NUWC-NPT Technical Document 12,118
10 December 2012

Serious Game and Virtual World Training: Instrumentation and Assessment

Jason H. Wong
Anh B. Nguyen
Lauren Ogren
Undersea Warfare Combat Systems Department

NAVSEA
WARFARE CENTERS
NEWPORT

Naval Undersea Warfare Center Division
Newport, Rhode Island

PREFACE

This report was prepared under Project No. DWAM10900 Basic, "Game-Based Training Systems," principal investigator Jason H. Wong (Code 2516). The sponsoring activity is the Office of the Deputy Assistant Secretary of Defense for Systems Engineering (ODASD(SE)).

Reviewed and Approved: 10 December 2012

Marie E. Bussiere

Marie E. Bussiere
Head, Undersea Warfare Combat Systems Department

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE (*DD-MM-YYYY*)	2. REPORT TYPE	3. DATES COVERED (*From – To*)
10-12-2012	Technical	

4. TITLE AND SUBTITLE

Serious Game and Virtual World Training: Instrumentation and Assessment

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Jason H. Wong
Anh B. Nguyen
Lauren Ogren

5.d PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Naval Undersea Warfare Center Division
1176 Howell Street
Newport, RI 02841-1708

8. PERFORMING ORGANIZATION REPORT NUMBER

TD 12,118

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING/MONITOR'S ACRONYM

11. SPONSORING/MONITORING REPORT NUMBER

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The goal of this report is to understand the current and future state of evaluating the effectiveness of training technology in terms of learning objectives, system architecture, and user interfaces. Training systems are of interest to system engineering for several reasons. As these systems are meant to mimic real-world platforms, training systems are ideal for exploring new technologies, tactics, and procedures in a laboratory. Additionally, these systems enable human-in-the-loop experimentation, so that system and warfighter performance can be examined as a system. Finally, training systems are moving beyond the warfighter into the acquisition workforce, so understanding how to create engaging material that captures student attention and increases retention will be critical.

Training is a large field and two specific trends will be examined: serious games and virtual world training. There will be a focus on the current methods of data collection and analysis that are available in current systems. These data can include system-generated data (e.g., user movement in the environment, chat logs, and object interaction) and human generated data (e.g., exam scores, eye tracking, and motion tracking). Additionally, this report will extrapolate the current trends and propose how training evaluations could be improved. The effect on cost and accuracy of these evaluations will be discussed, along with the major technical challenges that must be solved to reach the future state.

15. SUBJECT TERMS

Embedded Sensors   Eye Tracking   Virtual Reality   Interactive Simulation   Brain-Computer Interface
Digital Training   Virtual World Training

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Jason H. Wong
(U)	(U)	(U)	SAR	49	19b. TELEPHONE NUMBER (*Include area code*) (401) 832-4817

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39-18

20130325015

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

Figure **Page**

LIST OF TABLES

Table **Page**

LIST OF ABBREVIATIONS AND ACRONYMS

3-D	Three-dimensional
AI	Artificial intelligence
BIS	Bohemia Interactive Studio
BCI	Brain-computer interface
CAD	Computer-aided design
CONOPS	Concept of operations
COOPEX	Concept of operations exercise
CROWS	Common remotely operated weapon station
DARWARS	Game simulation trainer
DFAS	Defense Finance Accounting Service
DLI MTT	Defense Language Institute Mobile Training Team
DoD	Department of Defense
EEG	Electroencephalogram
FCT	Flood Control Trainer
FCVW	Federal Consortium for Virtual Worlds
GSR	Galvanic skin response
HRV	Heart rate variability
HSI	Human systems integration
HTTP	Hypertext Transfer Protocol
IEEE	Institute of Electrical and Electronics Engineers
IITSEC	Interservice/Industry Training, Simulation, and Education Conference
ITQ	Immersion Tendencies Questionnaire
LED	Light emitting diode
LFVT	Live fire virtual target
LMS	Learning Management System
MIT	Massachusetts Institute of Technology
N/A	Not applicable
NUI	Natural user interface
NUWC	Naval Undersea Warfare Center
PDA	Personal digital assistant
PE	Practical Exercise
PQ	Presence Questionnaire
SCORM	Sharable Content Object Reference Model
SOCOM	Special Operations Command
TMA	Target Motion Analysis
UCF	University of Central Florida
USC-ICT	University of Southern California-Institute for Creative Technologies
USSOCOM	United States Special Operations Command
VBS2	Virtual Battlespace 2
VNC	Virtual network computing
VW	Virtual world

SERIOUS GAME AND VIRTUAL WORLD TRAINING: INSTRUMENTATION AND ASSESSMENT

1. INTRODUCTION

The goal of this report is to understand the current and future state of evaluating the effectiveness of training technology in terms of learning objectives, system architecture, and user interfaces. Training systems are of interest to system engineering for several reasons. As these systems are meant to mimic real-world platforms, training systems are ideal for exploring new technologies, tactics, and procedures in a laboratory. Additionally, these systems enable human-in-the-loop experimentation, so system and warfighter performance can be examined as a system. Finally, training systems are moving beyond the warfighter into the acquisition workforce, so understanding how to create engaging material that captures student attention and increases retention will be critical.

Training is a large field, and two specific trends will be examined: serious games and virtual world training. There will be a focus on the current methods of data collection and analysis that are available in current systems. These data can include system-generated data (e.g., user movement in the environment, chat logs, and object interaction) and human-generated data (e.g., exam scores, eye tracking, and motion tracking). Additionally, this report will extrapolate the current trends and propose how training evaluations could be improved. The effect on cost and accuracy of these evaluations will be discussed, along with the major technical challenges that must be solved to reach the future state.

The current state of game-based and virtual world training is that the technology itself is still very new. Only in the past 10–15 years has computer graphics been able to render virtual environments in a relatively realistic manner, so much of the effort in this area has been towards improving the fundamental technology. Due to this, few resources have been available to devote towards evaluating the teaching effectiveness of these new training platforms. Some work has been done in collecting both system and human data, but studies are few and far between. Traditional games and virtual worlds have not been designed with data recording in mind, though some efforts have begun to automate this process of data recording. Human data collection has included likeability surveys, pencil-and-paper exams, and some preliminary work in measuring physiological responses to being immersed in the virtual environment. The foundation for thorough data collection is nearly there, but work needs to be done.

The future of assessing training effectiveness is one that lies in the development of additional technology, focused research, and a commitment to open systems. Virtual world and gaming technology will likely be used for more than just training in the future, and this additional development will serve to improve all aspects of these technologies, including funding studies that compare new training methods with traditional curricula. Improved human and system data collection will better inform different types of evaluations. In-world assessments will motivate students to perform well and more accurately discern how well a student absorbed or retained information, and improved electronic data records can show how new training methods have a broader impact on an entire organization. Better methods of collecting physiological data from students will automate human data collection, and these technologies will become more accurate and less intrusive in the future as well.

2. EMERGING TECHNOLOGIES FOR DIGITAL TRAINING

Simulations have been used in military training for nearly a hundred years, when wooden horses were used to train World War I soldiers in the skill of mounting and dismounting (figure 1). With the advent of computers, simulation environments have become virtualized, with the need for greater realism pushing the state of the art in computer graphics. One area where simulations have not seen rapid development is with instrumentation for data collection. These data must be collected in order to assess student performance and the efficacy of the training platform itself.

Figure 1. Wooden Horse Simulator from World War I

Better data collection and analysis from training systems to understand overall trainer-student performance is crucial in many ways. Within the traditional training domain, assessment is required for existing and new training systems. This is done to understand how new training systems bring improvements in student learning. Additionally, adaptive training systems are gaining momentum. Adaptive training systems are designed to adapt to the student – if a student is having difficulty with a certain concept, the system can detect the issue and spend extra time helping the student understand the idea. This kind of rapid feedback loop requires the proper instrumentation of the system to collect and analyze the right data.

Better instrumentation also supports the expanding role of warfighter trainers. Training systems are designed to mimic real-world equipment that the warfighter will use on the battlefield. This means that training systems are ideal for use as near-real-world platforms to explore new missions, technology, and concept of operations (CONOPS) in a laboratory environment. Any new capability is a combination of the system and the warfighter. Training systems enable developers to understand warfighter performance in the context of the systems as a critical component of the overall capability. A project at the Naval Undersea Warfare Center (NUWC) Division, Newport, RI, named the Analytic Attack Center is examining the issues of integrating a multitude of components, from training systems to data collection devices. Collecting data from both the training system and the warfighter can provide a complete picture of how missions, technology, and CONOPS interact.

Finally, training systems are being focused on more missions beyond the warfighter. Training considerations must now extend into the acquisition workforce. For example, the Defense Finance Accounting Service (DFAS) recently sought to use virtual worlds to demonstrate how training could occur for identifying and remediating unmatched disbursements. Traditionally, this training was primarily text-based and retention of information was low. Creating training that is more immersive will hopefully engage students, help them learn and retain more knowledge, and better support the organizational mission. Preliminary anecdotal observations with a virtual world training system for members of the acquisition workforce demonstrated students re-entering the content during their lunch breaks to review material; they felt that the virtual content helped them learn in an engaging fashion. Instrumenting these systems will provide an improved understanding of training effectiveness and return on investment.

2.1 TRAINING CAPABILITIES IN SERIOUS GAMES AND VIRTUAL WORLDS

Traditionally, virtual training has occurred through simulation, which was primarily focused on the environment and had little to do with other facets of training, such as the content. Instead of having to build an actual tank or enemy village, a computer could simulate the environment, and students would hopefully find the simulation realistic enough to gain value out of practicing tactics, missions, and procedures in a much safer and cheaper simulated environment instead of something with a physical presence. Recent trends have taken training in simulated environments beyond focusing on the environment itself. Instead, social aspects are being included, along with goals, points, and competition. These compelling and engaging aspects of the curriculum should aid in learning, increasing motivation and readiness to learn (Murphy, 2011).

Recently, an increased interest in serious games has pushed digital training beyond focusing on the simulated environment alone. Serious games share many properties of commercially available video games, such as scoring, points, and competition, but the primary focus of a serious game is to teach. Today's generation of warfighters is predominantly made up of digital natives (Prensky, 2001). They are comfortable with computers, and many have embraced the Internet, video games, and other digital technologies. Due to this, they are used to stimulating environments with clearly stated goals, such as scoring the most points, killing the most enemies, and reaching the target in the least amount of time. A serious game differs from a majority of the commercially available video games in that the primary purpose of a serious game is not to entertain. Instead, they aim to train students, encourage them to investigate a concept, or explore a scenario. Despite this, game concepts such as points are still present in serious games (Sadagic, 2011). Essentially, these products seek to add some dimensions of video gaming to a simulated training environment in the hopes of increasing motivation and engagement with training.

