Introducing Emotion into Military Simulation and Videogame Design: *America's Army: Operations* and VIRTE

Russell Shilling and Michael Zyda
MOVES Institute
833 Dyer Road, Room 254
Naval Postgraduate School
Monterey, CA 93943-5118
USA
Email: russ@shilling.us

E. Casey Wardynski
Office of Economic and Manpower Analysis
670 Cullum Road
U.S. Military Academy
West Point, NY 10996-1798
USA
Email: je2743@usma.army.mil

KEYWORDS
America's Army, VIRTE, emotion, videogames, sound design, audio, physiology, memory, learning, training

ABSTRACT

Emotion is a key component for sound design in movies and videogames. We believe that it is also a key component in virtual environments and simulation. The following paper summarizes work at the MOVES Institute's Immersive Audio Laboratory which demonstrates the emotional impact of sound in interactive media and also shows that emotionality evoked in a simulation has a positive impact on learning for events that occur in the simulation. Our research methods employ objective measures such as physiological recordings and memory recall testing rather than the more commonly used subjective questionnaires and surveys. It is our belief that these objective measures are more easily replicated and generalized to a wide variety of simulations and situations. We discuss our research in terms of the parallel development in the MOVES Institute of the videogame "*America's Army: Operations,*" which we use as an experimental test bed and tool. Applications of this research are discussed in terms of high-end simulation projects like the Virtual Technologies and Environments (VIRTE) program sponsored by the Office of Naval Research.

INTRODUCTION

Both in videogames and movies, the entertainment industry has long recognized the role of emotion in immersing viewers in the story portrayed on the screen. However, military simulation has focused almost entirely on improving the quality and accuracy of visual representations to the exclusion of producing an engaging and emotional experience. The philosophy has been that emotion is irrelevant and is not instrumental to the learning process.

We believe that emotion is a critical component of learning in virtual environments. We have been working with the entertainment industry to adapt techniques used in movies and videogames to produce systems that engage users on the visceral level as well as the intellectual. We also believe that it is critical to produce simulations that participants want to use and enjoy using. In addition, research conducted in our laboratory is showing that emotional arousal has a positive impact on learning, performance, and sense of immersion in virtual environments. This research has been aided by the development of *America's Army: Operations (AA:O)*, a professional videogame created and developed at the MOVES Institute at the Naval Postgraduate School in Monterey, CA. It is managed by the U.S. Military Academy's Office of Economic and Manpower Analysis at West Point. The current paper will summarize the techniques used in *AA:O* to produce emotion as well as the research conducted in parallel to determine the importance of emotion in training and to measure emotional response provided by different audio techniques. Our research differs from most previous lines of research, because we rely on objective rather than subjective measures for determining emotion and immersion in simulation.

AMERICA'S ARMY: OPERATIONS

America's Army: Operations (*AA:O*) is a multiplayer online first person shooter videogame developed in-house by the MOVES Institute. *AA:O* was built on Unreal's latest engine technology and designed by a group of professional game developers, simulation researchers, and graduate students. *AA:O* was not designed to be a training system, but rather a tool for introducing people to the goals and values of the U.S. Army. The development team's goal is to balance realism and entertainment in ways that are not seen in either traditional military simulation or videogame communities.

For instance, in order to produce realism in the game, the development staff visited over 19 Army bases during the construction of the game. The artists, level designers and programmers have fired weapons, participated in training exercises and taken detailed photographs, films and recordings of training facilities and weapons platforms. Actual soldiers were used in the motion-capture sequences. As a result of this attention to detail, weapons are modeled with extreme accuracy. Players must proceed through detailed reloading and jam clearance sequences. Weapon accuracy changes depending on whether the weapon is used in the supported or unsupported position. Accuracy is impacted by a combination of player experience, health, if they are under fire and whether the player is walking or running.

Figure 1. Screenshot from *America's Army: Operations*

Prior to commencing multiplayer games, participants must complete Army basic training, which is modeled directly from the actual training bases used by the U.S. Army. Training includes obstacle courses, rifle and sniper ranges, weapons instruction and the U.S. Airborne School, complete with 250 ft jump tower. Players also proceed through Military Operations in Urban Terrain (MOUT) training complete with the Multiple Integrated Laser Engagement System (MILES), the military's version of laser tag. Thus, in *AA:O*, even the simulators are simulated.

SOUND DESIGN AND EMOTION

One of the primary ways of introducing emotion into a movie, simulation, or videogame is through the proper use of audio cues and ambiences. In conversations with experts at THX, Lucasfilm Skywalker Sound, and Dolby, we were repeatedly told, "sound is emotion". A game or a simulation without an enriched sound environment is emotionally dead and lifeless.

