Winter 2003

An Examination of Personality as a Predictor of Guard Behavior in a Virtual Environment

Jean M. Catanzaro
Old Dominion University

AN EXAMINATION OF PERSONALITY AS A PREDICTOR OF GUARD BEHAVIOR IN A VIRTUAL ENVIRONMENT

by

Jean M. Catanzaro
B.A. December, 1995, University of South Florida
M.A. December, 1998, East Carolina University

A Dissertation Submitted to the Faculty of
Old Dominion University in Partial Fulfillment of the
Requirement for the Degree of

DOCTOR OF PHILOSOPHY

INDUSTRIAL/ORGANIZATIONAL PSYCHOLOGY

OLD DOMINION UNIVERSITY
December 2003

Approved by:

Mark W. Scerbo (Director)

James P. Bliss (Member)

Glynn D. Coates (Member)

Frederic D. McKenzie (Member)

ABSTRACT

AN EXAMINATION OF PERSONALITY AS A PREDICTOR OF GUARD BEHAVIOR IN A VIRTUAL ENVIRONMENT

Jean M. Catanzaro
Old Dominion University, 2003
Director: Dr. Mark W. Scerbo

Military personnel need access to realistic training tools that can provide a safe environment in which to acquire skills that will generalize to real world tasks. A virtual environment (VE) is one such tool. The focus of the present study was to evaluate a VE as a training tool for military guards. The first goal was to examine the potential of VE technology to provide effective training for standing watch at a military checkpoint. The second goal was to study a set of personality traits that might predict performance. Participants completed the NEO Personality Inventory and were trained to perform the role of a military checkpoint guard within a CAVE Automatic Virtual Environment. Trainees interacted with virtual drivers and determined whether drivers exhibited suspicious behavior and met identification requirements for entry onto a fictional base. Results indicated that participants were able to use VE technology to (a) familiarize and immerse themselves in a military checkpoint task, (b) improve performance on training scenarios, and (c) transfer their knowledge from one session to a subsequent session. Examination of personality traits yielded significant results only for openness as a predictor of performance. Collectively, these findings suggest that VEs show potential for scenario-based training.

In grateful appreciation of my friends and family who have been there to support and encourage me.

ACKNOWLEDGMENTS

The longer I work towards my goals and the more I learn about the fortunate experiences in my life, the harder it is to summarize how grateful I am to the people who matter most. First, my thanks to God for giving me the strength, appreciation and opportunities that led up to my graduate career. Second, I would like to acknowledge the unfaltering support of my mother, father and brother who have always had encouraging words, patience and more encouraging words, not to mention a collectively outstanding sense of humor that put hard times in perspective. (It is a blessing to have parents who believe my biggest weakness is chocolate and a brother whose positive attitude is very contagious.) Third, my friends who have hung in there with me and were patient when I was busy with my studies. Finally, but importantly, my committee members, Dr. Scerbo, Dr. Coates, Dr. Bliss and Dr. McKenzie as well as Dr. Silver who provided great insight into the realm of personality and military behavior. I am grateful for their input and encouragement in my professional and personal growth. I am especially thankful for my advisor Dr. Mark Scerbo who provided a great deal of support throughout the duration of my graduate career and Dr. Glynn Coates whose wit and support were also valuable to me. I realize professors are empowered with the potential to greatly influence the paths of their students and I am grateful to have had professors who truly had a passion for their field, a passion that undoubtedly rubs off on others.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Providing the appropriate resources and training for military guards allows the United States to be a world leader and promoter of world peace. The United States trains more than 100,000 foreign police and soldiers annually, both within the U.S. and in over 100 countries around the world (Garcia, 2002). This leadership role is maintained through an ongoing exploration and development of advanced training tools.

On September 11, 2001, the United States encountered the most extensive, damaging act of terrorism in history. While the majority of U.S. military activity occurs outside of our borders, the unexpected attack on the World Trade Center Towers, resulting in over 2800 fatalities, presents us with a dramatic example of the increased need to train and prepare for threats on our own soil.

There is already an extensive array of training platforms for our military, involving self defense combat tactics, sniper training, and wide scale simulated group war games. However, much of military training is focused on war fighting in open terrain, using weaponry in rural areas and still lacks scenarios that address the characteristics and constraints of operating in urban environments (Hills, 2002). Resolving this vulnerability is an important objective at a time where, as Hills (2002) states: "Problems develop more rapidly, operations are more resource intensive" and the difficulty discriminating between friend and foe is "as great as ever" (p. 5). Most military action occurs abroad and our troops are less acquainted with homeland defense. In essence, much is still missing from the range of training necessary to prepare for new terrorist challenges (Hills, 2002).

The journal model format used for the preparation of this dissertation is the *Journal of Applied Psychology*.

Technology can play a role in providing a safe means for filling some of the gaps. Technology supports a more extensive and versatile training regimen for military personnel by providing a safe environment to apply new knowledge and means for transferring important concepts to novel situations.

The focus of the present study is to evaluate a virtual environment (VE) as a training tool for military guards. The project had two primary goals. The first was to examine the potential of VE technology to provide effective training for people standing watch at a military checkpoint. Second, personality factors that potentially contribute to performance will also be examined. Specifically, a set of personality traits considered as potential predictors of performance on a military checkpoint task will be studied.

This section is organized into three main topics. First, the nature of a military checkpoint and guard tasks will be discussed. Second, an overview of VE technology, its applications and applicability to military tasks will be provided. Last, personality predictors will be discussed as potential predictors of performance will be discussed.

Guarding the Checkpoint

Military checkpoints, established at strategic entry points to a base, town, or property are staffed by military guards who play a pivotal role in the protection of military assets and public safety. These checkpoints allow personnel to screen individuals and vehicles for appropriate entry requirements. Consequently, they serve an important function as a potential first line of defense and provide an early opportunity to identify and disengage threats. The ability to properly select the best qualified personnel and provide proper resources and training for these guards is critical to safeguarding lives and property.

