd Army Science Conference.

Fowlkes, J.E., Washburn, D., Eitelman, S., Daly, J., & Cohn, J.V. (2005). Use of haptic displays to enhance training in a virtual environment. *Proceedings of the 11th International Conference on Human- Computer Interaction*, 22-27 July. [CDROM]

Grant, S. C. and Magee, L. E. (1998) Contributions of proprioception to navigation in virtual environments, *Human Factors*, vol. 40, pp. 489-497.

Greenwald, T. (2002). *An analysis of auditory cues for inclusion in a virtual close quarters combat room clearing operation.* Master's Thesis, Naval Postgraduate School, Monterey, CA.

Hale, K.S., & Stanney, K.M. (2004). Deriving haptic design guidelines from human physiological, psychophysical, and neurological foundation. *IEEE Computer Graphics and Applications*, 24(2), 33-39.

Harris, L. R., Jenkin, M., & Zikovitz, D. C. (2000) Visual and non-visual cues in the perception of linear self-motion. *Experimental Brain Research*, 135(1), 12-21.

Hays, R. T. & Singer, M. J. (1988). *Simulation fidelity in training system design: Bridging the gap between reality and training*. New York: Springer-Verlag.

Lathan, C.E., Tracey, M.R., Sebrechts, M.M., Clawson, D.M., & Higgins, G.A. (2002). Using virtual environments as training simulators: Measuring transfer. In K.M. Stanney (Ed.), Handbook of virtual environments: Design, implementation, and applications (pp. 403-414). Mahwah: NJ: Lawrence Erlbaum Associates.

Lee, D.N. (1976) A theory of visual control of braking based on information about time-to-collision, *Perception*, vol. 5, pp. 437-459.

Milham, L., Hale, K., Stanney, K., Cohn, J., Darken, R., & Sullivan, J. (2004). When is VE training effective? A framework and two case studies. Proceedings of *48th Annual Human Factors and Ergonomics Society Conference*, New Orleans, LA

Milham, L., Jones, D., Bell, M., Chang, D., Delos Santos, K., Stanney, K., & Cohn, J. (2006). Beyond Training Effectiveness Evaluation: Guiding the Enhancement of Military Training. Paper submitted to *the Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC) Annual Meeting*. Orlando, FL.

Muller, P., Cohn, J.V., & Nicholson, D. (2003). Developing and Evaluating Advanced Technologies for Military Simulation and Training. Proceedings of *25th Annual Interservice/Industry Training, Simulation and Education Conference*, Orlando, FL.

Murdock, B. B. (1957). Transfer design and formulas. *Psychological Bulletin*, 54(4), 312-326.

Nielsen, J. (1993). Usability engineering Academic Press, San Diego: CA

Osgood, C. E. (1949). The similarity paradox in human learning: a resolution. *Psychological Review*, 56, 132-143.

Roscoe, S.N. (1971). Incremental transfer effectiveness. *Human Factors*, 13, 561-567.

Sellen, A., & Nicol, A. (1990). Building user-centered on-line help. In B. Laurel (Ed.), *The art of human-computer interface design* (pp. 143-153). Reading, MA: Addison-Wesley.

Templeman, J. N., Denbrook, P.S. & Sibert, L. E. (1999). Virtual Locomotion: Walking in Place through Virtual Environments. *Presence*, 8(6) 598-617

Thorndike, E. L. (1906). *Principles of learning*. New York: A. G. Seiler.

Whitton, M., Cohn, J., Feasel, J., Zimmons, P., Razzaque, S., Poulton, S., McLeod, B., Brooks, F., (2005) "Comparing VE Locomotion Interfaces," *Proceedings of IEEE Virtual Reality 2005, (Bonn, Germany March, 2005), 123-130*, IEEE Computer Society.