The film industry has allocated significant resources to developing techniques for the design of sound effects and ambient sounds that evoke a sense of realism and manipulate the emotional response of the viewer. It is difficult to imagine that all sound heard in the battle scenes of *Saving Private Ryan* were added in layers after the film was shot. Yet, in the opening scenes depicting the Normandy invasion, the audio effects, including the actors' voices, are completely synthetic; added to the film after it was shot. The audio effects were spatialized using a surround-sound system to immerse the audience in the sound field.

Using this philosophy, the sound design for *AA:O* is incredibly rich and textured. Weapons sounds are modeled for a combination of sonic accuracy and emotionality. However, flat recordings of weapons fire were not used. Traditional recording and sound reproduction methods cannot capture the full dynamic range of high decibel weapons fire. A flat recording is not only emotionally flat; it also sounds unrealistic (Yewdall, 1999). Instead, flat recordings were mixed with other explosive sounds to compensate for the weaknesses of the reproduction media. Great care was taken when creating sounds to correspond with weapon animation sequences to make the sounds of jam clearance and reloading as accurate and compelling as possible. Since there is no tactile response involved in handling weapons in a videogame, it is important that the sound convey the feeling and emotion of handling the weapon in lieu of touch and feel.

In order for sound to impart emotion in a combat scenario, you need to capture the wide variety of sounds which are present in combat. Hence, we modeled the sounds of bullets whizzing by your ears, the sounds of bullet impacts in different types of materials (wood, metal, concrete, etc), and the sounds of debris resulting from bullet impacts. Thus, it is common to have bullets cracking by your ear and ricocheting or impacting on a concrete wall or wooden frame behind you. Meanwhile, the sounds of wood and concrete fragments shower down around your feet. Additionally, footsteps and other impacts have texture specific sounds associated with them. You hear your own footsteps and the footsteps of the players around you. We employed the movie sound designer's creed "see a sound, hear a sound" when we were designing the environment (Holman, 1997). Within the limitations of the game engine, if you see an action on the screen, you hear a corresponding sound. These details are crucial for immersing a player in the scene.

Finally, *AA:O* is a Dolby Digital certified game using the NVIDIA Nforce platform and is 5.1 and 6.1 compliant on non-Dolby sound applications as well.

Environmental effects are created using Creative Lab's EAX 3.0, an API used to induce numerous types of audio effects, including reverberation, occlusion, obstruction, and exclusion. The goal of the API is to mimic effects that approximate modeling the acoustics of rooms, buildings, and other audio environments. It does this without the expensive CPU requirements of actually modeling geometry and audio ray tracing. Future efforts in our lab will concentrate on using real-world interactive acoustic models to see how these impact users' perceptions of the environment.

The overall result of these many audio details is a highly immersive auditory experience which enhances the gaming experience and draws the player into the action. The first question becomes, can we prove that entertainment audio actually increases emotionality or is this folklore?

PHYSIOLOGICAL RESPONSE MEASURES

In order to determine the role audio plays in evoking emotion in videogames, we measured physiological responses during videogame play while subjects were playing a combat sequence with and without sound, using headphones or a THX certified 5.1 surround speaker system. Speakers and headphones were compared because of the hypothesis that a system employing a subwoofer might evoke more of an emotional response than a system using headphones alone (Shilling & Shinn-Cunningham, 2002). Temperature, Electro Dermal Response (EDR), and heart-rate measures were collected during action sequence game play. Results indicated increased physiological responses on all measures in the sound versus no sound condition. There was only an increased temperature response in the speakers versus headphone condition. These results clearly indicate that the audio component of a videogame or simulation contributes significantly to the emotional response of the participants (Scorgie & Sanders, 2002). The increased physiological response between speakers and headphones is probably due to the increased bass response derived from a subwoofer system that provides a more dynamic and "whole body" response to the sound. However, the effect may not be great enough to justify the increased footprint of a speaker-based system for simulations that must be placed in spaces with a small footprint (Shilling & Shinn-Cunningham, 2002).

EMOTION AND TRAINING

Given that audio design boosts emotionality, can we prove that emotionality actually is an important aspect of training in simulation? To answer this question we turned to physiological models of human memory. Adrenalin is a key hormone in emotional arousal and fight-or-flight responses. In animals, it has been shown that injections of adrenalin (a key hormone in emotional arousal) can enhance memory (McGaugh, 2000). It stands to reason that emotional arousal (in moderation) may also have a positive impact on human learning. After all, the limbic loop in the human brain modulates both emotional response and memory consolidation. The purpose of this research was to attempt to create a "virtual injection" of adrenalin to enhance learning in virtual environments.