Standing watch at a military checkpoint is a critical and often perilous job because, as a first line of defense, the checkpoint can also be the first opportunity for intervention against threats. Performing guard duty in Iraq has proved fatal for both civilian and soldier alike. Among other tactics, terrorists are capitalizing on soldiers' inclination to cater to cars who deceptively indicate the presence of an injured passenger (Cable News Network, 2003). A group of women and children were killed in a suicide car bombing at a checkpoint near Najaf when troops encountered a van that failed to halt despite commands and warning shots ("Battles Get Close-Up," 2003). People who wish to infiltrate or do harm to property and humankind have demonstrated tactics that have become more inventive in technique. Consequently, those assigned to these posts conduct tasks that require many skills. These include observation, psychological and social skills (e.g., the ability to read body language and respond appropriately to emotional drivers), the need to remain alert over extended periods of time as well the ability to handle matters of situational urgency (e.g., injured passengers, traffic waiting), the necessity to deal with diverse populations, the ability to identify procedures that apply in a given situation, and the ability to maintain a running log in memory of people, vehicles, and characteristics deemed dangerous or problematic. The *Navy Bluejacket's* manual states that performing a security watch is one of the first duties of a military recruit and requires that the sentry "keep alert, attend to duty, report all violations, and preserve order" (Cutler & Herdt, 1999, p. 90). Not only do guards have to rely on their many skills, but they often have to do so for long periods of time. If suspicious activity is detected, the guard must handle the situation by executing proper procedures often under varying levels of stress and workload. In addition, guards must perform all these activities without escalating problems. Although most of the time guards face innocuous situations, the ability to be prepared for critical events at any given point is a necessity of the job. Also, many times there are nondangerous events that still require some assertiveness and social know-how, (e.g., handling angry drivers who are missing a base decal or visitor pass and are told to turn around). Other times, a guard may detect an item (e.g., an unmarked package, knife, or hunting rifle) lying on the car seat that raises cause for concern. When on duty, a military guard must know that the possibility of threat can arise at any given moment and when least expected. It is his or her job to be trained and prepared.

Nature of the Job Environment

Military guards work in a very challenging environment marked by tedious, innocuous periods and highlighted with periods of extreme arousal and alarm resulting from potential dangers. As shown by action and gathered intelligence data, enemies are formulating new strategies, generating novel methods of threat that potentially threaten the security of the installation that military guards serve to protect.

Threats to a base may or may not be obvious. A threat to a base can come in any number of forms, from a person using a false identity or documents to gain access, to a person carrying illegal or dangerous items. Guards must inspect cars for odd markings or suspicious contents. As "America's open borders make tracking terrorists a daunting exercise" (Thomas, 2001, p. 2), there is perhaps no better way to describe the primary task of a military guard than to be ready for the unexpected.

A Day in the Life of a Gate Guard

At a typical checkpoint, traffic volume varies. For example, base entry points may have infrequent traffic every 10 minutes or heavy periods of traffic that result in lines of vehicles. This queue can inappropriately result in a guard rushing through procedures to appease waiting vehicles. Certain entry points allow the screening of trucks and other entry points may be designated for entry by cars only. Guards must watch for persons and vehicles wanted by military or state authorities documented on their alert lists. The guards are in continuous contact with the base dispatch, able to immediately assess license plate and driver license information for persons with outstanding warrants or for which there is an alert issued for suspicious behavior. While on watch, guards are vigilant about identifying disjointed aspects of dialogue from drivers, (i.e., comments that just "don't add up") and maintaining general alertness for the passersby and the activities of fellow guards who also are working at the checkpoint.

The task of screening a vehicle at a military checkpoint usually begins by surveying the approach of the vehicle for odd behaviors and characteristics such as swerving driving patterns, missing decals, or missing license plates. Once the vehicle has approached, the guard evaluates whether the vehicle and its occupants have appropriate identification for entry into the base. In addition to identifying inappropriate responses to simple questions, guards look for odd, possibly threatening behavior, suspicious identification, registration, or other paperwork. It is important that guards be able to work effectively with fellow guards. Throughout their daily tasks, guards must maintain military bearing in all interactions. Military bearing refers to an overriding manner of conduct requiring guards to handle themselves with dignity in a manner that reflects credit on the military service.

Importance of Training

Being a military guard requires quick thinking and immediate action in the event of a dangerous situation. When a dangerous event arises, the guard must have a response already in mind and the execution of that response must be second nature. To avoid danger, the guard must draw on his or her training and be prepared to react quickly.

Quality of training is of utmost concern to security. Consequently, training procedures must be adapted to meet the new situations that guards may encounter. Training for security personnel must be more elaborate than ever. Prior to Sept. 11, 2001, few would have thought that a single man who "has no throne, no armies, not even any real territory" could create such a intense machine of threat by merely possessing the "power to make men willingly go to their deaths for the sole purpose of indiscriminately killing Americans" (Thomas, 2001). Today, our enemies are bold, empowered by a perceived sense of martyrdom as a motivation to audaciously encroach upon our space and willfully die in exchange for delivering threat. Many former strategies have been based on the assumption that, as a general rule, the enemy will work in such a manner as to provide the optimization of attack while, where possible, minimizing the opportunity for injury to oneself or fellow soldiers. Of course, there are situations where a soldier will run across a field, in absence of any other timely, viable alternative toward meeting a goal (e.g., bridging a communication gap amid absent/failed radio technology). However, generally, one would expect an attacker's personal survival instincts to be a significant motivator in the strategy adopted. Without such a motivator, identifying aberrant behavior or characteristics becomes more important in the detection of persons who, once deployed toward a target, will exert maximal aggression and resulting damage.

The military must develop adaptable training tools that are dynamic and can address emerging threats. For example, new procedures and technology must be developed for fighting in urban terrains, detecting and confronting enemies underground or within caves in extreme weather conditions. As Hills (2002) states, issues such as force protection in urban areas may not be that different from those areas in mountains or jungles, however, designing training mechanisms that will address the magnitude and unpredictability of the situation is a challenge. Thus, there is a need for new and innovative technologies to better model the complexity of human behavior in dynamic changing scenarios that mimic real world counterparts.

