

Training in Affectively Intense Virtual Environments

Lawrence M. Wilfred
Department of Information Science and Technology
University of Missouri – Rolla
USA
lmwwdb@umr.edu

Michael G. Hilgers
Department of Computer Science
University of Missouri-Rolla
USA
hilgers@umr.edu

John M. Hortenstine
Department of Computer Science
University of Missouri – Rolla
USA
jhort@umr.edu

Madhu Reddy
Department of Information Science and Technology
University of Missouri – Rolla
USA
mreddy@umr.edu

Richard H. Hall
Department of Information Science and Technology
University of Missouri – Rolla
USA
rhall@umr.edu

Ming C. Leu
Mechanical Engineering Department
University of Missouri – Rolla
USA
mleu@umr.edu

Christopher P. Walker
Department of Computer Science
University of Missouri – Rolla
USA
cpwalk@umr.edu

Abstract: The purpose of this research was to examine the impact of affective intensity of a virtual reality (VR) training environment on learning, as demonstrated by performance within a “real life”, affectively intense environment. Participants completed a VR “training” scenario in which they were required to locate victims of a terrorist attack either in an “affectively intense” or “neutral” environment. Participants then attempted to locate the rooms containing the victims within the actual building the VR environment was modeled on, as they listened to affectively intense audio. The major findings were: 1) Those who trained in the affectively intense environment performed substantially better in the “real” environment; 2) Participants in the two environments did not differ with respect to autonomic arousal or perceived presence; and 3) Those more experienced with computer games reported a higher degree of presence in the virtual environment and performed better in the “real” environment.

Introduction

Context

A virtual reality system was developed by the University of Missouri-Rolla (UMR), with the support of the U.S. Army’s Tank-automotive and Armaments Command (TACOM, grant # DAAE07-02-C-L068) to help train first responders in dealing with terrorist attacks involving the use of weapons of mass destruction. This is referred to as the First Responder Simulation and Training Environment (FiRSTE) project (Berry & Hilgers, 2004; Hall et al., 2004; Leu et al., 2003; Misra, Decker, Barker, & Hilgers, 2004; Tichon, Hall, Hilgers, Leu, & Agarwal, 2003) (also, see <http://campus.umr.edu/firste>). As a part of this project, research was conducted to examine basic psychological and physiological processes associated with virtual reality based training for learning in affectively intense environments. Affectively intense learning refers to learning in which the task to be learned is performed in a highly stressful, emotionally intense environment (such as first responders to weapons of mass destruction) (Hall et al., 2004). There is very little research on the factors that contribute to effective learning in these types of environments. The main goal of this research was to examine the optimal degree of affective intensity within the virtual environment that lead to the most effective learning in a “real world” affectively intense performance environment.

The research used a first responder scenario for training and testing, utilizing a specialized version of the virtual environment developed as part of the FiRSTE project.

Affectively Intense Learning

The term “Affectively Intense Learning” was coined by researchers at the University of Missouri – Rolla (Hall et al., 2004; Tichon et al., 2003) to denote the type of learning associated with strong emotional states where the environments tend to be stressful, such as military training. Intuitively, it would seem that this type of learning would be particularly difficult to teach effectively using a traditional classroom approach. Presumably, training would be much more effective in an environment where stress conditions are heightened. Such affectively intense environments are very difficult to recreate and control in a systematic manner without virtual reality technologies.

Our contention is that the affective intensity congruence between training and performance is important for effective learning. Some support for this is provided by the “mood congruence” and “state-dependent learning” effects. Mood congruence refers to the aspect of human memory where persons are more likely to remember facts that coincide with their mood. In other words, if the person were in a good mood, then she would be more likely to remember pleasant information (Eich & Macaulay, 2000). The same pertains to unpleasant information. If the person is in a bad mood, then she will be more likely to remember unpleasant information. A similar phenomenon called state-dependent learning refers to the phenomenon in which people recall information easier if they can return to the same emotional and physical state when they learned the information. A basic contention of the present research is that this phenomenon can be extended to include affectively intense situations - the research addressed the effects of high affect training environments and subsequent performance in a corresponding high affect real world environment. Would subjects trained in an affectively intense environment display better performance compared to those trained in a control environment, where affective intensity was lower, while other factors remained the same?