The other trend is towards using virtual worlds as the training platform. A virtual world is a multi-user environment where users are represented by digital avatars; these avatars can easily communicate and interact with one another and the surrounding environment. This social aspect enables teacher-student and student-student interactions, which are critical to learning (Swan, 2003). Additionally, users can be in geographically disparate locations. Since users log in as avatars and appear to be co-located with other instructors and students who could be very far away, there is a sense of unity among the training participants.

Additionally, multiple virtual world applications can occur in a single environment. Traditional simulations and many serious games only simulate a village or a single seaport; typically, a simulation cannot tie multiple geographic areas together. The virtual world can have a larger scope, allowing varied environments to be tied together. In any given virtual world, a naval exercise could be occurring in one sector while the next sector over contains a schoolhouse inhabited by hundreds of student avatars. An offshoot of these multiple applications is that information is persistent across users and time. Once information is constructed in the virtual world, users can log in at any given time and observe the newly created content, which enables instructors to collaborate and improve content with others while keeping logs of the evolution of the content for future reference. New instructors can then view the most up-to-date content while learning about the changes that were made to the curriculum. This is a powerful characteristic of virtual worlds because any changes that are made persist to all users across all sessions.

The final advantage of virtual worlds is the ease of content creation. Traditional high-fidelity computer simulations used for training can cost millions of dollars, and training scenarios and serious games must often be designed by the company that is also developing the environment itself. This creates a long design loop where changes to training scenarios must complete a lengthy cycle as the modifications travel between the schoolhouse and the developer. Certain virtual worlds can enable instructors and curriculum developers to create new materials quickly and easily, often inside of the virtual world environment itself.

It is important to recognize the distinction and similarities between virtual worlds and game engines like Unity 3D, Havok, and the Unreal Engine. Both can provide effective training capabilities in a social, immersive environment. However, gaming platforms are usually rigidly designed to deliver a single specific "serious game" that is heavily constrained in terms of user movement and actions. Virtual worlds, on the other hand, can be used to deliver that same serious game while including the flexibility to step out of the game into another immersive environment, as well as providing the ability to create or change the environment in real time. This document will focus on the similarities between serious games and virtual worlds, and references to "virtual worlds" will be at a more general abstraction that includes gaming platforms and other immersive 3-D environments for training.

2.2 METHODS OF EVALUATING TRAINING EFFECTIVENESS

Evaluating new training programs, platforms, and technologies has traditionally been categorized into four levels (Kirkpatrick, 1998). They are often visualized as building on top of one another (figure 2), but a higher level does not necessarily rely on a lower one. Level 1 is

Reaction—the simplest (and cheapest) level of evaluation, which is a measure of how much students enjoyed the training. This is most often measured through surveys after a student has completed a course. Level 2 is commonly referred to as Learning, which is a measure of how much knowledge had been gained and retained from training. Exams are the most common method of assessing training at this level and are traditionally used to determine the effectiveness of a new training program.

Level 3 is Behavior (also known as Performance), which is an indication of how training has affected on-the-job behavior. For example, a Level 3 evaluation of the effectiveness of a safety training video would involve examining how workers apply the material from the safety video to their everyday work duties (e.g., wearing a hard hat at all times). Finally, Level 4 is Results, which is measured at an organizational level. This evaluation can come in the form of fewer safety accidents, improved operational efficiency, and lower costs. These outcomes are often the highest-level goals of any training program, but they are often hardest to measure. All of the levels of evaluation are important in assessing the effectiveness of gaming and virtual world training technologies, as they are all dependent on one another. If students like the training, they will be motivated to use the training and, hopefully, they will learn the material well. This will impact their job performance, leading to improved return-on-investment by the overall organization. Collecting all of these data is the challenge and is an area ripe for future development.

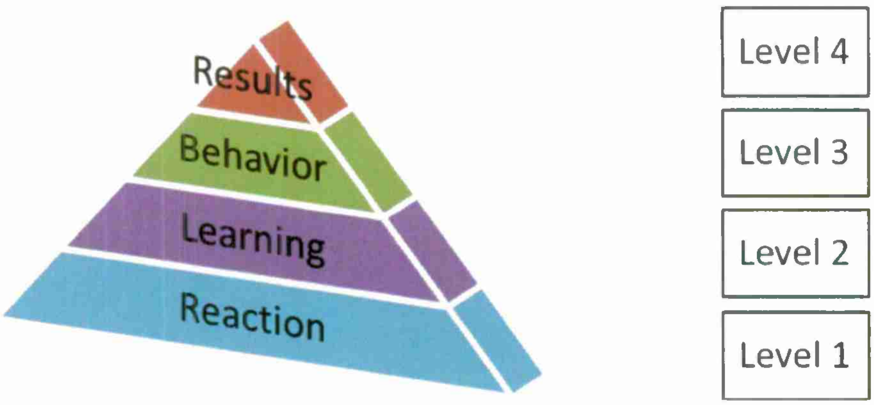

Figure 2. Four Levels of Kirkpatrick's (1998) Model for Evaluating Training Efectiveness

2.3 WAYS OF CLASSIFYING KNOWLEDGE

The goal of training is the transfer of knowledge, and it is important to understand how different kinds of knowledge may be taught differently. Bransford, Brown, and Cocking (2000) classified knowledge in three broad categories: declarative, procedural, and conceptual. Declarative knowledge is individual facts such as "The derivative of x^2 is 2x". Procedural knowledge is a series of steps designed to accomplish a particular goal, such as calculating a derivative. Finally, conceptual knowledge attempts to convey the underlying concept of a set of facts, such as understanding why the power rule for calculating the derivative of polynomials works (i.e., understanding the mathematical proof). Conceptual knowledge is rich in relationships and understanding as opposed to declarative knowledge, which is a set of loosely connected statements of fact. It should be noted that this is only one conceptualization of knowledge and is not definitive. Some researchers believe conceptual knowledge is simply a subcategory of declarative knowledge, while others argue for an affective knowledge category that includes emotions and other affective states.

Each of these different types of knowledge must be taught in different ways. Declarative knowledge is classically taught via memorization (e.g., with flash cards) and rehearsed through repetition while procedural knowledge is typically learned and then practiced multiple times. Conceptual knowledge is the trickiest to teach because it provides the foundation upon which other knowledge is derived. Therefore, simply memorizing conceptual knowledge is not enough. True comprehension and understanding is required, which makes concepts difficult to teach and to verify that students have learned.

2.4 CATEGORIZING TYPES OF DATA FOR COLLECTION AND ANALYSIS

Data that can be used for training evaluation can be organized in several different ways. One useful dichotomy is to categorize data as either system-based or human-based. System-based data are recorded by the interface and involves system parameters, such as physical movement of the avatar, text chat logs, number of enemies killed, and total score. These data are all a result of user interaction and can be used to determine how engaged a student seems to be in the environment. The data can also be used to evaluate the strategies students used to solve problems presented during training. One issue with system data is that the raw data (e.g., avatar movement) are abstract and often far from a useful metric that can be used to evaluate training effectiveness. This will be an important consideration in future systems.

Human data are recorded from the students (and instructors) via surveys, exams, behavior observation and recording, and even physiological methods such as heart rate variability and electroencephalogram (EEG), or brain wave, recordings. These data recording methods can support measures of engagement, attitude, immersion, and initial learning and later retention. All of these are important to evaluating the effectiveness of these new digital training applications, as the user must be engaged and efficiently learning to ensure a positive return on investment.

A combination of system and human data will provide an overall picture of new training technologies. This can facilitate a more thorough evaluation of training effectiveness. However, it must be understood that evaluating the usefulness of training is not a simple dichotomy of effective versus ineffective. Many tradeoffs can be made between initial investment cost, recurring maintenance cost, student time commitment, instructor effort, amount of student learning, and a multitude of other factors. This report does not attempt to rank these factors by importance. Instead, this report examines what kind of data can be collected to support a holistic understanding of a training technology that can be used to make informed decisions.

3. CURRENT STATE OF GAMING AND VIRTUAL WORLD TRAINING TECHNOLOGY

A literature review of the military serious game and virtual world-based training environments was conducted to identify projects relevant to the current landscape. Focusing on projects that have undergone a training effectiveness evaluation, the initial literature search produced a variety of projects that spanned across military branches, from the Army Virtual Battlespace to the Navy Immersive Target Motion Analysis (TMA) training modules and the Air Force 367th Training Support Squadron Simulator. The current state of data collection and analysis from the serious games and virtual world platforms is reviewed, along with human and system data collection and analysis abilities.

3.1 SURVEY OF GAME-BASED AND VIRTUAL WORLD-LEARNING PLATFORMS

Gam e-based and virtual world technologies are moving very rapidly with noticeable advancements in the current landscape. There are a variety of Department of Defense (DoD)-related projects that could be mentioned. However, due to the sheer number of these projects, a subset was chosen and is discussed below because these projects have had large impacts on the training community and have also produced useful data about training effectiveness (table 1).

Table 1 below identifies 19 projects and categories, the work based on the branch of the military where the project is used, the kind of project (serious game or virtual world), the type of knowledge being taught by the system, whether the serious game or virtual world was collecting human data, system data, or both, the kinds of data collected from either the serious game or virtual world, and the Kirkpatrick Level where the published evaluation took place.

Table 1. Serious Games and Virtual World Training Projects Surveyed

Project	Branch	Game or Virtual World	Knowledge Category	Human or System Data	Data Collected	Kirkpatrick Level	Reference Keyword
367th Training Support Squadron Simulator	Air Force	Game	Conceptual	Human	Likeability Surveys	1	Beus
America's Army: Live Fire Virtual Target	Army	Game	Procedural	N/A	N/A	N/A	America's Army/LFVT
Common Remotely Operated Weapon Station	Army	Game	Procedural	N/A	N/A	N/A	America's Army/CROWS
Danger Close: Virtual Experience Immersive Learning Simulation	Army	Game	Procedural	Human System	Likeability Surveys Game Performance	2	McFadden
DARWARS Ambush!	Army	Game	Conceptual	Human	Questionnaires Performance Evaluations Psychometrics	2	Raybourn

Table 1. Serious Games and Virtual World Training Projects Surveyed (Cont'd)

Project	Branch	Game or Virtual World	Knowledge Category	Human or System Data	Data Collected	Kirkpatrick Level	Reference Keyword
ELECT BiLat	Army	Game	Procedural	System	Game Performance	2	McAlinden
Enterprise Resources Planning Training Simulator (ERPTS)	Navy	VW	Procedural Declarative	Human	Pre-/Post Tests in-World Comments Knowledge Exams	2	McKay
Game After Ambush!	Army	Game	Conceptual	N/A	N/A	N/A	Stephens
Game Technology in Institutional and Unit Training	Army	VW	Conceptual	Human	Pre/Post Tests Exam Scores Questionnaires	2	Goldberg
Immersive Target Motion Analysis (TMA) 3-D Training Modules	Navy	VW	Conceptual	Human	Tests in-World Exercises Usability Survey	2	Nguyen
Javelin Weapon System Trainer	Army	Game	Procedural	Human System	User Feedback	1	Cubic
Moral Combat	Army	Game	Conceptual	Human System	After Action Review	1	Dempsey
Nuclear Biological Chemical Trainers	Army	Game	Procedural	N/A	N/A	N/A	Berry
Rapid Response Missile Simulation	Army	Game	Procedural	N/A	N/A	N/A	Berry
Tactical Iraqi	SOCOM	Game	Procedural	Human System	Likeability Surveys Game Performance Post Mission Survey	2	Surface
Troop Warrant Officer Course	Canadian Army	Game	Conceptual	Human	Course Pass Rates	2	Roman
VESSEL: Damage Control Trainer	Navy	Game	Conceptual	Human System	Usability Surveys Game Performance Battle Station Performance	3	Murphy
Virtual Battlespace 2 Practical Exercises	Army	Game	Procedural	Human System	Student Exams Exercises Questionnaires Interviews	2	Topolski
Virtual-Integrated MOUT Training System	Army	VW	Conceptual	Human System	Likeability Survey Game Proficiency Measures Knowledge Exam	2	Knerr

One widely recognized and accepted game-based simulator in use by the U.S. Army is called Virtual Battlespace 2 (VBS2), created by Bohemia Interactive Studio (BIS). The power of this simulation platform is the BIS Real Virtuality game engine that allows developers to rapidly create scripted scenarios with custom terrain, characters, weapons, and vehicles. This fosters a community of users that are leveraging the customizable and extendable power of this simulator (Bohemia Interactive Simulations, 2011). The ease of content creation enables scenarios to be created quickly based on recent training requirements, which is a major benefit (figure 3).