Existing Training Mechanisms

The current training for checkpoint guards is typically conducted on the job. In addition, guards draw on other forms of training (e.g., simulated training tasks, field exercises, and classroom lectures) and attempt to generalize that training to their checkpoint duties. Prior to specific security training, all military personnel go through "Basic Training," consisting of general physical, psychological, and institutional preparation for becoming a soldier. Basic training provides knowledge relevant to military guard duties (e.g., team work exercises, strategic positioning of one's body, and concise communication skills), however, none of these activities are specifically designed for checkpoint duty activities. Furthermore, there is a historical lack of standardization in the amount and type of training guards receive as training varies widely from instructor to instructor (Heacock, 1999).

Training is contextually dependent and can differ depending on the military base for which the soldier serves. In other words, guards rely on on-the-job training that appears to be a nonstandardized, passing down of information from more experienced guards. Another problem with on-the-job training is that unless an event comes up during one's shift, one may never be trained to handle that event. Further, people assigned to guard duty are often the more inexperienced younger personnel. For those assigned to checkpoint duty, it is often considered an auxiliary task as opposed to one's primary job assignment[1] (E.S. Ankney, personal communication, October 2001).

Difficulties in Training

The operating environment for the military guard may be hostile or hazardous, often stressful, and is regularly time sensitive. As such, the environment provides less than optimal conditions for training. Some components of military training are not always conducive to traditional teaching methods such as textbooks, interactive software, and hands-on (live) training. Rose and Foreman (1999) suggest that virtual reality (VR) offers a more interactive medium than videos and CD-ROMs and has the potential to be more realistic than diagrams and manuals. Also, traditional methods fail to adequately simulate the situational pressure and team work critical for task performance. As Smith (1995) states: the enemy "is unpredictable and does not always operate as the books say he should" (p. 63). Further, it would be difficult and impractical to direct 100 volunteer drivers to simulate traffic conditions for the purpose of allowing guards to practice security and traffic routing procedures. It would be equally difficult to assemble those 100 volunteer drivers to simulate various breaches of security or subtle deviations from

[1] It should be noted that the nature of this job has changed within the time frame this paper is being written due to the terrorist acts of September 11, 2001.

an established norm or profile. The response to a given hostile event is often difficult to train without the availability of the event, e.g., an attacker. To complicate the issue, it is not always possible to reproduce certain events for as many trials as would be needed to achieve training objectives. Further, the physical location, facilities, and instructor may not be available to train someone who is assigned to leave immediately for a mission. This raises the question of how the military can train large groups of personnel to become effective military guards in a way that minimizes the use of resources (e.g., staff and space), while maximizing training effectiveness.

Training for military personnel is limited to activities that can be conducted under relatively safe conditions. As noted earlier, live training is also constrained by practical limitations; Pew and Mavor (1998) note that limits on physical space and the operational costs for major weapon systems confine the number of units that can take part in live simulations. However, simulation and VR offer many advantages for training over live or classroom approaches. Trainees can benefit from interaction with virtual instructors and experience virtual scenarios in a simulated environment designed to meet readiness training objectives. The training necessary to meet readiness objectives is extensive and must be reinforced through practice. The complex nature of the military environment coupled with the extensive amount of rapidly changing technology create the need for comprehensive initial training and nearly constant retraining in order to maintain proficiency (Driskell, Hogan, Salas, & Hoskin, 1994). Skill retention is critical since there may be long periods of time between skill acquisition and actual deployment of military personnel to a mission that will utilize those skills. Maintaining knowledge gleaned from training sessions may be difficult if there is a long time interval between the training exercise and the application of the knowledge, especially if there are similar tasks that generate interference factors. Generally speaking, the longer the length of the retention interval, the greater the skill decay (Arthur, Bennett, Stanush, & McNelly, 1998). Although research has indicated that initial skill retention from VE training may not be superior to conventional electronic media for certain intellectual skills (Hall, Stiles, & Horwitz, 1998), VE has the potential to offer participants a practical refresher course and consequently may provide a mechanism to reinforce knowledge.

The military recognizes the potential of VEs to provide a cost effective and safe method for training and has become a leading user of VEs for training (Pew & Mavor, 1998). The military's use of VE technology is quite extensive. As a sampling of applicability, Lampton et al. (1995) wrote: "The U.S. Army has made a substantial commitment to the use of VE technology, such as networked simulators to create virtual battlefields for combat training and mission rehearsal, development of military doctrine, and evaluation of weapon system concepts prior to acquisition decisions" (p. vii). Caird (1994) notes the potential and/or application for VE technology in flight training, satellite positioning, air traffic control, underwater recovery, mission rehearsal, ground combat training, tactical airspace visualization, and tank combat maneuvers. Macedonia (2002) notes that VE technology has radically changed the way the military prepares for war. For this reason, it is important to evaluate VE technology, its benefits and drawbacks, as a tool appropriate for training military guards.

VIRTUAL ENVIRONMENTS

The purpose of using VE technology for military training is to reproduce the kinds of situations soldiers may encounter in the field, to train responses, emphasize capabilities, and ultimately to provide a vehicle for making and learning from mistakes in an environment where the consequences are not fatal (Smith, 1995). As an aid to decision making, simulations allow soldiers to evaluate the consequences of using one strategy over another and adopt various risk levels in order to determine how a scenario could play out in the real world.

There are several types of simulations used by the military today that can be divided into three categories: live, virtual, or constructive (Pew & Mavor, 1998, Smith, 1995). Live simulations are based on traditional 'role-playing' where trainees operate real equipment but often with simulated weapon firing. Live simulations have limits. It is difficult to cross rivers, terrain, and destroy buildings without a high cost and risk (Smith, 1995). A VE is a computer-generated environment that either simulates the real world or a fantasy world in which people are provided an opportunity to visualize and interact with images and sounds of events that may otherwise be difficult, dangerous, or unavailable to use. Constructive simulations are the most widely used simulations in the military (Pew & Mavor, 1998). Constructive simulations involve virtual humans operating virtual equipment with real people providing inputs and parameter settings. For example, BATTLEMODEL is a simulation framework used to support tasks including fighter combat, strike missions, airborne early warning and control, as well as maritime operations (Heinze et al., 2002). Minimally, military simulations are used for three reasons: (a) to train individual combatants, leaders, or teams; (b) to analyze systems, doctrine and tactics; and (c) to answer questions related to the improvement of command and control and the interoperability of joint forces (Pew & Mavor, 1998). For example, the Marines used a Marine Tactical Warfare Simulator (MTWS) to supplement field training with computer simulated war games that graphically portray, land, air, and sea forces in a simulated war that has become a backbone to combat training (Ewing, 1998). In this case, admittedly soldiers would still need to practice coinciding skills such as traversing 20 miles with an 80-pound rucksack over rough terrain, however, the 10-day maneuvers saved millions of dollars in jet fuel, ammunition costs, and costs associated with planning and executing amphibious landings (Ewing, 1998).