Pilot Experiment: Virtual Terrorist Attack on the Computer Science Building

The present experiment was preceded by a pilot study, based on a model, which posits four classes of variables as influencing the virtual reality and learning outcome relationship: learner variables, virtual environment factors, perception of presence, and affective intensity (Hall et al., 2004). An experimental environment was developed, based on a virtual model of the University of Missouri – Rolla’s Computer Science building, adding special affectively intense effects that might be associated with a terrorist attack. The virtual environment was created using the game engine *Half-Life*© in accordance with their licensing agreement.

The environment consists of a 3 floor building, which includes classrooms, offices, computer labs with appropriate furniture and textures to match the actual Computer Science building at the University of Missouri – Rolla. Affectively intense effects were added such as random fires and explosions which would be triggered in the general vicinity of the avatar controlled by the user. The building included fire extinguishers located at each end of the floor. The objective of the user was to locate the fire extinguishers and put out the fires and trigger fire alarms located around the building. There was limited foam in the fire extinguishers and the user would “die” if he/she stood close to the fire for too long. “Death” was signified by a 30 second idle time when the user’s control over the avatar was immobilized. Another feature was that the lights in the various rooms could be toggled between on and off.

The main goals of this pilot experiment were:

1. To determine the measures of perceived affective intensity corresponding to the affectively intense events in the environment, through quantitative comparisons of galvanic skin response (GSR).
2. To determine the degree of presence experienced in the environment through survey tools, qualitative observation, and participants’ self report.
3. To identify methodological problems and/or other issues involving the efficacy of the research environment, via collection of qualitative data.

The pilot study conducted with 5 subjects showed that there was a marked increase in the GSR readings associated with explosions in comparison to periods when explosions were not occurring. There was also increased GSR activity due to frustration caused by inhibited locomotion and lack of lighting after it was turned off by a user in the virtual environment. Most participants stated that they recognized rooms where they had classes; one student found the environment so realistic that he was reluctant to put out the fires because “I never much liked the

Computer Science building anyway”. One subject who had experience with computer games found the controls counter-intuitive since all the controls were on the 5 button mouse (Hall et al., 2004).

Present Experiment

The present research followed a similar methodology, while extending the pilot study in a number of ways. Some of the fundamental differences were: a) The training environment was experimentally manipulated and participants randomly assigned to one of two conditions (affectively intense vs. affectively neutral); b) Five times more subjects participated, allowing for some limited use of inferential statistics; c) A performance/testing component was added where users were tested in a non-virtual environment (i.e., the “real” Computer Science building); d) Participants used a head mounted display; and e) The task to be performed in the virtual environment was more apparent and result oriented.

The research focused on finding the factors that influenced affectively intense learning (Figure 1.1). The numbers in the figure represent the experimental questions.