Figure 3. Virtual Battlespace 2 Simulation Engine

VBS2 has served as a game-based training platform for the creation of scenarios specifically designed to be measured for training effectiveness compared to traditional training methods (Goldberg, 2010). The effectiveness of the game technology was investigated by using VBS2 to create two U.S. Army related practical exercises (PEs): Route Reconnaissance and Vehicle Tactical Movement.

The project measured performance from the control group that utilized the current terrain board exercises and an experimental group that utilized VBS2 and an existing game simulator called DARWARS Ambush! that also uses the Real Virtuality game engine. The assessment was done with pre- and post-tests along with exercises, questionnaires, and interviews with the students. No further assessment on the job was completed, so this project attained a Kirkpatrick Level 2 for training effectiveness evaluation.

Results showed that the VBS2 and the game-based simulator were favored by the instructors and students over the existing terrain board exercises because of their tactical realism and representation of operational risks, engaging students, and fostering a teamwork environment through communication. Even though both groups improved from the pre- and post-test, VBS2 was proven more effective that the terrain board exercise in the Tactical Vehicle Movement PE. Students in the game-based training group also increased the unit process and unit cohesion ratings after going through the VBS2 training.

Another recent game-based simulator that the U.S. Special Operations Command (USSOCOM) is currently using to train military personnel is called the Tactical Language and Culture Training System, created by Alelo, Inc. This training system teaches languages including Iraqi Arabic, Afghanistan Pashto and Dari, French and Indonesian, along with cultural knowledge from these regions (Alelo Inc., 2011). USSOCOM conducted an experiment through Alelo, Inc. that examined the effectiveness of the training and the trainees' views of the system (figure 4).

Figure 4. Screenshot from the Tactical Iraqi Simulation

For this experiment, researchers focused on the "Tactical Iraqi" simulation with several groups of students. The first Marine group went through the Iraqi Arabic language and culture training through the Defense Language Institute Mobile Training Team (DLI MTT) and then with the Tactical Iraqi training system. A second Marine group only utilized the "Tactical Iraqi" training system. All groups had approximately 40 hours of training time. Results show that the DLI MTT group who also went through the tactical Iraqi system had the greatest increase in knowledge. However, all groups, including the Tactical Iraqi-only group, had statistically significant increases in language and culture knowledge. The perception of the usability and the satisfaction level from the participants of the Tactical Iraqi training system were positive as well. These two elements of the assessment indicate a Level 2 Kirkpatrick evaluation. The conclusion was drawn that while it is feasible to teach only with the Tactical Iraqi training system, data show greater learning improvement by utilizing the existing training along with the Tactical Iraqi game-based simulator (Surface, Dierdorff, and Watson, 2007).

In 2009, the Navy Flooding Control Trainer (FCT) was developed in order to supplement classroom instructions and to reduce the demand on the training resources at the Navy Recruit Training Command (figure 5). FCT was developed with an open source Delta3D game engine, and the focus was to provide a 3-D environment of a simulated *Arleigh-Burke* class destroyer for individual flooding control training. A Kirkpatrick Level 2 usability study and a validation study were done for this trainer. Based on the information from these studies, nobody in the virtual training group went to the wrong deck, demonstrating that they have familiarized themselves during the FCT compared to 33% of the control group who went to the wrong deck. 36% of the treatment group described the leak correctly versus 16% of the control group, and only 28% of the treatment group entered the flooding compartment without appropriate personal protective equipments compared to the 67% in the control group. Statistics show that the FCT provides sailors with the ability to acquire a strong ship damage control foundation for the rest of their Navy career (Murphy and Hussain, 2009).

Figure 5. Screenshot of the Navy Flood Control Trainer

3.2 SYSTEM DATA COLLECTION

Most work in the development of serious games and virtual worlds has been on the fidelity of the system. The realism of graphics and ability to create immersive content is imperative to keep users engaged. In this area, developers have been successful. One area that is lacking is the area of system data collection. The ability to extract wanted information out of these environments is a difficult process. Of the projects surveyed, 8 out of 15 serious games and only 1 out of 4 virtual worlds had some method of collecting system data, and each game or world has its own method to collect data. There are some preliminary efforts to create common methods of data extraction across platforms while others are pioneering methods to analyze the data that are extracted from the systems.

There are various existing data collection methods that the DoD is utilizing, including the Learning Management System (LMS) that can conform to the Sharable Content Object Reference Model (SCORM). SCORM is a digital library consisting of training materials ranging from system documentation to maintenance and operation procedures for warfighters to learn about various pieces of equipment (Friess, 2011). The key innovation of SCORM is that all compliant training materials adhere to a defined standard, making these materials interoperable.

This allows for a common layer of communication between different learning platforms. Therefore, data from various learning systems can be interoperable, enhancing the combined value of using different platforms for different needs.

TechWizards, Inc. recently developed and integrated a SCORM-compliant module into the OpenSim virtual worlds, allowing an LMS to monitor and track student progress. OpenSim is designed as an open-source version of SecondLife, so it has properties and features that mimic the closed-source, commercial product. Students log in to the virtual world and are presented with a virtual personal digital assistant (PDA) that shows the student's progress in real time. This virtual PDA also tracks the completion of the available training in the virtual world, and these data are fed into a web-based LMS. This LMS can then aggregate the students' completion from various sources, including the virtual world. Additionally, the LMS allows anyone with the right credential access to log in and monitor the progress of students in real time. Metrics such as time of completion, engagement statistics, and competency evaluation can be gathered for students' feedback. This kind of technology has also been developed as a wrapper around the SCORM system to collect similar data (Gallagher and Berking, 2011).

Other individual projects have looked at collecting specific metrics from virtual worlds like SecondLife, such as avatar motion (La and Michiardi, 2008). The authors wanted to examine user mobility and understand if movements in SecondLife mimicked those of real-world human movement patterns. The researchers built a customized crawler avatar to monitor user movements and recorded snapshots of users' positions every ten seconds to develop a movement map. Three venues were examined: (1) Apfel Island (German User Starting Ground), (2) Dance Island (Indoor Dance Club), and (3) Isle of View (Outdoor Organized Event). As with real-world interests, users tend to cluster around a central point of interest. For example, the indoor dance club showed users interacting for 100 seconds or more. This crawler avatar was able to observe other avatars in a natural way, demonstrating that this is an effective method of data collection inside virtual worlds.

Dixit and Youngblood (2009) attempted to visually represent data that are gathered through user monitoring in SecondLife. The aim was to provide insight into patterns and user behavior within virtual environments and to improve the user's interactive experience through data mining results. The researchers logged avatar movement to discover what actions people were commonly taking and to find patterns in player behavior. They recorded five separate user behavior methods, varying from when a user was lost to fine avatar control to an indication of learning. The trace processing algorithms could be run very quickly, which is beneficial when analyzing large amounts of data. Therefore, these algorithms are close to analyzing user behavior in real time and are applicable not only to games but to training/learning scenarios within virtual environments as well.

3.3 HUMAN DATA COLLECTION METHODS FOR KNOWLEDGE ABSORPTION

A majority of studies examining new training effectiveness have been done in the form of user surveys and exam scores—Levels 1 and 2 of Kirkpatrick's levels of evaluation rubric.

These are often the easiest and most cost-effective methods of evaluating training effectiveness. However, likeability surveys are rarely reflective of how much material was actually learned, though it is useful to understand what students did and did not like about a new training method. Exam scores are slightly more useful, but even this seemingly simple method of evaluation can be tricky.

A training effectiveness study can use exams already written for the current curriculum, but they may not test the skills that the new training technology is designed to teach. Therefore, scores from the already-existent exams may not show improvement. On the other hand, new exams can be written to allow students to demonstrate their newfound knowledge, but it can be difficult to remain unbiased when writing and grading the custom exams. There is also the criticism that, if students score well on a custom exam, all that is proven is that the technology teaches the desired information. The larger effects—whether the additional knowledge is useful in a broader context—remain unexamined.

Broader questions of improved on-the-job behavior and greater organizational benefits (Kirkpatrick Levels 3 and 4) are rarely examined. This is in part due to the difficulty of tracking students beyond the classroom into the workplace. Additionally, it is difficult to isolate the effects of training on performance and organizational benefits when other factors (e.g., employee incentive programs) could also have an effect. Of the various serious games and virtual world training environments reviewed, three projects only used surveys to evaluate the training platform. Eight projects conducted a Level 2 assessment and examined student performance through a combination of exams, game performance, or overall course performance.

A novel method for assessing student knowledge through virtual environments is performing knowledge assessments in the virtual environment. Many studies of learning and long-term memory have provided evidence for the theory of context-dependent learning (Godden and Baddeley, 1975). This research proposes that people learn information together with context, and that memory retrieval works best if someone is in the same state during learning and retrieval. For example, a student who studies in a noisy environment scores higher on an exam if the exam room is also noisy (Grant et al., 1998). This holds true for internal mood states (Eich and Metcalfe, 1989) and even if students learn lists of words underwater or on land (Godden and Baddeley, 1975).

These principles can be applied to game-based and virtual world training by designing assessments that occur in the virtual environment. An added benefit is that curriculum designers are no longer tied to multiple choice, fill-in-the-blank, or essay questions. Instead, games can be made that motivate students during the assessment to get the high score, most enemy kills, or compete to beat their fellow students. This higher motivation should overall increase learning enjoyment and possibly increase overall knowledge absorption as well.