Virtual environments have different levels of immersion. Immersion is a function of the system's ability to generate sensory information (Mazuryk & Gervautz, 1996). Virtual environments can be categorized by one of three levels of immersion. Desktop VEs are the simplest form of VE that uses a conventional PC monitor to display an image of the world. Fish Tank VE is an improvement upon desktop VEs in the sense that fish tank systems support head tracking and enhance the participant's sense of 'being there' in the virtual world. These systems are often paired with LCD shutter glasses that enable stereoscopic viewing. Immersive systems allow participants to become immersed in the VE via head mounted displays that allow tracking and support stereoscopic view. Immersive systems may involve aural, haptic, and sensory interfaces (Mazuryk & Gervautz, 1996). The CAVE Automatic Virtual Environment (CAVE) is an immersive environment that offers suspension of disbelief and a viewer-centered perspective (Cruz-Neira, Sandin, DeFanti, Kenyon, & Hart, 1992).

History of Virtual Environments

The concept of VE is not new. Ulate (2002) indicates that the military's use of instruments that mimic reality can be traced back to the use of flight simulators built by the Link Company in the late 1920s through the 1930s. These simulators were built so that they resembled the front portion of an airplane and were attached to a pedestal that could be used to simulate motion to train World War II pilots how to fly by instruments for night missions. In Blade and Padgett's (2002) compilation of the history of VEs, they argue that VEs date back to the *The Veldt,* a 1950s science fiction story in which Bradbury describes a children's playroom designed to immerse children in an African landscape. Years later in 1961, a multi-sensory simulator named *Sensorama* boasted of the ability to immerse a person in another world by incorporating stereopsis (3D), wide vision, motion, color, stereo sound, aroma, wind, and vibration (Blade & Padgett, 2002), however, it was not interactive. Interactive devices soon followed including Philco's head mounted display (1961) and the dataglove, created by VPL research in 1977, that may have inspired Mattell's powerglove to be the best selling childrens' toy in 1988 (Blade & Padgett, 2002).

Virtual environments have a history of being used for military applications (Pew & Mavor, 1998). In 1983, the U.S. Defense Advanced Projects Agency (DARPA) funded the SIMulator NETwork program (SIMNET) with the goal of creating a proof-of-concept demonstration of interactive training for battle engagement and war-gaming including tanks, aircraft, and command and control structures (Alluisi, 1991; Caird, 1994). This real-time, person-in-the-loop battle engagement simulation was intended to enable individuals to train collectively and interactively. This combat skills tool was aimed at providing lower cost training by using a modular, incremental development process (Caird, 1994). SIMNET marked a deviation from single user systems and applied VE technology in a team context whereby military personnel could practice tank combat maneuvers against other manned tank crews via a network of a lower fidelity simulators (Caird, 1994). SIMNET enabled the training of commanders, familiarization of soldiers with enemy terrain, and the evaluation of alternative weapon systems prior to acquisition. The first operations allowed tank combatants to practice operations including a hasty attack, deliberate attack, hasty defense and passage of lines (Caird, 1994). The Virtual Environment Debrief Interface (VEDI) was developed to assist pilots in visualizing large scale spatial relationships with respect to other aircraft and allowed pilots to practice critical air intercept scenarios (Caird, 1994). The MARS Virtual Reality Simulator was developed with the goal of training Naval Officers in decision making and spatial skills (e.g., space-time trajectories) necessary to acquire ship positions and maneuver a ship appropriately (Caird, 1994; Magee, 1997). The military has also used VE technology to train maintenance practices of military equipment (Caird, 1994).

The CAVE, originally designed by the University of Illinois at Chicago, is a tool for immersive environment research. An immersive environment is one in which the participant feels present in an alternative reality, as compared to non-immersive PC screen-based systems, e.g., use of a pc and a joystick to navigate (Rose & Foreman, 1999). The CAVE is an immersive environment where images can be projected on the walls, floor, and ceiling of a room that surrounds the viewer (Blade & Padgett, 2002), thus reducing feedback from the real world surroundings. The CAVE is often used in conjunction with a magnetic tracking device that manipulates the person's field of view based on head position. The loss of immersion in PC screen-based VEs may be less emotionally engaging and consequently, may translate into a detrimental effect on learning. Ulate (2002) suggests that the lack of consideration for emotional arousal in the design of such VEs may reduce training effectiveness.

Advantages of Virtual Environments

Virtual environments provide an ideal method for training because they provide a natural compromise between performance on real-world tasks and on those in a controlled laboratory setting. Furthermore, as Ulate (2002) states, "the military's motto 'train like we fight, fight like we train' can be put into action without incurring significant fiscal cost, or more importantly, endangering human life" (p. 3).

Virtual environments are also a beneficial research tool. Rose and Foreman (1999) noted VR's ability to isolate participants from their normal sensory environment, by placing them in computer-generated environments that permit infinite flexibility and controllability. This environment provides an advantageous setting for psychologists to monitor human behavior in great detail. Virtual environments "can be a complex large-scale environment, yet, crucially, the experimenter retains complete control over the sensory array that the participant experiences" (Rose & Foreman, 1999, p. 550).

The U.S. Army uses simulation to decrease training time and expenses as well as to increase the realism of training events (Pew & Mavor, 1998). Simulators and VEs are attractive because of their potential to save resources when the nature of the task as well as space, time, and financial limitations dictate the need for alternatives to real world assessment procedures. In fact, for certain occupations the ability to make poor decisions in a simulator and learn from them instead of making the same decision in the real world is ideal. For example, in the case of firefighters, simulated training scenarios are a critical alternative to sending a junior person into a dangerous situation with the knowledge and ability to perform the task without suffering injury. Virtual environments are able to serve as performance indicators and provide an element of safety.