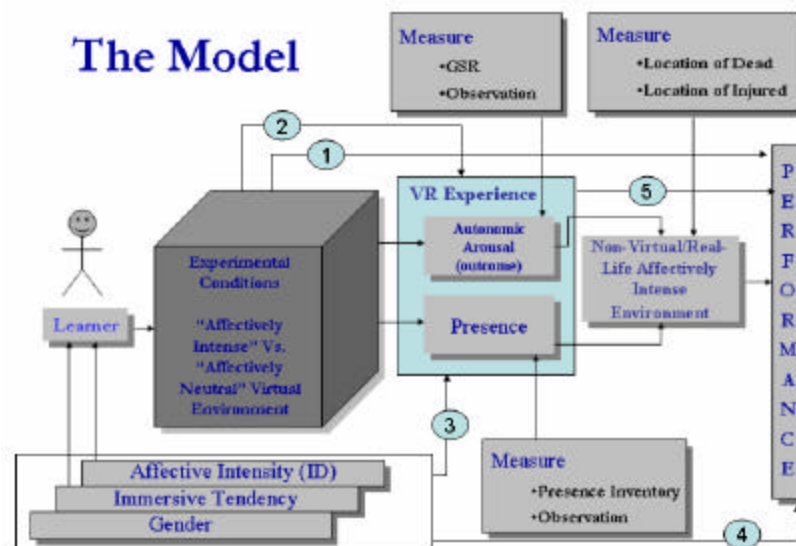


Figure 1: Research Model for Current Research

The specific experimental questions were as follows.

1. How do learning outcomes differ as a function of experimental condition?
2. How does VR experience differ as a function of experimental condition?
3. What is the relationship between individual differences and VR experience?
4. What is the relationship between individual differences and learning outcomes?
5. What is the relationship between VR experience and learning outcomes?

Method

Participants

The participants were 30 subjects selected from the University of Missouri – Rolla, mostly recruited from an undergraduate class in Web Development and Design, in return for extra credit. Twenty two participants completed the entire experiment. Five participants did not complete the experiment due to equipment failure and three participants did not complete due to VR sickness.

Materials

Individual Difference measures included Affective Intensity and Immersive Tendency (gender was also recorded). The virtual reality experience measures included galvanic skin response (GSR) and the Presence Questionnaire. The learning outcomes metric was based on the correct dead and injured count after the participant had completed the inspection of the actual Computer Science building at the University of Missouri – Rolla. Brief descriptions of the metrics used for the experiment are given below.

Immersive tendency is the psychological state characterized by perceiving oneself to be enveloped by, included in, and interacting with an environment that provides a continuous stream of stimuli and experiences (Witmer & Singer, 1998). It may be noted that the immersive tendency measure had subscales to measure focus, gaming tendency, and involvement. Affect Intensity is defined as the stable individual differences in the strength with which individuals experience their emotions (Larsen & Diener, 1987). Skin conductance is one of the fastest responding measures of stress response and has been found to be one of the most robust and non-invasive physiological measures of autonomic nervous system activity (ANS) (Helander, 1978). Presence is the degree to which a participant feels that a virtual environment is not mediated by any technological interface. A five item questionnaire, developed by Slater and colleagues, was used in this experiment (Slater & Steed, 2000).

Three virtual environments were created using the *Half-Life* © game engine, one was an acclimation environment. The other two were virtual reality environments based on the Computer Science building at the University of Missouri – Rolla. One of them was affectively intense with explosions and fires, while the other was affectively neutral.

The bulk of the experiment was conducted at the Laboratory for Information Technology Evaluation (LITE) which is especially equipped for usability testing. The various inputs from the video channels like the head mounted display, the graphical representation of GSR change on the Biopac control computer, the camcorder output, and the auditory effects are processed by the mixer which is output to the VCR for recording purposes and displayed on a television for the tester in the experiment room

PROCEDURE

The participants were briefed on the rationale behind the test and the basic testing procedure and presented with a consent form. An informal interview was then carried out, to ascertain general information on the individual gaming backgrounds, their gaming interests, gaming experience, virtual reality experience, and any relevant history of motion or VR sickness. Baseline GSR measurement was taken while the participant completed the Immersive Tendency and Affect Intensity questionnaires. The participants were then presented with instructions on how to operate in the virtual environment. The head mounted display was adjusted into place on the participant and the acclimation VR environment was started up. The participants were encouraged to navigate in the virtual environment using the 5 button mouse.