This method of in-world assessment has not often been used in any widely used game-based and virtual world training systems to date. However, data are currently being collected for a pilot study using content developed jointly between NUWC Division Newport and the Submarine Learning Center. This content is hosted in the SecondLife virtual world for use by the Naval Submarine School, Groton, CT, in teaching techniques. The virtual content has been structured as a series of exhibits along a trail, with each "stop" designed to teach new concepts, built upon those learned previously (Nguyen, 2010).

Each exhibit on the trail is accompanied by an interactive assessment, taking the form of a game. For one game, students are presented with a multiple-choice question, and they must run to the platform corresponding to their chosen answer (figure 6). After a short time, the instructor presses a "Check Answer" button, and the platforms with incorrect answers disappear, dropping student avatars standing on those platforms into the water below. Answers chosen by each student can be hidden from the class until the last second, so individual student responses can be recorded and graded. The addition of the platforms and dropping avatars into the water are unnecessary to the actual assessment, but anecdotal evidence has shown that students become immersed in the game and are focused on getting the right answer. However, more research will need to be done in this area to formally examine the effect of motivation on assessment outcomes.

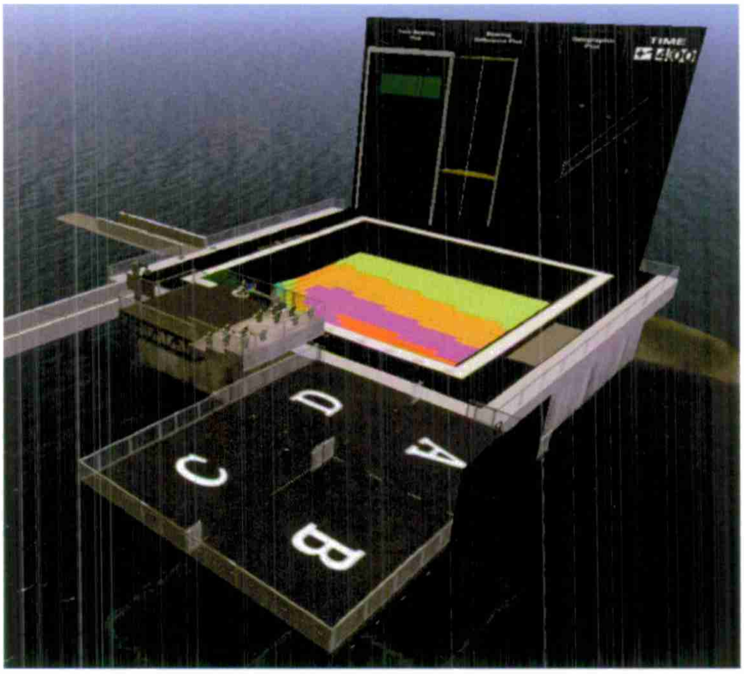

Figure 6. Screenshot of the In-World Assessment from Nguyen (2010)

3.4 HUMAN DATA COLLECTION FOR IMMERSION

Immersion in a virtual world can have major implications on human performance and interactions (Slater, Linakis, Usoh, and Kooper, 1996). One must understand how the immersion will affect a person's reaction to a situation presented to them within a virtual environment. It also is important to understand the extent to which a person must be immersed and engaged to achieve the full benefit of their interaction in the virtual world.

Immersion describes the entirety of the user's experience within virtual reality or the virtual world (Coelho, Tichon, Hine, Wallis, and Riva, 2006). Lombard et al. (2000) separated immersion into six subsections to further describe each facet of immersion. The facets included:

- social richness,
- realism,
- transportation of self,
- inclusion in environment,
- social actor within medium, and
- medium as a social actor.

Of these, it is often felt that transportation and the feeling of being there, best describe the essence of immersion (Schuemie, van der Straaten, Krijin, and van der Mast, 2001). The feeling of being there can sometimes seem general and, as a result, others have sought to further describe what being there means. A person may have a sense of personal immersion (feeling part of the virtual environment), social immersion (other people or things existing within the virtual environment), and environmental immersion (reaction of the virtual environment to the person) (Heeter, 1991).

The above components of immersion can be consolidated into three major variables that make up the overall experience. The components that can be manipulated to alter the feeling of immersion are (1) inclusion of self, (2) content, and (3) engagement. Self-inclusion is characterized by a person's ability to perceive themselves as being incorporated in and interacting with the surrounding environment (Witmer and Singer, 1998). Engagement or involvement is the effect on a person when they focus their attention on a defined set of stimuli, in this case a virtual environment. Content describes the information that is being presented to the user. A level of immersion can be constructed through increasing or decreasing each variable accordingly. It is hypothesized that depending on the learning task, different degrees of immersion must exist until the student reaches the point at which their performance levels off.

Currently, the most common method to measure immersion is with the use of questionnaires. Two questionnaires have been developed and a version of each is included in the appendix. The first is the Presence Questionnaire (PQ), which measures how much immersion individuals experience while using the environment or the degree to which immersion is experienced (Witmer and Singer, 1998). The second is the Immersion Tendencies Questionnaire (ITQ), which measures the tendency of individuals to become immersed in virtual environments. This serves to understand how likely it is for someone to be immersed in another activity in general, which provides a baseline level of immersion for each person. These questionnaires provide only a single data point for how the person felt during the entirety of the observation period. In addition to this, the use of questionnaires is a highly subjective measure of a person's experience (Darken, Bernatovich, Lawson, and Peterson, 1999).

3.5 MOTION-BASED USER INTERFACES FOR VIRTUAL ENVIRONMENTS

An interesting technology that is rapidly developing is new input methods for virtual environments. While there have been great advances in the sensory fidelity of software that creates virtual 3-D spaces, little work has been done on the realism or ease of use of the actual user interface. Research is being done to examine different possible user interfaces. One is to use a controller that lends itself to controlling virtual environments more naturally. This could be done with a commercially-available product, such as a game console controller. The idea is that an intuitive controller layout, combined with improved button placement and functionality, could greatly improve operator efficiency (Ogren and Wong, 2010).

In addition to better-designed controllers for virtual environments, innovative technology is being developed to try to remove the controller entirely. One method of doing this is through motion tracking. The Microsoft Kinect system is meant as a companion to the Microsoft Xbox gaming platform. The device is comprised of two cameras and four microphones. This is paired with software that can recognize up to two bodies and track them with six degrees of freedom. It is similar to the Nintendo Wii in that it can track arms and hands as they move around, but unlike the Wii, the body is the controller. This allows the body to be tracked, so a soccer game can be controlled purely using the legs and feet.

The Microsoft Kinect could be used to control any virtual 3-D environment intuitively, creating a natural user interface (NUI). The goal of a NUI is to be designed so that users can walk up to a system and immediately begin using it. Therefore, the user feels like a "natural." The Microsoft Kinect has been used to control many Xbox 360 games since its release, so there are many commercial examples of making gestures to accomplish certain tasks. Sreedharan, Zurita, and Plimmer (2007) sought to implement a NUI for controlling avatars in the SecondLife virtual world using the Nintendo Wii controller. They found some success in creating gestures that were natural to users, but many gestures (e.g., laughing, standing up) did not have a direct analog to actions you could make with the Wii controller, so there was little consensus from participants on how to accomplish those basic tasks. As the Kinect continues to mature in capability, there are likely more projects to come that integrate the Kinect into increasingly complex 3-D environment scenarios.

3.6 HUMAN AND SYSTEM DATA ANALYSIS

Analysis of the human and system data collected from training systems is currently restricted to students in a single class. Game performance, test scores, and survey data can be applied to individual students and averaged at the class level, but this is the main extent of data analysis. One reason for this is the limited scope of many training evaluations—resources only allow for analysis at the classroom level. It is difficult to understand the impact of training on on-the-job behavior or larger organizational goals without longer-term data. Oftentimes, these data do not even exist or are difficult to quantify. For example, on-the-job performance is sometimes documented only as a written performance review. This can be difficult to turn into quantitative data that can be analyzed. Of the 18 projects surveyed for this report, only one conducted an analysis of on-the-job behavior.

System data analysis can be especially cumbersome because of the difficulty in translating raw system data into higher-level concepts. For example, avatar position over time is a stream of raw system data that a serious game or virtual world could provide. However, turning that positional data into something useful (e.g., that the student is visiting the right people in the right order) requires complex pattern recognition and data mining algorithms that are currently in the early stages of development. Data mining large datasets is a topic of increasing interest to any organization with huge stores of data, and these algorithms will get better with time, though the degree that they can be generalized across fields and organizations is unknown.

4. FUTURE TRENDS OF GAMING AND VIRTUAL WORLD TECHNOLOGY

As the serious game and virtual world technology trend progresses, private companies, universities, and government entities are teaming up to understand the future directions of these technologies. Providers such as Linden Labs (creator of Second Life), Teleplace, VastPark, and others are pushing the boundaries of virtual world technologies. Universities, including the University of Southern California Institute for Creative Technologies (USC-ICT), University of Central Florida (UCF), and military laboratories, including NUWC Division Newport are on the forefront exploring these technologies. These emerging technologies will continue to grow rapidly as the academic, industry, and government organizations continue to collaborate and develop better and more functional systems.

Currently, there are many collaboration venues including various conferences such as the Interservice/Industry Training, Simulation and Education Conference (IITSEC), Defense GameTech Users Conference, Federal Consortium for Virtual Worlds (FCVW) Conference, and the Advanced Distributed Learning Implementation Fest conference that allow researchers to showcase their work and to obtain feedback and gather partners across the spectrum. As these technologies become more readily available to new users and become more open, they will attract innovations from collaborative projects among researchers and developers. The virtual world usage will grow as the collaborative infrastructure becomes more readily available for users to access and more training applications take advantage of the unique capabilities that virtual worlds offer.

As the serious games and virtual world technologies improve, they will be used for purposes beyond training, such as collaborative engineering or human systems integration (HSI) research in order to fulfill the gaps presented by the current analysis tools. Currently, there are various emerging technologies that satisfy many of the technological requirements like providing large-scale gaming platforms, having agent-based modeling for crowd analysis, or providing an Analytic Attack Center, virtual classrooms, virtual laboratories, and a robust AI engine to aid in the training environment.

As educational organizations feel increased pressure to provide more services with fewer resources, continual technological improvement will drive down the cost of creating and maintaining the virtual environments that host game-based and virtual world-learning platforms, and the methods used for assessing student performance from a system, knowledge, and human perspectives will improve as well. The current state of assessment in virtual environment training platforms shows that these assessment methods lag behind other areas of development. As the rapid pace of innovation in content creation tools begins to slow, greater attention will be paid to developing assessment methods. In parallel to this, there is a strong consumer-level push for physiological measurement, so technologies such as cost-effective EEG recording and eye tracking will allow for less-invasive human performance measurement.