Another potential advantage of using VE technology relates to skill retention. As the length of time between learning and application increases, skills may deteriorate. For example, Navy ships typically go on extended deployments in order to practice, reinforce, and maintain skills in preparation for war that may occur soon or years later. It is expensive to deploy troops in order to maintain skill sets. Because extended practice of skills is recognized as an important factor in skill retention (Hall, Stiles, & Horwitz, 1998), it is noteworthy that VR has the ability to offset skill decay by offering alternative methods practice, experimentation, and application of the skill set repetitiously if necessary.

Advantages and disadvantages of VE training systems have been summarized by Caird (1994), and are shown in Table 1. In sum, VEs offer advantages over real environments as they allow for training, encompassing observation, evaluation, redundancy and safety in a controlled environment (Rose & Foreman, 1999).

Applications of Virtual Environments

Virtual environments have been used in a wide range of applications in academia, games, commerce, weapons systems, military exercises, engineering simulations, medicine, and human factors to name a few. A review of human factors issues by Stanney, Mourant, and Kennedy (1998) indicates that VR can be used as a training tool in the medical field, a system analysis and design tool for the human factors field, an aid for

Table 1.

Advantages and Disadvantages of Virtual Environment Training

Advantages	Disadvantages
Allows user to see and interact with approximations of real, abstract, and imaginary worlds in real-time.	Has too many expectations placed on it to succeed in immediate term.
Commercially available in various forms.	Access to and development of high fidelity technology requires money and highly skilled interdisciplinary team of technicians.
Cost of VE technology continues to decrease.	Cost is still relatively expensive.
Provides flexible training medium.	Flexibility incurs a programming cost.
Flexibility of tool with potential to revolutionize training.	Due to its relatively early stage of development, technical difficulties still limit the clarity of sight and range of possible interactions.
Enables individualized, self-paced learning.	Required degree of physical fidelity between virtual tasks/environments and their real world counterparts is unknown.
Potential to enhance training for a range of job tasks.	Lack of evaluation information within individual and organization needs that would otherwise allow an indicator of effectiveness.

predicting market trends and financial gains in the stock market, a visualization tool for complex data in the scientific community, and a visualization and training system for war scenarios in the military. Pew and Mavor (1998) note that VEs are "applied to address human engineering concerns, the design of systems and their interoperability with other services or multinational forces, and option prioritization and risk assessment decisions, as well as to examine survivability, vulnerability, reliability, and maintainability" (p. 36). Although a vast majority of the public's exposure to VEs has been in the area of entertainment and high fidelity video games, there are also numerous applications in which VEs are being used as a training tool. Virtual environments have been used to train ground control personnel on the operability of the Hubble telescope (Loftin & Kenney, 1995; Loftin et al., 1997), to training battlefield engagement and war-gaming skills (e.g., SIMNET), baseball skills (Andersson, 1993), naval ship maneuvers (Magee, 1997), air traffic control (ATC) operations, situational training with applicability to law-enforcement, ATC, emergency response units (Stansfield, Shawver, Rogers, & Hightower, 1995), and resolution of hostage situations (Reintzell, 1997). Due to the extensive use of VEs in the military, it is important to assess the advantages, limitations, effectiveness, and appropriateness of using VE as alternatives to traditional methods of training.

Virtual Environments for Training

Aside from interactive computer games, training is arguably the fastest growing application for VR (Psotka, 1995). The utility of VEs for training purposes depends on a number of factors. Stanney et al. (1998) provide a number of human factors questions that pertain to how the VE is received by the user and how effective the VE is for training purposes.

Figure 1 illustrates a number of important considerations. First, developers of VEs must consider whether the environment will cause the user to feel discomfort, adverse effects, or simulator sickness. Prolonged exposure to a VE (e.g., over 45 minutes) may cause simulator sickness to be a greater concern. Kolasinski (1995) suggests there is more to simulator sickness than simply concluding its cause to be inconsistent information about body orientation and motion. Other factors may also influence the occurrence or intensity of simulator sickness.

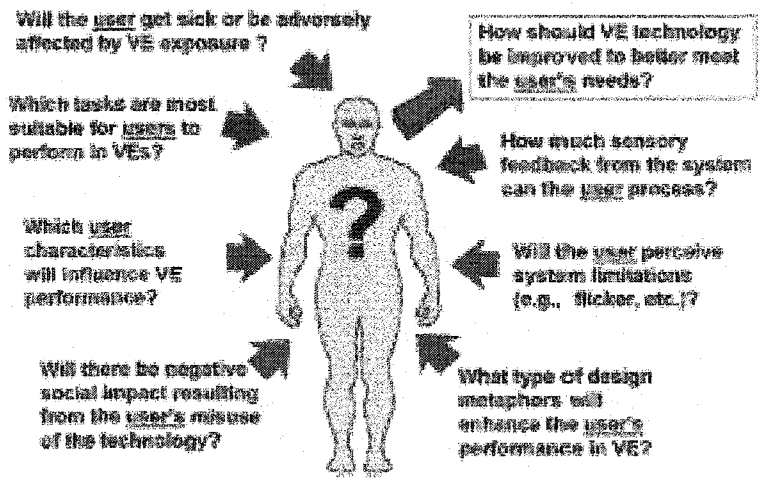

Figure 1. The human factor in virtual environments: Design considerations for an effective virtual environment training tool (Stanney, Mourant, & Kennedy, 1998, p. 328).