On successful completion of the acclimation environment, the participant was asked to assume the role of a first responder fire fighter and presented with a scenario and task directions. Participants were randomly assigned to either the affectively intense or affectively neutral environments with the constraints that: a) An effort was made to counter-balance the order as much as possible; and b) An effort was made to have an equivalent number of participants in the two groups. GSR measurement was started and the video of the session recorded. The inspection was to be conducted while the participants defused situations occurring around them such as fires and explosions (note that the explosions only occurred for those in the affectively intense group). The virtual reality session lasted 12 minutes and the video of the session, used for observational analysis, consisted of 3 separate videos merged into one via the mixer – facial expressions of the participant with the head mounted display, continuous graphical GSR readings and the virtual reality simulation of the Computer Science building. The participant then completed the presence questionnaire. On completion, an informal interview was conducted to ascertain the state of the participant, the impact of the virtual environment and what possible improvements could be made to it.

The next phase of the experiment was conducted at the actual Computer Science building at the University of Missouri – Rolla, which is located approximately 200 yards from the LITE laboratory. The participant was given a CD player with headphones, which played tracks that correspond to random explosions recorded from the affectively intense VR session. This was used to recreate as closely as possible an affectively intense environment in the real world with as little inconvenience to other users of the building. The participant was given 12 minutes to locate the rooms where they found dead/injured personnel to the best of their memory and asked to record this on a piece of paper.

The experiment timeline is summarized in the following table.

Time(Minutes)	Task
0-10	Briefing, Consent Form, Informal Interview
10-20	GSR Baseline, Immersive Tendency Questionnaire, Short Affect Intensity Questionnaire
20-25	VR Instructions
25-30	VR acclimation environment
30-40	VR scenario and Instructions
40-57	VR test environment of Computer Science building, GSR measurement, Video Recording
57-60	Stop Recording, Removal of electrodes
60-65	Presence questionnaire
65-70	Debriefing, Informal post interview
70-82	Inspection of Computer Science building

Table 1: Experimental Timeline

RESULTS

Quantitative Results (This section details the outcomes that were statistically significant.)

Learning Outcomes as a Function of Experimental Condition

In order to compare the effects of experimental condition on injured count correct, a Univariate Analysis of Variance was computed with experimental condition as the independent variable and the injured correct as the dependent variable. This Analysis of Variance was marginally significant $F(1, 20) = 3.178$, $p = .09$, $eta^2 = 0.137$. Note that an eta^2 of .137 is a medium to large effect size based on Cohen's criteria (Cohen, 1969). The descriptive statistics associated with these analyses are displayed in Table 2.

	Affectively Intense	Affectively Neutral	Significance
Dead Correct	2.250(0.462)	3.100(0.506)	Not significant
Injured Correct	3.250(0.397)	2.200(0.435)	$p = .09$, $eta^2 = .137$

(Standard deviation in parentheses)

Table 2: Mean Learning Outcomes as a Function of Experimental Group

Relationship between individual differences and VR experience

In order to determine the degree of relationship between Immersive Tendency and the VR Experience parameter presence, Pearson's correlations were calculated pairing each of the Immersive Tendency sub scale scores with Presence. These correlations and statistical significance are displayed in Table 3.

Subscale	r
Involvement	.34+
Focus	.329+
Games	.54*++

* $p < .05$; + medium to large effect size; ++large effect size

Table 3: Pearson Correlations between Immersive Tendency and Presence

Relationship between Individual Differences and Learning Outcomes

In order to assess the degree of relationship between immersive tendency and learning outcomes a series of Pearson correlations were computed pairing each of the three immersive tendency subscales with both dead correct and injured correct. The results of these analyses are displayed in Table 4.