4.1 OPEN SYSTEMS THAT SUPPORT SYSTEM DATA COLLECTION

It is important for serious games and virtual worlds to provide as much openness as possible in order to provide flexibility to anyone who wishes to utilize the platform. Virtual world technology is meant to be a versatile platform that can support a wide variety of user cases, so an open architecture is critical in allowing people to be creative in using the virtual world. With the current trend moving toward open-source software for easy access to various data, the format of the data is also converging toward common standards and specifications like the Sharable Content Object Reference Model (SCORM). The P1828 Institute of Electrical and Electronics Engineers (IEEE) virtual world standard working group is currently developing standards for importing and exporting data across virtual worlds as well. In order to track the changes in the data, documents, and reports, users are now using various SCORM-compliant LMS as a centralized, scalable platform to administer these changes and to track training programs.

One such open virtual environment system is the OpenSim virtual world environment, a 3-D open source server for multiple users and operating systems. OpenSim utilizes the core of SecondLife messaging protocols and can be accessed with any SecondLife viewer (Overte Foundation, 2011). Since one of the goals of OpenSim is to be extensible, it is excellent at allowing various system connections for data gathering and monitoring. Along with the ability to support multiple simulators at once, OpenSim can also support multiple clients and protocols, allowing developers to create innovative virtual world clients such as the one created by TechWizards. Since OpenSim allows clients to create 3-D content in real time and provides in-world scripting through various programming languages, monitoring and gathering systems can be directly built into the virtual world and then export this information to an offsite-hosting provider.

Along with OpenSim, the SecondLife virtual world has the ability to connect to various websites through Hypertext Transfer Protocol (HTTP). This capability allows users of these virtual worlds to input data and gather outputs from various sources including web servers, databases, and any external simulators that can be connected to a web interface. Another capability that these virtual worlds can support is the ability to connect remotely to any available simulator and workstation through virtual network computing (VNC). This is an extremely powerful tool because it allows users from remote locations to access the computer systems that they need through the virtual world.

There are also various other third party tools—Second Inventory, Imprudent Viewer, and CopyBot—to export and import data between the different virtual worlds. These tools allow users to back up the virtual contents, including objects and scripts, and some even allow users to restore the content in a separate virtual world environment. The information can then be retrieved and metrics can be collected if the data are lost or irretrievable. This provides an important avenue for content backup and migration and contributes to the openness of virtual worlds in allowing for content import and export.

4.2 INNOVATIVE METHODS FOR EVALUATING HUMAN LEARNING

Training assessment will continue to occur, especially at Kirkpatrick's Levels 1 and 2 (reaction and learning), since those are simple and traditional methods of assessment. New ways of assessing knowledge will continue to be developed as well, such as creating more engaging assessments in-world. This idea, while discussed as a current state technology in section 3.3, requires a great deal of technological and pedagogical development. The technology to create in-world engaging assessments, such as competitive games, must be introduced into the virtual environment. For example, the SecondLife virtual world does not have built-in mechanisms to create games, keep track of points, etc. This must all be developed before a production-quality game-based assessment can be implemented. In addition to the technology, pedagogical principles must be understood as to what makes for a successful assessment and to ensure that the addition of a competitive gaming aspect does not interfere with the primary goal of ensuring student learning.

Another in-world method of assessing student knowledge is to utilize the in-world building tools. It is often said that teaching material to others is one of the best ways to learn the material. Therefore, students could be tasked to use the virtual environment to build a demonstration of a concept. This is similar to creating a diorama in science class, except that virtual tools are being used. This is an idea that has not been implemented in any training system, but it is an example of using the distinctive features of the virtual world technology to improve assessment methods.

System data can also serve to understand how well students are learning. System data can include avatar-avatar interactions and avatar movement. While these low-level pieces of data are not useful in themselves (e.g., "Student Avatar chatted with Manager Avatar"), these data could be useful in a wider context. For example, a game instantiated in a virtual world could be designed to teach a student the proper channels for a travel request order to flow through. Students, as their avatars, would then have to walk from office to office, getting the proper authorizations. Knowing who the student talked to and if they were unsure of where to go next indicates how well students understand the process. Therefore, collecting system data is important, but context must be given to the raw data to be able to evaluate training.

Advanced levels of Kirkpatrick's assessment rubric will be fleshed out in the future as well. Kirkpatrick's Level 3 of training effectiveness looks at the extent that on-the-job behavior is affected by training, and Level 4 examines high-level organizational outcomes, such as cost savings and improved ship operations with fewer accidents. In order to perform a training effectiveness evaluation at these levels, more data must be collected from individuals as they mature from students into warfighters. These data must be input into a series of open, accessible databases that track student performance throughout their coursework, performance ratings on the job, and other similar data.

This kind of tracking requires a large, fundamental change within the military information technology infrastructure. New policies must be put into place that allow for this detailed level of tracking yet seek to protect the privacy of the warfighter. In addition to policy changes, huge technological changes must occur. A majority of exam scores, performance reports, and other necessary data are stored on paper. Digitizing this information and storing it in easily accessible but secure repositories would require a major investment in new information technology.

While there are barriers and costs associated with evaluating training effectiveness at a higher, outcome-based level, there are also extraordinary benefits. Relationships between performance in certain classes in the schoolhouse and future on-the-job performance can be explored and quantified more deeply. Longitudinal studies can be carried out to determine how the implementation of a new curriculum affects future Fleet performance. Funding decisions can be made based on data instead of on intuition. Oftentimes, a new learning technology is used because it appears to be a good idea. However, tracking warfighters' performance throughout their careers now enables the collection of data that can provide evidence to what training methodologies and principles are actually effective.

4.3 IMPROVED HUMAN PHYSIOLOGICAL DATA COLLECTION METHODS

Several possible physiological responses could be measured to support training effectiveness from students. Currently, most of these methods reside in the laboratory due to their invasiveness and expense. In the future, though, physiological data collection will become more accurate, cheaper, and transparent to the student, which will be crucial to the acceptance of the recording technologies. Four of the more promising physiological data collection technologies are the (1) electroencephalogram (EEG), (2) galvanic skin response (GSR), (3) heart rate variability (HRV), and (4) eye tracking. Example technologies can be seen in figure 7.

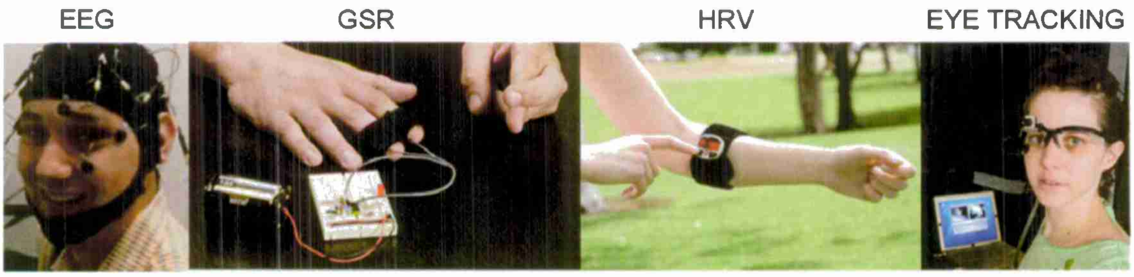

Figure 7. Different Physiological Measures for Collecting Human Data

4.3.1 Electroencephalography (EEG)

One of the more common methods of physiological data collection is through electroencephalography or EEG. The EEG signal is made up of the electric responses of neurons during brain activity, and these "brain waves" are categorized by the frequency of spikes that occur within the signal, known as bands. Each band can be thought of as corresponding to a human's cognitive processing of external information (Luber, 1995). For example, delta waves are seen primarily during slow wave sleep, while gamma waves are indicative of synchronous neuronal firing and, therefore, coordinated cognitive activity.

Studies have shown a correlation between EEG activity in certain frequency bands and cognitive engagement with various tasks (Freeman, et al., 1999). This engagement index has been shown to be a valid measure of determining cognitive engagement (Freeman, et al., 1999; Pope, et al., 1995). This is useful while evaluating a new training technology, as improving student motivation and student engagement are often some of the primary goals in creating new methods of learning. There is no evidence for an EEG-based measure of immersion itself. However, it is important to note that all EEG measures are broad patterns of neural firing at a fine-grained time scale—it is the interpretation of those patterns that purports to measure engagement. It is reasonable to hypothesize that, with regard to interacting with a virtual environment, immersion and engagement are strongly related. Therefore, EEG measures of engagement can speak to how immersed someone feels inside the virtual environment.

EEG recording, while accurate, can be expensive and complicated to set up. Consumer-level EEG systems, such as those included with brain-computer interface (BCI) "games", are relatively cheap and include only four or eight electrodes, which are sensors placed on the scalp of the head that can detect electrical activity. Research-quality EEG systems require at least 64 high-quality electrodes for accurate recording, with 128 or even 256 electrode systems as the preferred configuration (figure 8). In addition to increasing the cost of these systems, the complexity of affixing this many electrodes to the head can be quite time-consuming—at least 1 hour per participant for a 64-channel system. Additionally, a saline gel is typically used to improve conductance from the scalp to the electrode. While typically harmless, the gel can be messy, and participants often need to shower after an EEG session to remove the gel. As technology improves, both the cost and complexity should be reduced, especially as interest in BCIs increase.

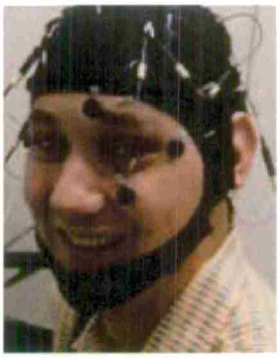

Figure 8. Participant Wearing a 32-Channel EEG Cap

The future of physiological data collection, including EEG recording, lies along the same dimensions of many technologies—more, cheaper, and better. Research-level EEG systems can have hundreds of high-quality electrodes, while commercial systems have dozens at most. Technological progress will lead to the development of cheaper, higher-quality electrodes that omit a less noisy signal. New electrodes for EEG systems will hopefully improve the scalp-electrode connection and eliminate the mildly irritating saline gel that is commonly used in high-end EEG systems. Additionally, wireless technology could be used to remove the bulky wires that currently tie the EEG helmet and all the electrodes to other equipment that cleans up and amplifies signals. It is unlikely, however, to have remote EEG recordings, where the electrodes sit far away from the scalp. The electrical signals given off by the brain are too weak, especially once they dissipate off the scalp. Future generations of commercial EEG recording equipment can use these improved sensors as a foundation for recording brain waves.

4.3.2 Galvanic Skin Response (GSR) and Heart Rate Variability (HRV)

Other physiological measures can be gathered with a lower cost and less complexity at the cost of less validity and accuracy in the data. For example, heart rate variability (HRV) and the galvanic skin response (GSR) have both been shown to indicate stress. Both HRV and GSR are measured with sensors applied to the skin and are relatively inexpensive. The GSR is a measure of skin conductance, which varies based on moisture (i.e., skin sweat). Both of these physiological measures are indicative of stress. While being engaged in a task is not necessarily stressful, heart rate typically increases when someone is engaged in a task versus disengaged (Chaouachi, 2010, Meehan, et al., 2002). This measure could also be especially useful for game-based training environments, where stress is a positive outcome of successful training. However, there are other factors besides engagement that can cause stress—a difficult user interface, for example, could raise stress, introducing confounding variables into the data.