Second, research is needed to identify which tasks may be more conducive to VE training and the limitations that help human factors specialists design effective training tools. Conceivably, there are tasks that are a better fit for VEs. Human factors specialists play an important role in identifying elements that characterize tasks suited for VEs, (e.g., elements related to the capabilities and limitations of users). A third important research area is the identification of user traits that contribute to VE performance. Fourth, Stanney et al. (1998) advise developers to be aware of the potential for negative social impact as a result of misuse of technology. For example, researchers argue that interacting with simulated humans may have a negative impact in terms emotional detachment. Others argue the social impact of violent video game technology is the propagation of violent behavior in children (Anderson & Bushman, 2001; Miller, 2001). A fifth noteworthy issue concerns human sensory and perceptual limits in processing feedback provided in VEs. In addition to the user's own sensory limits, there is also an issue of the user's awareness of the limitations of the VE system and how this affects his or her interaction and consequent performance. Finally, designers of VEs should consider the type of metaphors that will provide a baseline familiarity of the elements and interaction they can expect during their virtual experience. Although all of these factors are important considerations, the current study is especially concerned with the identification of tasks for that are well-suited for the VE.

TRANSFER OF TRAINING

Transfer of training concerns the manner in which previous learning impacts new learning and performance (Swezey & Andrews, 2001). This transfer is based on the idea that learning one task affects learning on another task and the greater degree of similarity among tasks, the greater degree of training transfer (Holding, 1965). Traditionally, transfer of training has been discussed in the context of stimulus-response theories of learning with several factors implied as predictors of training transfer (Holding, 1965). This similarity concerns two issues: (a) the similarities between the two tasks, and (b) the similarities in the related responses. Holding (1965) states that if two tasks present identical stimuli and require the same responses, then transfer of training should be ideal. Alternatively, if the task stimuli and response sets are different, training transfer will not occur. Thus, dissimilar tasks such as basketball and cooking each have different response sets. Consequently, transfer of training between cooking and basketball will most likely fail. Task difficulty affects training transfer in that there is differential training transfer between tasks of varying levels of difficulty. For example, Holding (1965) suggests that playing an organ may result in better transfer to learning the piano than the reverse (i.e., moving from piano to organ). Greater training transfer results when going from a more difficult task to an easier task than when going from an easy task to a more difficult task. Other factors that affect transfer of training include the relationship between retrieval cues and encoded information, the type of instructional techniques used to enable the integration and gleaning of information, organizational strategies used for information processing, and automation of performance with consistent training (Swezey & Andrews, 2001).

There are two theories related to transfer of training: identical elements and transfer-through-principles. According to Thorndike and Woodworth's (1903) identical elements theory, training transfer occurs in situations where the training environment and the transfer (operational) environment contain identical elements (Lintern, 2001). This theory is based on the notion that training on one type of task or activity will transfer to another activity as long as there are common elements or features of the task such as aims, methods, and approaches (Ellis, 1969). Accordingly, when a novel situation presents itself, a person capitalizes on the common elements between the novel situation and previous experiences (Swezey & Andrews, 2001). Support for this theory depends on the ability to identify the extent of similarity between the stimulus and response mechanisms in the transfer condition and the operational environment condition. Since most environments are complex, this distinction is not always easy to infer. Judd (1908) scrutinized the identical elements theory by arguing that the critical factor in transfer of training is the ability to abstract general rules or principles, a theory Judd referred to as generalization or transfer-through-principles (as cited in Ellis, 1969).

The transfer-through-principles theory differs from the identical elements theory because there is no underlying assumption that a learner needs to be consciously aware of similar elements for transfer of training to occur. Instead, it is assumed that the learner can use previously acquired principles and apply them to the operational setting (Swezey & Andrews, 2001). Transfer of principles is typically called upon in tasks of a problem-solving nature (Holding, 1965). For example, it would be reasonable to assume that teaching students the principles of addition and subtraction will result in transfer to accounting tasks where maintaining the books involves tracking credits and debits. In a military context, teaching soldiers about defensive positioning that allows team members to avoid being hit by cross fire should result in transfer to a military checkpoint task or other scenarios such as urban warfare where defensive positioning is critical.

There are two types of training transfer, positive and negative. Positive transfer occurs when knowledge acquisition facilitates performance and negative transfer occurs when the performance or experience interferes or inhibits performance on a second task thus resulting in a decrease in performance (Ellis, 1969).

The principles underlying transfer of training rely on the generalizability of knowledge to other tasks or contexts. When humans respond to a particular stimulus or set of stimuli, they tend to respond similarly the next time they are in a similar situation. Consider a simple household example by Holding (1965). If in the rush of cooking, a chef splatters the oven with tomato sauce, it becomes evident that the more appropriate response would have been to use a lid or aluminum foil to cover a pot. Thus, the chef learns to use a cover in subsequent cooking sessions when tomato sauce is involved. In another situation, the chef may intend to cook tomato sauce in a microwave oven. Past experience will tell him or her to cover the dish, however, covering that dish with a steel lid or aluminum foil would be a mistake in context of microwave cooking. That is, if a different response becomes required for the event, the old response may be inappropriately executed resulting in negative transfer (Holding, 1965).

Swezey and Andrews (2001) state that in many environments, it is not a question of whether there is positive or negative transfer. Instead, sometimes a hybrid created by the interaction between elements fosters a situation where there can be positive transfer of some elements and negative transfer of others. It is also possible that zero transfer can occur in which earlier performance has no effect on subsequent performance (Ellis, 1969).

Obviously, one of the benefits of simulation is presumed to be positive transfer from an artificial environment to the real world. For example, flight simulators have been used to train commercial and military pilots since WWII (Rolfe & Staples, 1986). Research is still being conducted to examine which tasks and real world conditions are more suitable to be simulated in comparison to more traditional methods. It is hoped that what the trainee learns in the simulated environment will generalize to a different situation. Of course, it is also possible that a set of factors (e.g., varying cues, lack of detail, environmental factors or otherwise different or conflicting cues) can result in negative transfer of learning in which the goals are not achieved, or even worse, the training results in a decrement in performance.

One goal of a VE is to immerse the participant in an alternate version of 'reality' such that they feel they exist in the simulated world and are able to interact with people and objects in that environment as if that world existed. This goal is especially important in certain training applications where it is important to have a trainee's attention focused on the task at hand and not on artifacts within the environment or interface that detract from realism. Skill transfer is dependent on the commonality between the training environment and operational environment (Lintern, 2001). If the virtual world contains cues that are inconsistent with a person's mental model of the real world (e.g., incompatible depth cues and/or temporal stimulus-response relationships) it may have a negative impact on training. Designing a goal-oriented VE that matches the user's mental model and is compatible with the human sensory system is critical to training effectiveness.