Scale	Learning Outcome	
	Dead Correct	Injured Correct
Immersive tendency: involvement	.14	.18
Immersive tendency: focus	-.02	-.57**++
Immersive tendency: games	.55**++	.10
Affective intensity (ID)	-.37(*)+	-.37(*)+

(*) $p < .10$; ** $p < .01$; +medium to large effect size; ++large effect size

Table 4: Correlations of immersive tendency subscales and affective tendency with learning outcomes

In order to assess the impact of gender on learning outcomes two one-way analyses of variance were computed. In each, gender (male vs. female) was the independent variable and in one dead correct was the dependent variable and in the other injured correct was the dependent variable. The first ANOVA, in which dead correct was the dependent variable, was not statistically significant. However, the second ANOVA was statistically significant $F(1, 20) = 7.11, p < .05, \eta^2 = .26$. Descriptive statistics associated with these two analyses are displayed in Table 5.

Outcome	Gender	
	Male	Female
Dead Correct	2.73(1.79)	2.43(1.27)
Injured Correct*++	3.27(1.39)	1.71(.95)

(Standard deviation in parentheses)

* $p < .05$; ++large effect size

Table 5. Mean Learning Outcomes as a function of Gender

Qualitative Results

The qualitative data consisted pre-experimental interviews, observations of the participants during the experiment, observations of video recordings of the experiments and post-experimental interviews. Three major themes arose from the analysis of the data. They were user performance, measurement tools and methodology as shown in figure 2.

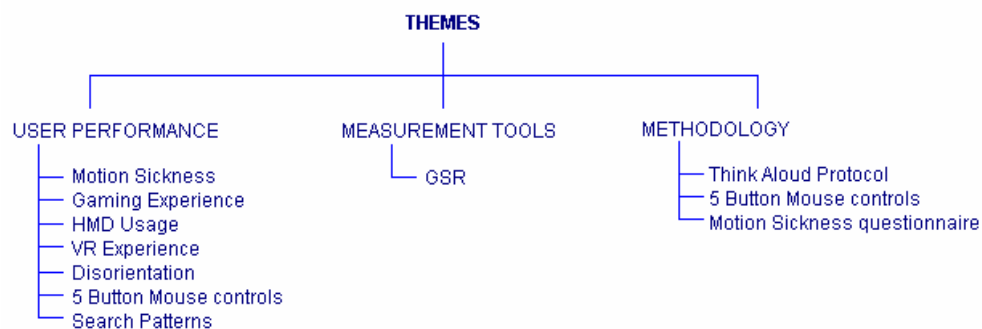


Figure 2: Themes from qualitative data analysis

We observed a multitude of factors that affected user performance to some degree. These factors either improved or degraded a user's performance. For instance, the search pattern employed by the participants made a substantial impact on user performance. Those participants who located the injured and dead first before putting out fires performed better than their counterparts who tried putting out the fires as they encountered them.

The measurement tool was also commented on by the users. The GSR (Galvanic Skin response) as a measurement tool was slightly cumbersome due to the error induced by slight movements by the user. Also, users had to keep one hand immobile during the course of the experiment because the GSR measuring device was attached to it. The users complained about this encumbrance.

Finally, methodology related issues like the effectiveness of the Think Aloud Protocol were highlighted. It was difficult to get users to verbalize their thoughts as they were navigating in the environment. Also, the 5 button

mouse for navigation affected users' mental model of navigation in a virtual environment. The users assumed that they would use both a mouse and keyboard and had trouble only using a mouse.

Discussion

There is evidence that the degree of affective intensity within a training virtual environment is, indeed, related to learning performance within a "real world" affectively intense environment. Though this was found to be true for only the injured correct count as the learning outcome, it may be explained by the "reality factor" associated with it. Participants formed a very realistic and practical approach to finding the injured personnel over the dead. This goal very much translates into the real world in what is required of a learner from an affectively intense environment. For example, if fire fighters were trained in a virtual environment and they did not give precedence for an injured person as far as evacuation action or first aid is concerned, then the practical purpose of such simulations would be lost.