Galvanic skin response and heart rate variability are likely not to have a huge change in terms of hardware technology advancement. For example, heart rate monitors are relatively unobtrusive and can be worn on a daily basis without issue. GSR monitors are not currently designed for long-term wear, but the sensors are comfortable to wear for several hours. The major advancement for these technologies must come from the software side, where the processing algorithms must advance to understand the data more clearly. Any number of factors could influence GSR and heart rate in addition to someone's level of engagement with a virtual environment. Frustration with the user interface, an uncomfortable chair, or a need to use the bathroom can all have an effect. An improved understanding of all these confounding effects plus the technology to filter them out of the signal will increase the usefulness of these measures.

4.3.3 Eye Tracking

One final measure that could be useful to measuring human responses to game-based and virtual world training is an eyetracker, which tracks the pupil of the eye as users make eye movements to gather information in the environment. Eyetrackers can be worn like a helmet for high accuracy or attached to a computer monitor for remote eye tracking that is less accurate but significantly more convenient. Either way, fixation data are commonly gathered by tracking the eyes. A fixation is when the eye stops moving and fixates on a specific location in space, presumably to gather information. Jennett, et al. (2008) found that the number of fixations participants make during a long task will decrease as they focus on specific areas for longer. This measure does not have high-temporal fidelity, especially when compared to EEG recordings. However, eye tracking is often employed for other purposes, so using the technology for another purpose means that the participant does not have to be wired up with as many devices.

Eye tracking technology can also be used to measure engagement. Eye tracking technology has been in development for longer than EEG recordings, but the future trajectory is similar to EEG recording systems. Eyetrackers will become smaller and cheaper as the camera technology improves, and higher-end eyetrackers will become more accurate, sampling eye position at increasingly faster rates. Remote eye tracking systems (with cameras that do not sit on headgear but instead near a computer monitor) have the most to benefit from these developments, as it is more difficult to track the eye when the head is free to move around and the pupil appears much smaller. Eyetrackers as a whole are far less invasive and easier to set up than EEG systems, and remote eyetrackers can be used with minimal effort. As these systems develop, remote eye tracking may serve as an important technology for measuring engagement and immersion.

4.3.4 Wearable and Embedded Sensors

As technology progresses, improved methods of physiological data collection will be investigated. Currently many laboratory sensors are being incorporated into wearable systems. The Future Force Warrior program from the U.S. Army Soldier Systems Center currently has an initiative underway that includes sensors in combat uniforms (figure 9). These sensors include physiological monitoring systems that observe heart rate, blood pressure, and hydration. Portions of this armor were deployed in 2010 for testing in the field. Additionally, Massachusetts Institute of Technology (MIT) Media Laboratories is working on wearable human sensing devices. This includes motion sensors that measure rotational force and torque on limbs. They also are investigating devices that record social interactions, as well as audio and video data. These wearable sensors could someday provide the ability to increase a person's immersion into a serious game or virtual world while enable unobtrusive physiological data collection.

Figure 9. Future Force Warrior Example Combat Armor

4.4 FUTURE USER INTERFACES FOR VIRTUAL ENVIRONMENTS

Many of the alternative input technologies into virtual environments are in their beginning states and require significant technological development before they are viable for widespread adoption. As mentioned previously, the Microsoft Kinect is the first commercial product that allows for full motion control of a computer system, providing an inexpensive solution to a problem that researchers and engineers have been trying to solve for over a decade. However, it is possible to envision how technology such as this might be used in the future.

The Kinect has a robust motion-tracking system due to the setup of multiple cameras. The combination of multiple Kinect systems could provide full motion-tracking capabilities without impeding the user's movements. Today's motion-tracking systems often require the user to wear multitudes of tracking LEDs or dots that cameras pick up and then interpret into movement. The ability of the Kinect to create a 3-D image of a person could remove the need for additional equipment when tracking people around a room.

Other uses for the Kinect could include 3-D facial recognition. For example, if an operator walks up to a console or system, and they would like to use the 3-D cameras within, the Kinect could be used to digitally recognize the person and load specific settings or preloaded content that is tailored to that user. Additionally, avatars in the virtual world currently do not effectively display emotions, and reading facial cues is one of the more common methods of nonverbal communications. If the Kinect were able to accurately detect emotions and map them onto an avatar's face, all users would likely be more engaged in the virtual environment due to the significant increase in avatar realism.

Similar to the Kinect is a product called OrganicMotion, which utilizes a grid of cameras set to cover a room. These cameras capture the 3-D position of the user along with live video of the person. This information is then fed into the virtual environment. At this time, if the user moves in real space, the avatar will move accordingly; this kind of motion control is similar to what is available with the Microsoft Kinect. The novel technology that Organic Motion provides is that instead of a user's avatar looking like a computer-generated person, the live video is overlaid so it appears that the real-world user inhabits the virtual world (figure 10). This has the potential to increase greatly the level of immersion that the user feels in the environment. With this technology, not only do the user's motions correspond to what the avatar does, but the on-screen avatar looks like the user.

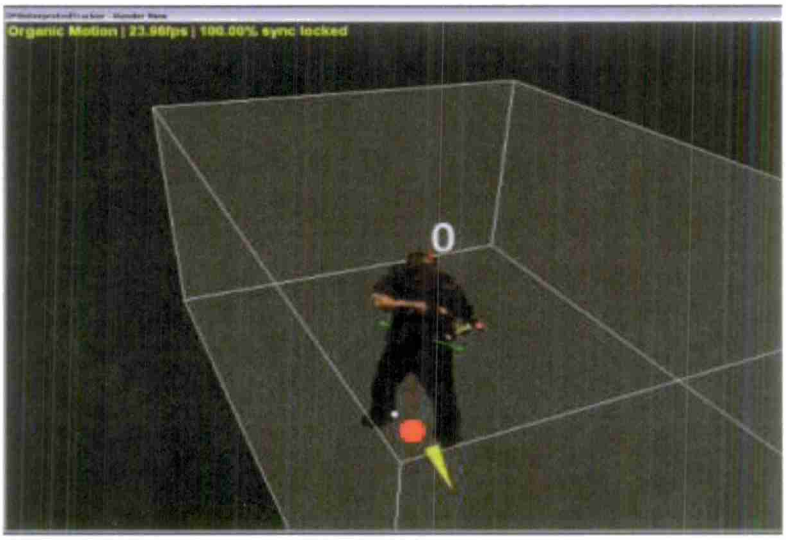

Figure 10. Screenshot of the Organic Motion Interface with Live Video of a User in the Virtual World

Another potential NUI technology comes in the form of the BCI. This technology is still in the embryonic stage, but the promise is to be able to read the EEG signals (i.e., "brain waves") that are emitted as a byproduct of thought using a low-cost, consumer-grade EEG system. From this, a virtual environment could be controlled simply by consciously thinking about what action to take next. These systems have been successfully deployed for only the simplest tasks, such as moving a cursor on a screen in a single direction. Detecting the electrical signals that are given off as a result of neurons firing is very difficult, and being able to turn those signals into useful data requires a great deal of work.

Currently on the market are demonstrations that this kind of "mind-reading" is possible. To be truly useful, there are several areas where BCIs must grow before becoming a viable technology. The first is greater recognition accuracy of a larger number of commands. One part of achieving this goal lies in the hardware. Electrical activity in the brain is very difficult to detect, so the number of electrodes, quality of scalp-electrode connection, and the clarity of the output signal are important factors in being able to classify brain wave patterns. As the EEG hardware provides a cleaner, amplified signal, the software algorithms that perform pattern recognition also must improve. Neural networks are the primary software drivers of pattern recognition, and the algorithms that power these networks will require development. For example, improved training methods for neural networks will greatly improve initial brain wave pattern recognition and reduce initial setup time.

Brainball (Hjelm, 2003) is an example of how this EEG reading technology can be applied. In this simple game, two users face off on opposite sides of the table (figure 11). The objective of the game is to remain more relaxed than the other user. As one's stress levels rise (which is the opposite of the game's objective), a metal ball rolls towards that user's goal area. Eventually, the ball will roll into the stressed user's goal area, and the more relaxed player will score a point. This is a very basic game, but it is an appropriate depiction of the current state of

BCI technology. Only mood states or very simple commands can be detected. For a BCI system that provides input to interact with a virtual environment, this technology must progress a great deal in order for it to be viable.

Figure 11. Sample Game of Brainball

Between improved hardware and software, BCI systems will be able to be trained on and recognize a greater number of inputs. Only then will BCI be ready to interact with virtual environments. Even so, until BCI systems can truly read minds, a keyboard interface may still be required for functionality, such as chat. While voice chat provides a hands-free method of communication, voice is not always ideal (e.g., in group settings). A keyboard provides the most efficient method for text entry, and a BCI system that can understand natural language will take a very long time to develop.

4.5 ENHANCED ALGORITHMS TO IMPROVE DATA ANALYSIS

Many of the future trends for instrumentation and data collection point toward collecting more data—avatar movements, communications, physiological data, etc. However, collecting more data is not necessarily useful unless it can be turned into relevant information. Better algorithms for data mining will be necessary in the future to automate raw data processing. These algorithms will likely take the form of a trainable network (e.g., a Bayesian network) that is able to recognize patterns of input and probabilistically generate meaningful output.

This kind of data mining is being explored by industries that have many disparate data—Wal-Mart, for example, could analyze shopper's purchasing habits to better organize their stores. The government intelligence community has also heavily invested in these kinds of analysis tools to augment their ability to analyze huge amounts of raw data. One potential issue lies in how generalizable these algorithms can be towards other domains and organizations. The intelligence community would likely be investing in text-based data mining, while the data collected in serious games and virtual worlds would be of a different nature. As new algorithms, networks, and analytical techniques are developed, questions of reuse will become important to understand.

An example that applies to serious games and virtual world training could include mouse click data as reported by the system. A number of clicks that took place inside of an area would indicate that the student expected to find an object but was unable to. This could prompt the system to provide the student with a hint as to where to find the desired object. An example involving human data could involve EEG recordings. In real time (or after a learning session), brain wave patterns could indicate that a student has not engaged with the material. This could indicate boredom or else frustration with the content. Both of these examples require collecting a large amount of data to be analyzed for larger, relevant patterns.

5. CONCLUSIONS

The future is promising for serious game and virtual world training technologies because of the advantages that these platforms offer. Students will become more engaged and immersed in the learning environment, so they can absorb more information and retain it for longer. However, many of the claims made by proponents of these technologies must be scientifically evaluated. Groundwork is currently being laid to collect data from the system and the student that will help in assessing training effectiveness, but there is a great deal that must be done. Oftentimes, developing the system itself receives the bulk of time and money, but effectiveness evaluations must be considered as well.