Military Checkpoint Training Transfer

Transfer of training using VE technology has been successfully established in a military context. For example, naval ship officers who used the Maritime Surface / Subsurface Virtual Reality Simulator (MARS VRS) demonstrated positive training transfer in comparison to those officers who had not used the system (Defence and Civil Institute of Environmental Medicine (DCIEM) as cited by Caird, 1994, p. 47). Virtual reality has potential as a valuable training tool for numerous military tasks. It is important to explore tasks and contexts for which VR training is most appropriate, identify those that are a good fit to existing technological capabilities, and document shortcomings or mismatches between a VR and its applicability as a training tool for a particular task.

The present study examined the application of VE technology for military checkpoint training. The VE in this study was designed to simulate checkpoint tasks. The trainees were tasked with maintaining a checkpoint. Drivers approached the checkpoint seeking entrance to a fictitious base. The trainee inspected each vehicle, obtained identification for all occupants, and made decisions as to whether the drivers could enter the base.

Participants were presented with a set of scenarios developed to represent common events at a checkpoint. Specifically, the scenarios addressed the trainee's ability to: (a) handle matters of situational urgency according to procedure; (b) contend with social pressures that conflict with procedure; (c) recall and identify vehicles, people, and license plates from a predefined target list; (d) detect inappropriate features such as contraband and missing elements; and (e) carry out procedures while simultaneously maintaining situation awareness.

The categories included events that are aimed at evaluating whether a trainee was vulnerable to social and environmental time pressures. For example, in one scenario, an ambulance arrived without proper authorization via radio alert and the driver advised the trainee that he did not have time to go through the normal identification verification routine because he had an injured passenger. Guards must uphold the security of the base and should know to follow proper procedure and perform an identification check on both the driver and passenger even if confronted with an urgent situation. Second, trainees were evaluated on their ability to identify and resist social pressure when it conflicts with procedure. Social pressure was created by introducing people of social status, but whose situation creates a conflict with standard procedures. For example, the city mayor appeared in one scenario requesting entrance to the base without a proper visitor's pass. The mayor argued that due to her status she did not need a pass. The third category of evaluation was related to memory skills. In each session, trainees were presented with specific people or vehicles to remember and watch out for during their shift. Consequently, this task required long-term retention of a facial image, license plate number, or vehicle description. Fourth, trainees were evaluated on perceptual issues. For example, trainees had to recognize the absence of a feature such as a license plate, base sticker (decal), or visitor pass. Trainees were also instructed to identify contraband items, such as guns, knives, or unmarked packages. Fifth, trainees were evaluated on their ability to maintain situation awareness. Situation awareness was tested when the trainee attended to a driver who demonstrated suspicious behavior implicating suspicious behavior in a second vehicle. In order to demonstrate good situational awareness, the individual had to attend to the first vehicle, while monitoring the activity of the second vehicle. In this particular scenario, participants who failed to ask for backup from their virtual partner encountered gun fire from the second vehicle. There was a sixth category that was integral to all tasks; the ability to follow proper procedure. For example, the trainees were required to know that a military vehicle must have a base sticker, whereas a civilian vehicle requires a visitor's pass. The present study focused on VE training as a means of effective transfer of skills for military watchstanders. A simulated virtual checkpoint containing multiple scenarios was designed to address decision-making ability in the six categories described above. These categories were designed to be representative of basic skills required of a checkpoint guard.

The current study had two specific objectives. The first objective was to create a realistic simulation of a military checkpoint in a CAVE in an effort to determine whether VE technology is effective for training individuals to handle critical events at a military checkpoint. The CAVE was chosen because of its immersive nature (Blade & Padgett, 2002). As Rose and Foreman (1999) indicate, sense of immersion is enhanced through the use of stereoscopic glasses, increase in screen size and reduction of background lighting. The CAVE supports all of these features, thus reducing distracting feedback from the real world.

The experiment was run over two sessions. The first session contained scenarios designed to measure skills in the categories described above and the second session contained scenarios that were similar to events contained in the first session. For instance, in Session 1, a scenario addressing social influence involves a mayor who has proper identification and makes a plea to enter the base without the required visitor pass. In the second session, transfer of training is tested when an admiral's son claims to live on base but fails to show proper identification. If the trainee made an error and showed vulnerability to social influence in Session 1, positive transfer of training would be demonstrated in Session 2 if they did not repeat the same mistake. It was expected that individuals who participated in a VE training session and received feedback on that session would commit fewer errors than those who did not receive training and feedback.

PREDICTORS OF PERFORMANCE

Most performance research has focused on two classes of variables as predictors of job performance: ability and personality factors (Sackett, Gruys, & Ellingson, 1998). Ability factors have received much attention because they indicate whether a person has the physical or mental ability to perform a task; however, personality factors can also be an important determinant of performance because of their possible connection to constructs such as a person's motivation, attitude, and approach to performing a task. Although ability is important with respect to whether a person 'can do' a job, personality provides an indication of whether a person 'will do' a given job (Mount & Barrick, 1995; Sackett et al., 1998). Further, it could be said that personality affects a person's strategy in how she or he approaches a problem (e.g., being proactive, passive, impulsive, etc.). A common selection model of performance is often based on the idea that performance is a multiplicative function of motivation and ability (Maier, 1965). Ability refers to an individual's potential performance whereas a person's performance refers to what an individual actually does in a given condition (Maier, 1965).

Personality traits are said to offer potential insight into one's motivational processes at various stages (Hollenbeck, Brief, Whitener, & Pauli, 1988; Mount & Barrick, 1995; Phillips & Gully, 1997). Hollenbeck and Whitener's (1988) literature review indicates that personality measures are valid only when used in conjunction with ability tests and then, only when there is a strong theoretical foundation to support a relationship between a particular trait and motivation. Hollenbeck and Whitener (1988) suggest that personality traits are more strongly related to an individual's motivation to perform the task than the individual's ability to perform the task. Hollenbeck and Whitener's (1988) illustration of the relationship between personality traits, motivation, ability and performance is depicted in Figure 2 below.