The affective intensity of the environment did not affect perceived presence or autonomic response, as measured by GSR. A few possible explanations: 1) The individual differences effected presence so much that they overwhelmed any effect for experimental condition; 2) Users in both the experimental conditions had a high level of presence; 3) Factors like frustration effected GSR so much that it overwhelmed the effects of explosions and other characteristics of the affectively intense environment; and 4) GSR turned out to be a less than optimal measure of autonomic arousal (to be described in more detail below).

Immersive tendency and gender had a substantially strong impact on presence. All of the immersive tendency measures, especially games, were strongly related to presence. In addition, males rated their degree of presence substantially higher than females. Adjunct analyses indicated that males tend to have a much greater preference for computer games than do females. This preference for video games, which are in fact quite similar to the virtual reality simulations, caused them to be more receptive to the virtual world. Computer games today are modeled quite closely to a perceivable world to make game playing as realistic an experience as possible. Most games try to increase the feeling of being there and having performed in that fantastic world. There is a general tendency of males to have a greater affinity for games, which can be noted from the increased "Games" score. This tendency was correlated with presence. Hence males tend to exhibit increased presence over their female counterparts.

On the other hand, the affective intensity individual difference measure was not significantly related to either of these VR experience measures, and none of the individual difference measures were related to GSR. Possible reasons: 1) The affective intensity measure may not have been appropriate for this type of experiment as discussed above; 2) Variations in GSR could have been flawed or sometimes misleading due to various error inducing factors; 3) External factors such as ambient humidity, temperature and other factors such as recent physical activity among others could have changed the physiology of the skin which in turn could have changed the GSR readings. GSR readings are taken by means of electrodes attached to the tips of two fingers; the equipment measuring it is extremely sensitive and erratic hand movements can cause sharp peaks. From a practical stand point, heart rate may be better measure of autonomic arousal, which will be noted for future research.

Some immersive tendency subscales strongly related to learning outcomes, though in somewhat unusual and sometimes unexpected ways. With immersive tendency, focus was very strongly negatively related to injured personnel located, while games was strongly positively related to dead correct. This brings out distinct dimensions of the tendency to be present. Focus deals with the ability to focus on the task at hand while sub scale "Games" refers to the tendency to be present within the context of video games. In both cases, the total score determines the tendency of one to be immersed. This can be explained if one were to assume that the group scoring high on "Focus" got extremely involved in the environment such that the environment parameters like visual detail, affective cues (only in the affectively intense virtual environment), and navigation complexity, were so overwhelming that the simulation objective eluded them or did not hold their required attention. Computer games today seem to emphasize increasing one's kill count to win. Such games often require one to "kill" a lot of "bad guys" or "be killed". Although avid gamers may have a tendency to become highly immersed, they would know how to distinguish a "dead person" so that he/she would not have to "waste ammunition on it" or "would know, 'its going to stay dead and not come after me to kill me'". Therefore it would follow that avid gamers or subjects who score high on the immersive tendency subscale "Games" would find more dead personnel in the virtual environment.

The individual difference, affective Intensity was fairly strongly and negatively related to both learning outcomes. This may also help to explain the negative "Focus" relation with learning outcomes, in that one can picture a person who tends to react in a very affectively intense manner and tends to be very focused on the reality of the environment and, as a consequence, is distracted from the learning task.

Males performed substantially better than females in what appeared to be the most important learning outcome (injured correct). Their mean is almost twice as high as the females. This is consistent with qualitative observations, where they adopted a practical objective to save the injured first and their general tendency to avoid locating the dead altogether.

Degree of presence perceived was largely unrelated to learning outcomes. This is consistent with previous studies (Mania & Chalmers, 2001; Moreno & Mayer, 2002), so it appears that this finding extends to affectively intense learning, based on this experiment.

Autonomic response was not significantly related to learning outcomes, though the correlations were much stronger than with presence and effect size for the correlation between GSR and dead correct is medium to large. So, there is some evidence that the affective intensity experienced in the VR environment, as represented by GSR, is more strongly related to performance in an affectively intense environment, than is presence. If there had been a large sample (more power) these correlations would very likely have been statistically significant.