Figure 2. Screenshot from memory experiment using *AA:O*

An experiment was conducted to observe learning differences in low-arousal conditions and high-arousal conditions (Ulate, 2002). *AA:O* was used as the virtual environment (Figure 2). In the low-arousal condition, participants wandered peacefully through a scenario, memorizing objects encountered while searching buildings on a mission to free POWs. High-arousal participants wandered through the same environment, but were required to fight their way through the scenario. Immediately after finishing the game, participants were tested for their memory of objects inside the buildings. An additional test was given 24 hours later. Results indicated that participants in the "high-arousal" condition were significantly better at encoding and recalling objects presented in the virtual environment immediately after experiencing the videogame and 24 hours post exposure. Thus, memories for events in a virtual environment are enhanced in situations where there are moderate levels of arousal. These findings also indicate that simulators used for mission rehearsal should not be dry, emotionless systems, but should elicit an emotional response from the user rather than a purely intellectual response. Further research is needed to determine if it is possible to over stimulate a user in a simulation, thus negating the positive effects.

CONCLUSIONS

What does this mean for the design of simulators and mission rehearsal systems? Since, emotional response has traditionally been an overlooked detail in the construction of simulations and virtual environments; we need to consciously consider emotion when designing simulations. Mission rehearsal systems which allow pilots to fly through terrain maps might be more effective if the pilot is engaged in combat while flying though the map. This needs further study.

These findings also impact research and development on other projects being pursued at the MOVES Institute. For instance, the Office of Naval Research sponsored Virtual Technologies and Environments (VIRTE) program envisions a multi-user multi-platform simulation for the Marine Corps. The simulation will include squads of Marines interacting in MOUT environments. Based on this research, we know that audio design is critical for creating the emotional context and arousal needed for optimal human performance in simulation. In fact, we have recently conducted task analyses to determine which cues are necessary for both accurate performance in MOUT situations and also for producing emotion (Greenwald, 2002). Our research has concluded that these tasks require more auditory cues than can be provided by most videogame engines or VE systems. For instance, MOUT tasks probably require accurate room acoustics and physics instead of approximations. We also believe that care must be taken to ensure that sounds like footsteps and body noises (clothing, breathing, etc) are modeled accurately in terms of the distances at which they can be heard.

Live voice communication is also a problem for high-end simulation that has not been adequately solved by the gaming industry. Traditional techniques used in gaming (VoIP) have latency rates exceeding 200 msec. One solution we devised is to combine the strengths of low-cost OpenAL and DirectSound3D systems with high-end servers used specifically for simulation. One such system is the AuSIM GoldServer. The GoldServer provides non-networked spatialized live audio over headphones with exceptionally low-latency (Krebs, 2002). Of course, this is only a solution for headphone-based systems.

During the upcoming year, we will continue to develop new strategies for creating detailed audio environments and implement our findings in our videogame work and in our simulation programs. At the same time, we will continue to validate our work with objective measures of performance. Ideally, the research we are conducting will benefit both the entertainment and simulation community by helping to create environments that are more immersive and emotionally engaging.

REFERENCES

Greenwald (2002). An Analysis of Auditory Cues for Inclusion in a Virtual Close Quarters Combat Room Clearing Operation. Master's Thesis. The Moves Institute. Naval Postgraduate School, Monterey, CA.

Holman, T. (1997). Sound for Film and Television. Boston, MA: Focal Press.

Krebs, E. (2002). An Audio Architecture Integrating Sound and Live Voice in Virtual Environments. Master's Thesis. The MOVES Institute. Naval Postgraduate School, Monterey, CA.

McGaugh, J.L. (2000). Research Shows the Role of Emotions and Stress in Forming Long-Lasting Memories. Brain Frontiers, Autumn (2-3)

Sanders, R., and Scorgie, R. (2002). The Effect of Sound Delivery Methods on the User's Sense of Presence in a Virtual Environment. Master's Thesis. MOVES Institute. Naval Postgraduate School, Monterey, CA.

Shilling, R.D., Shinn-Cunningham, B. (2002). Virtual Auditory Displays. Virtual Environments Handbook, Kaye Stanney, New York, Erlbaum.

Ulate, S. (2002). The Impact of Emotional Arousal on Learning in Virtual Environments. Master's Thesis. MOVES Institute. Naval Postgraduate School, Monterey, CA.

Yewdall, D. (1999). Practical Art of Motion Picture Sound. Boston, MA: Focal Press.

BIOGRAPHY

LCDR Russell Shilling, Ph.D. is a U.S. Naval Aerospace Experimental Psychologist and the Technical Director for Immersive Technologies in the MOVES Institute at the Naval Postgraduate School (NPS) in Monterey, CA. and the lead audio engineer and sound designer for *America's Army: Operations*. He joined the Navy in 1992 after completing his Ph.D. in experimental psychology at the University of North Carolina at Greensboro where he studied neuroscience and auditory psychophysics. Prior to arriving at NPS, he conducted research on virtual environments at the Naval Air Warfare Center Training Systems Division and the U.S. Air Force Academy.