5.1 TECHNICAL CHALLENGES

One of the main impediments to improved data analysis is that algorithms that transform raw data into usable results are difficult to develop. These kinds of data mining techniques are receiving attention because of the number of organizations who have large datasets that they wish to turn into usable information. This kind of data can range from semantic data on the Internet that needs to be searched through to making connections between the items a typical grocery shopper purchases. Each data mining algorithm could possibly be context-specific (i.e., an Internet search algorithm would not be useful when fed student performance), so development efforts may not be able to leverage what other industries have already done. Data mining algorithms are a new field, however, and it is developing quickly.

Another technical challenge is that the larger questions of training effectiveness—modification of on-the-job behavior and better organizational outcomes—require better and longer-term electronic tracking. Collecting data about warfighters in the schoolhouse and beyond to their tours of duty will be necessary to begin to draw connections between training and warfighting performance. This requires open yet secure databases that store all kinds of data about individuals. This technological challenge also presents a policy challenge to ensure that any performance data are not unfairly used against any warfighter.

One final challenge is that more pervasive physiological data collection will require equipment that is easier to set up and less invasive. Currently, physiological data collection requires mildly invasive procedures on the participant, which limits the usefulness and external validity of the data. The idea of embedded sensors in uniforms is a positive trend towards wireless and remote recording that is transparent to the user. Data are automatically collected for analysis without the warfighter even needing to think about it. This field of remote sensing is growing; however, the medical applications and potential benefits are tremendous.

5.2 FUTURE RESEARCH DIRECTIONS

Additional research must be done on training effectiveness. Both initial learning and later knowledge retention must be tested. Examining retention is often difficult, especially with warfighters leaving the schoolhouse and moving into the field. However, a great deal of knowledge learned during training is lost afterward, so retention is a crucial aspect of effectiveness. More complex measures of effectiveness, such as on-the-job-performance and organizational outcomes, must begin to be examined. These kinds of evaluations are not easy or inexpensive to carry out, but they will actually be able to answer questions related to return on investment and training efficacy.

Better data collection and analysis techniques will contribute toward these goals. Innovative means of examining student learning and improved technology to measure learning will reduce the resources needed to perform effectiveness evaluations. As data collection and analysis occurs in a transparent fashion, automated data collection and analysis of system and human data will provide rapid feedback to the student, training system, and analysts. This will lessen turnaround time for students who want to know how well they understand the material, for instructional designers who want to know how well their training system can teach, and analysts who want to understand how users performed using new technology, performing new missions, or developing new CONOPS.

6. BIBLIOGRAPHY

Berry, M., "Army Game Studio: Serious Games Development," *Defense GameTech Users Conference*, Orlando, FL, 29–31 March 2010.

Beus, T. M., and C. T. Carde, "Game-Based Learning Prototype - Lessons Learned Paper," *Defense GameTech Users Conference 2011*, Orlando, FL, 22–25 March 2011.

"Bohemia Interactive Simulations," VBS2 VTK, http://www.bisimulations.com/products/vbs2

Bransford, J. D., A. L. Brown, and R. R. Cocking, eds., *How People Learn*, National Academy Press, Washington, DC, 1999.

Chaouachi, M., "Affect and Mental Engagement: Towards Adaptability for Intelligent Systems," *The Twenty-Third International Florida Artificial Intelligence Research Society Conference*, Daytona Beach, FL, 19–21 May 2010.

Coelho, C., J. Tichon, T. J. Hine, G. Wallis, and G. Riva, "The Sense of Presence in Virtual Reality Technologies," in *Communication to Presence*, G. Riva and F. David, eds., IOS Press, Amesterdam, Netherlands, 2006.

"Common Remotely Operated Weapon Station," America's_Army_Platform_and_Information_Site, http://info.americasarmy.com/projects.php?id=10.

Cubic Corporation, *Javelin Basic Skills Trainer*, San Diego, CA, 2009.

Darken, R. P., D. Bernatovich, J. Lawson, and B. Peterson, "Quantitative Measures of Presence in Virtual Environments: The Roles of Attention and Spatial Comprehension," *Cyberpsychology and Behavior*, vol. 2, no. 4, 1999, pp. 337–347.

Dempsey, M. G., "Moral Combat Facilitator Guide," Center for Army Profession and Ethic. West Point, NY, 2011.

Dixit, P., and M. Youngblood, "Understanding Playtest Data Through Visual Data Mining in Interactive 3-D Environments," *IEEE Computer Graphics and Applications*, March/April 2009, pp. 30-38.

Eich, E., and J. Metcalfe, "Mood Dependent Memory for Internal versus External Events," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 15, no.3, 1989, pp. 443–455.

Freeman, F., P. Mikulka, L. Prinzel, and M. Scerbo, "Evaluation of an Adaptive Automation System Using Three EEG Indices with a Visual Tracking Task," *Biological Psychology*, vol. 20, 1999, pp. 61–76.

Friess, J., "Selling the Virtual World for Navy Training," Southern Maryland Newspapers Online, http://ww2.somdnews.com/stories/03182011/entemil161415_32321.shtml .

Gallagher, S., and P. Berking, "Next Generation Learner: Design for Achieving Higher Order Learning Outcomes," *Advanced Distributed Learning Implementation Fest 2011*, Orlando, FL, 2–4 August 2011.

Godden, D. R., and A. D. Baddeley, "Context Dependent Memory in Two Natural Environments: On Land and Underwater," *British Journal of Psychology*, vol. 66, no. 3, August 1975, pp. 325–331.

Goldberg, S. A., "Game Technology in Institutional and Unit Training," *Defense GameTech Users Conference 2010*, Orlando, FL, 29–31 March 2010.

Grant, H. M., L. C. Bredahl, J. Clay, J. Ferrie, J. E. Groves, T. A. McDorman, et al., "Context-Dependent Memory for Meaningful Material: Information for Students," *Applied Cognitive Psychology*, vol. 12, no. 6, 1998, pp. 617–623.

Heeter, C., "Being There: The Subjective Experience of Presence," *Presence: Teleoperators and Virtual Environments*, vol. 1, no. 2, 1991, pp. 262-271.

Hjelm, S. I., "Research + Design: The Making of Brainball," *Association for Computing Machinery (ACM) Interactions*, vol. 10, no. 1, 2003, pp. 26–34.

Jennett, C., A. L. Cox, P. Cairns, S. Dhoparee, A. Epps, T. Tijs, et al., "Measuring and Defining the Experience of Immersion in Games," *International Journal of Human-Computer Studies*, vol. 66, no. 9, September 2008, pp. 641–661.

Kirkpatrick, D. L., *Evaluating Training Programs: The Four Levels*, Berrett-Koehler Publishers, San Francisco, CA, June 1998.

Knerr, B., and D. Lampton, "An Assessment of the Virtual Integrated MOUT Training System (V-IMTS)," United States Army Research Institute for the Behavioral and Social Sciences, Ft. Belvoir, VA, 2005.

La, C., and P. Michiardi, "Characterizing User Mobility in Second Life, *Proceedings of the First Workshop on Online Social Networks*, 2008, pp. 79–84.

"Live Fire and Virtual Targetry," America's_Army_Platform_and_Information_Site, http://info.armydev.com/projects.php?id=30.

Lombard, M., T. B. Ditton, D. Crane, B. Davis, G. Gil-Egui, K. Horvath, et al., "Measuring Presence: A Literature-Based Approach to the Development of a Standardized Paper-and-Pencil Instrument," *The Third International Workshop on Presence*, Delft, The Netherlands, 27–28 March 2000.

Lombard, M., R. Meddis, D. Pearson, and W. Ijsselsteijn, "Using Behavioural Realism to Estimate Presence: A Study of the Utility of Postural Responses to Motion-Stimuli," *Presence: Teleoperators and Virtual Environments*, vol. 9, no. 2, 2000, pp. 149–164.

Luber, L., M. Swartwood, J. Swartwood, and P. O'Donnell, "Evaluation of the Effectiveness of EEG Neurofeedback Training for ADHD in a Clinical Setting as Measured by Changes in T.O.V.A. Scores, Behavioral Ratings, and WISC-R Performance," *Applied Psychophysiology and Biofeedback*, vol. 20, no. 1, 1995, pp. 83–99.

McAlinden, R., "Games for Research: Lesson's Learned and the Road Ahead," *Defense GameTech Users Conference 2009*, Orlando, FL, 16 March 2009.

McFadden, J., "Danger Close," *Fires*, Department of the Army, Fort Sill, OK, September/October 2010.

McKay, T., "Defense Finance and Accounting Service Strategic Plan," Defense Finance and Accounting Service, Indianapolis, IN, 2011.

Meehan, M., B. Insko, M. Whitton, and F. Brooks, "Physiological Measures of Presence in Stressful Virtual Environments," *29th Annual Conference on Computer Graphics and Interactive Techniques*, San Antonio, TX, 23–26 July 2002.

Monte, P. V., "Applications of Virtual World Technology to Collaborative Engineering and Concept of Operation Exercise (COOPEX) Planning and Rehearsal," presented at the *Maritime Systems and Technology Conference*, Washington, DC, 2010.

Murphy, C., "Why Games Work – The Science of Learning," *2011 Interservice/Industry Training, Simulation, and Education Conference*, Orlando, FL, 28 November 2011.

Murphy, C., and T. Hussain, "Damage Control Trainer – A Serious Game Case Study," *2009 United States Coast Guard Human Performance Technology Workshop*, Williamsburg, VA, 23–25 September 2009.

Nguyen, A., "Undersea Warfare Training in a Virtual Environment," *Maritime Systems and Technology*, Washington, DC, 22–24 June 2010.

Ogren, L., and J. H. Wong, "An Analysis of Game Controller and Touchscreen Devices for Input Into a Complex High-Information Display," *54th Annual Meeting of the Human Factors and Ergonomics Society*, San Francisco, CA, 27 September–1 October 2010.

"Operational Language and Culture Training Systems," Alelo, Inc., http://www.alelo.com/tactical_language.html.

Overte Foundation, "OpenSimulator," http://opensimulator.org/wiki/Main_Page

Pope, A. T., E. H. Bogart, D. S. Bartolome, and T. D. Wilson, "Biocybernetic System Evaluation Indices of Operator Engagement in Automated Task," *Biological Psychology*, vol. 40, no. 9, May 1995, pp. 187–195.

Prensky, M., "True Believers: Digital Game-Based Learning in the the Military," *Digital Game-Based Learning*, McGraw-Hill, New York, 2001.

Raybourn, E., "Preliminary Results from a Study on a Real Game for Training: The Case of DARWARS Ambush!," *Defense GameTech Users Conference*, Orlando, FL, 9–12 March 2009.

Roman, P., "Serious Games - An Experiental Learning Renaissance," *Defense GameTech Users Conference 2011*, Orlando, FL, 22 March 2011.

Sadagic, A., "Deconstructing Game-Based Systems: What are They? What Really Matters?" *2011 Interservice/Industry Training, Simulation, and Education Conference*, Orlando, FL, 28 November–1 December 2011.