There are a number of possible ways that these results could be extended in future research. First, work could include feed back of autonomic nervous system measures to the participant in the affectively intense virtual environment to try and get them to regulate or control their current emotional state and then determine the impact that such loop systems have on the learning process. Second, this can be further extended where the task such as fire fighting effectiveness is regulated by conscious regulation of one's ANS measures through a process of affective biofeedback. Third, other measures to monitor autonomic nervous systems activity like heart rate can be used instead of GSR. Fourth, augmented reality systems can be used, with first responder gear to give a more enhanced feel for virtual reality based training.

References

- Berry, A. J., & Hilgers, M. G. (2004). Modeling First Responder Emergency Sensors for Virtual Training. *Proceedings of the 15th IASTED Modeling and Simulation Conference*, 234-239.
- Cohen, J. (1969). Statistical power analysis for the behavioral sciences. *London: Academic Press*.
- Eich, E., & Macaulay, D. (2000). Are Real Moods Required to Reveal Mood-Congruent and Mood-Dependent Memory? *American Journal of Psychology*, 11(3).
- Hall, R. H., Wilfred, L., Hilgers, M. G., Leu, M. C., Walker, C. P., & Hortenstine, J. M. (2004). Virtual Terrorist Attack on the Computer Science Building: Design and Evaluation of a Research Methodology. *Presence-Connect*, 4(4), <http://www.presence-connect.com>.
- Helander, M. (1978). Applicability of drivers' electrodermal response to the design of the traffic environment. *Journal of Applied Psychology*, 63(4), 481-488.
- Larsen, R. J., & Diener, E. (1987). Emotional Response Intensity as an Individual Difference Measure. *Journal of Research in Personality*, 21, 1-39.
- Leu, M. C., Hilgers, M. G., Agarwal, A., Hall, R. H., Lambert, T., & Albright, R. (2003). Training in Virtual Environments for First Responders. *Proceedings of the ASEE Midwest Section Meeting*.
- Mania, K., & Chalmers, A. (2001). The Effects of Levels of Immersion on Memory and Presence in Virtual Environments: A Reality Centered Approach. *CyberPsychology and Behavior*, 4, 247-264.
- Misra, A., Decker, A., Barker, B., & Hilgers, M. G. (2004). Modeling Chemical Events within Complex Buildings for First Responder Virtual Training. *Proceedings of the IASTED Modeling and Simulation Conference*.
- Moreno, R., & Mayer, R. E. (2002). Learning Science in Virtual Reality Multimedia Environments: Role of Methods and Media. *Journal of Educational Psychology*, 94, 598-610.
- Slater, M., & Steed, A. (2000). A Virtual Presence Counter. *Presence: Teleoperators and Virtual Environments*, 9, 413 - 434.
- Tichon, J., Hall, R. H., Hilgers, M. G., Leu, M. C., & Agarwal, A. (2003). Education and training in virtual environments for disaster management. *Proceedings of World Conference on Educational Multimedia, Hypermedia, & Telecommunications (EdMedia)*, 1191 - 1194.
- Witmer, B. G., & Singer, M. J. (1998). Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence*, 7, 225-240.

Acknowledgements

This work was supported in part by a grant from the U.S. Army's Tank-automotive and Armaments Command (TACOM, grant # DAAE07-02-C-L068); and University of Missouri – Rolla's Laboratory for Information Technology Evaluation (LITE) (<http://campus.UMR.edu/lite>), Virtual Environments and Object Modeling Labs (VEnOM) (<http://campus.UMR.edu/venom>), Virtual Reality and Rapid Prototyping Laboratory (VRPL) (<http://web.UMR.edu/~vrpl>), and Center for Technology Enhanced Learning (CTEL) (<http://campus.UMR.edu/ctel>).