Schuemie, M. J., P. van der Straaten, M. Krijn, and Charles A.P.G. van der Mast, "Research on Presence in Virtual Reality: A Survey," *Cyberpsychology and Behavior*, vol. 4, no. 2, pp. 183–202, April 2001.

Slater, M., V. Linakis, M. Usoh, and R. Kooper, "Immersion, Presence, and Performance in Virtual Environments: An Experiment with Tri-Dimensional Chess," *ACM Virtual Reality Software and Technology*, London, UK, 1996.

Sreedharan, S., E. S. Zurita, and B. Plimmer, "3-D Input for 3-D Worlds," *Proceedings of the 2007 Conference of the Computer-Human Interaction Special Interest Group (CHISIG) of Australia on Computer-Human Interaction: Design, Activities, Artifacts, and Environments*, ACM, New York, 2007, pp. 227–230.

Stephens, G. L., "PEO STRI Gaming Program Brief," *Defense GameTech Users Conference 2009*, Orlando, FL, 9–12 April 2009.

Strickland, D., and D. Chartier, "EEG Measurements in a Virtual Reality Headset," *Presence*, vol. 6, no. 5, 1997, pp. 581–589.

Surface, E. A., E. C. Dierdorff, and A. M. Watson, "Special Operations Language Training Software Measurement of Effectiveness Study: Tactical Iraqi Study Final Report," Surface, Ward, and Associates Technical Report No. 2007010602, Raleigh, NC, 2007.

Swan, K., "Learning Effectiveness: What the Research Tells Us," in *Elements of Quality Online Education, Practice and Direction*, J. Bourne and J. C. Moore, eds., Center for Online Education, Needham, MA, 2003.

Topolski, R., B. Leibrecht, S. Cooley, and N. Rossi, "Impact of Game-Based Training on Classroom Learning Outcomes," United States Army Research Institute for Behavioral and Social Sciences, Arlington, VA, September 2010.

Witmer, B., and M. Singer, "Measuring Presence in Virtual Environments: A Presence Questionnaire," *Presence*, vol. 7, no. 3, 1998, pp. 225–240.

APPENDIX
SURVEY AND QUESTIONNAIRE TEMPLATES

PRESENCE QUESTIONNAIRE

QUESTION	RESPONSE
1. How much were you able to control events?	1 2 3 4 5 6 7 COULD NOT CONTROL COULD ALWAYS CONTROL
2. How responsive was the environment to actions that you initiated (or performed)?	1 2 3 4 5 6 7 NOT RESPONSIVE VERY RESPONSIVE
3. How natural did your interactions with the environment seem?	1 2 3 4 5 6 7 NOT NATURAL VERY NATURAL
4. How much did the visual aspects of the environment involve you?	1 2 3 4 5 6 7 NOT INVOLVING VERY INVOLVING
5. How natural was the mechanism which controlled movement through the environment?	1 2 3 4 5 6 7 NOT NATURAL VERY NATURAL
6. How aware were you of events occurring in the real world around you?	1 2 3 4 5 6 7 NOT AWARE VERY AWARE
7. How aware were you of your display and control devices?	1 2 3 4 5 6 7 NOT AWARE VERY AWARE
8. How compelling was your sense of objects moving through space?	1 2 3 4 5 6 7 NOT COMPELLING VERY COMPELLING
9. How inconsistent or disconnected was the information coming from your various senses?	1 2 3 4 5 6 7 NOT CONSISTENT VERY CONSISTENT
10. How much did your experiences in the virtual environment seem consistent with your real-world experiences?	1 2 3 4 5 6 7 NOT CONSISTENT VERY CONSISTENT
11. How often were you able to anticipate what would happen next in response to your actions?	1 2 3 4 5 6 7 COULD NOT PREDICT COULD ALWAYS PREDICT
12. How completely were you able to actively survey or search the environment using vision?	1 2 3 4 5 6 7 COULD NOT SEARCH COULD ALWAYS SEARCH

PRESENCE QUESTIONNAIRE (Cont'd)

QUESTION	RESPONSE
13. How compelling was your sense of moving around inside the virtual environment?	1 2 3 4 5 6 7 NOT COMPELLING VERY COMPELLING
14. How closely were you able to examine objects?	1 2 3 4 5 6 7 NOT CLOSELY VERY CLOSELY
15. How well could you examine objects from multiple viewpoints?	1 2 3 4 5 6 7 NOT WELL VERY WELL
16. How well could you move or manipulate objects in the virtual environment?	1 2 3 4 5 6 7 NOT WELL VERY WELL
17. To what degree did you feel disoriented at the beginning of breaks or at the end of the session?	1 2 3 4 5 6 7 VERY CONFU.S.ED NOT CONFU.S.ED
18. How involved were you in the virtual environment experience?	1 2 3 4 5 6 7 NOT INVOLVED VERY INVOLVED
19. How distracting was the control mechanism?	1 2 3 4 5 6 7 NOT DISTRACTING VERY DISTRACTING
20. How much delay did you experience between your actions and expected outcomes?	1 2 3 4 5 6 7 VERY DELAYED NOT DELAYED
21. How quickly did you adjust to the virtual environment experience?	1 2 3 4 5 6 7 NEVER ADJU.S.TED VERY QUICKLY ADJU.S.TED
22. How proficient in moving and interacting with the virtual environment did you feel at the end of the experience?	1 2 3 4 5 6 7 NOT PROFICIENT VERY PROFICIENT
23. How much did the visual display quality interfere or distract you from performing assigned tasks or required activities?	1 2 3 4 5 6 7 ALWAYS INTERFERED NEVER INTERFERED

PRESENCE QUESTIONNAIRE (Cont'd)

QUESTION	RESPONSE
24. How much did the control devices interfere with the performance of assigned tasks or with other activities?	1 2 3 4 5 6 7 ALWAYS INTERFERED NEVER INTERFERED
25. How well could you concentrate on the required activities rather than on the mechanisms used to perform those tasks or activities?	1 2 3 4 5 6 7 NOT WELL VERY WELL
26. Did you learn new techniques that enabled you to improve your performance?	1 2 3 4 5 6 7 LEARNED NOTHING LEARNED A LOT
27. Were you involved in the task to the extent that you lost track of time?	1 2 3 4 5 6 7 NEVER LOST TRACK ALWAYS LOST TRACK
28. How completely were all of your senses engaged?	1 2 3 4 5 6 7 NOT ENGAGED VERY ENGAGED
29. How much did the auditory aspects of the environment involve you?	1 2 3 4 5 6 7 NOT INVOLVED VERY INVOLVED
30. How well could you identify sounds?	1 2 3 4 5 6 7 NOT WELL VERY WELL
31. How well could you localize sounds?	1 2 3 4 5 6 7 NOT WELL VERY WELL
32. How well could you actively survey or search the virtual environment using touch?	1 2 3 4 5 6 7 NOT WELL VERY WELL

IMMERSIVE TENDENCIES QUESTIONNAIRE

QUESTION	RESPONSE						
1. Do you ever get extremely involved in projects that are assigned to you by your boss or your instructor, to the exclusion of other tasks?	1 NEVER	2	3	4	5	6 ALWAYS	7
2. How easily can you switch your attention from the task in which you are currently involved to a new task?	1 HARD	2	3	4	5	6 EASY	7
3. How frequently do you get emotionally involved (angry, sad, or happy) in the news stories that you read or hear?	1 NEVER	2	3	4	5	6 ALWAYS	7
4. How well do you feel today?	1 NOT WELL	2	3	4	5	6 WELL	7
5. Do you easily become deeply involved in movies or TV dramas?	1 NEVER	2	3	4	5	6 ALWAYS	7
6. Do you ever become so involved in a television program or book that people have problems getting your attention?	1 NEVER	2	3	4	5	6 ALWAYS	7
7. How mentally alert do you feel at the present time?	1 NOT ALERT	2	3	4	5	6 ALERT	7
8. Do you ever become so involved in a movie that you are not aware of things happening around you?	1 NEVER	2	3	4	5	6 ALWAYS	7
9. How frequently do you find yourself closely identifying with the characters in a story line?	1 NEVER	2	3	4	5	6 ALWAYS	7
10. Do you ever become so involved in a video game that it is as if you are inside the game rather than moving a joystick and watching the screen?	1 NEVER	2	3	4	5	6 ALWAYS	7
11. On average, how many books do you read for enjoyment in a month?							
12. What kind of books do you read most frequently? (CIRCLE ONE ITEM ONLY!)	Spy novels Adventure Westerns Biographies	Fantasies Romance novels Mysteries Autobiographies	Science fiction Historical novels Other fiction Other non-fiction				

IMMERSIVE TENDENCIES QUESTIONNAIRE (Cont'd)

QUESTION	RESPONSE
13. How physically fit do you feel today?	1 2 3 4 5 6 7 NOT FIT … FIT
14. How good are you at blocking out external distractions when you are involved in something?	1 2 3 4 5 6 7 NOT GOOD … VERY GOOD
15. When watching sports, do you ever become so involved in the game that you react as if you were one of the players?	1 2 3 4 5 6 7 NEVER … ALWAYS
16. Do you ever become so involved in a daydream that you are not aware of things happening around you?	1 2 3 4 5 6 7 NEVER … ALWAYS
17. Do you ever have dreams that are so real that you feel disoriented when you awake?	1 2 3 4 5 6 7 NEVER … ALWAYS
18. When playing sports, do you become so involved in the game that you lose track of time?	1 2 3 4 5 6 7 NEVER … ALWAYS
19. Are you easily disturbed when working on a task?	1 2 3 4 5 6 7 NEVER … ALWAYS
20. How well do you concentrate on enjoyable activities?	1 2 3 4 5 6 7 NOT WELL … WELL
21. How often do you play arcade or video games? (OFTEN should be taken to mean every day or every two days, on average.)	1 2 3 4 5 6 7 NEVER … ALWAYS
22. How well do you concentrate on disagreeable tasks?	1 2 3 4 5 6 7 NOT WELL … WELL
23. Have you ever gotten excited during a chase or fight scene on TV or in the movies?	1 2 3 4 5 6 7 NEVER … ALWAYS
24. To what extent have you dwelled on personal problems in the last 48 hours?	1 2 3 4 5 6 7 NONE … LOTS

IMMERSIVE TENDENCIES QUESTIONNAIRE (Cont'd)

QUESTION	RESPONSE						
25. Have you ever gotten scared by something happening on a TV show or in a movie?	1 NEVER	2	3	4	5	6	7 ALWAYS
26. Have you ever remained apprehensive or fearful long after watching a scary movie?	1 NEVER	2	3	4	5	6	7 ALWAYS
27. How frequently do you watch TV soap operas or docu-dramas?	1 NEVER	2	3	4	5	6	7 ALWAYS
28. Do you ever become so involved in doing something that you lose all track of time?	1 NEVER	2	3	4	5	6	7 ALWAYS

INITIAL DISTRIBUTION LIST

Addressee	No. of Copies
Defense Technical Information Center	2