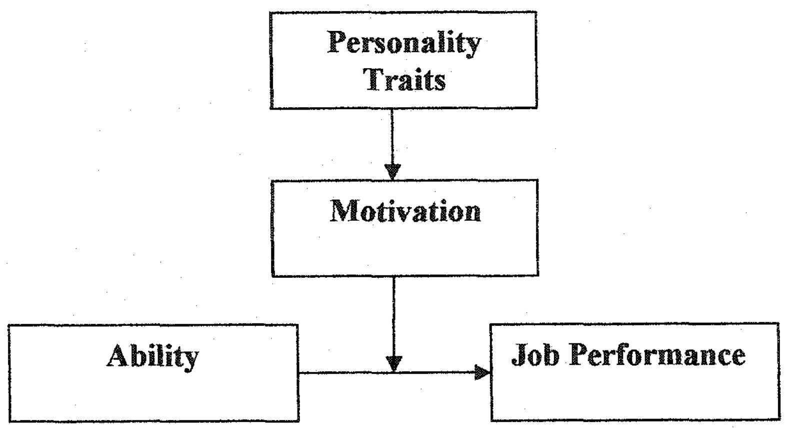

Figure 2. Personality and the moderating effects of motivation and ability (Hollenbeck & Whitener, 1988).

Personality traits and job performance are also moderated by the degree to which a job involves autonomy as well as the demand of the environment (Barrick & Mount, 1993). A weak situation involves few demands and pressures to conform whereas a strong situation involves significant demands and pressures to conform. Weak situations allow more room for an individual to use discretion in determining which behaviors to exhibit. In strong situations, the external environment hampers an individual's freedom to act in a less inhibited manner. Assembly line work is considered a strong environment because the procedures, sequences, and methods are strictly laid out for employees and there is limited opportunity to implement individualized methods. In contrast, weak environments offer increased flexibility in choice of action. For example, an individual may be instructed to complete a task, but the method of completion is left up to the individual. Barrick and Mount (1993) indicate individual differences are more likely to show up in weak situations because personality traits are able to play a stronger role in the choice of behaviors a person exhibits. This distinction between weak and strong environments is particularly applicable to the military because it is a very structured organization with an abundance of procedures that are trained to reduce individual variability. Identifying the influence of personality trait predictors will be more challenging in a very structured environment and must be kept in mind when considering the effect of potential predictors. As Barrick and Mount (1993) note, most research in this area has been done in controlled laboratory settings where the characteristics of strong and weak situations can be experimentally manipulated.

Personality

Personality is a difficult concept to define concretely. Much research has been devoted to dividing personality into component traits, a task that is not easy considering the complex interactions among the individual elements. Despite these difficulties, there is a general consensus among researchers on five broad factors of personality: extraversion, emotional stability (neuroticism), agreeableness, conscientiousness, and culture, also known as openness; (Barrick & Mount, 1991; Digman, 1990; Goldberg, 1990). Further, an abundance of research has provided growing evidence that the five factor model is robust across varying theoretical frameworks, instruments, and cultures using ratings from varying sources and with a variety of samples (Barrick & Mount, 1991; Digman, 1989).

Costa and McCrae's (1992) five factor model has evolved as one of the most widely accepted representations of personality and forms the basis of the commonly used NEO Personality Inventory (Costa & McCrae, 1992, Matthews & Deary, 1998). Costa and McCrae posit five areas of personality commonly referred to by the acronym, *OCEAN*: Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism. *Openness* consists of six subfacets: fantasy, aesthetics, feelings, actions, ideas, and values. *Conscientiousness* is represented by subfacets including competence, order, dutifulness, achievement striving, self-discipline, and deliberation. *Extraversion* is divided into the facets of warmth, gregariousness, assertiveness, activity, excitement-seeking, and positive emotions. *Agreeableness* is represented by the facets of trust, straightforwardness, altruism, compliance, modesty, and tender-mindedness. *Neuroticism* is composed of six facets: anxiety, hostility, depression, self-consciousness, impulsiveness, and vulnerability. Table 2 contains a brief description of each trait. Research suggests that nearly all personality variables in some form can be categorized under the five factor model (Goldberg, 1990), although it should be noted that there are still some researchers who feel that five dimensions are not adequate to represent the entire domain of personality (Barrick & Mount, 1991).

Personality and Performance Prediction

Several researchers have reported that at least three of the factors from the Big Five (extraversion, conscientiousness, and neuroticism) are associated with job performance and career success (Judge, Higgins, Thoresen, & Barrick, 1999). Conscientiousness, agreeableness, and neuroticism (emotional stability) have been shown to be positively related to jobs that require interpersonal interactions (Mount, Barrick, & Stewart, 1998). With regard to measures of cooperation and interaction,

Table 2

Summary Descriptors of the Big Five Trait Dimensions

Trait Name	Description
Openness	Curious, broad interests, creative, original, imaginative, untraditional
Conscientiousness	Organized, reliable, hard-working, self-disciplined, honest, clean
Extraversion	Sociable, active, talkative, optimistic, fun-loving, affectionate
Agreeableness	Good-natured, trusting, helpful, forgiving, gullible, straightforward
Neuroticism	Worries, nervous, emotional, insecure, inadequate, hypochondriac

(Costa & McCrae, 1989)

Hough's (1992) meta-analysis found that Agreeableness (r =.17), facets of Conscientiousness including achievement (r = .14) and dependability (r = .17), and Emotional Stability / Neuroticism (r = .13) were related to the criterion of teamwork. Mount, Barrick, and Stewart also found that Emotional Stability / Neuroticism, and Agreeableness were significantly correlated with performance in work teams (r = .27 and .33, respectively) moreso than in dyadic service jobs (r = .12 and .13, respectively). Research suggests extraversion is a valid predictor of job performance in contexts that involve social interaction (Barrick & Mount, 1991). However, this relationship is not always observed, in part, because of moderating effects of organizational contexts (e.g., reward structure and cooperative nature). Mount et al. (1998) found conscientiousness is more highly correlated with performance in nonteam settings (r = .29) than in team settings (r = .21). In contrast, they found that Openness to experience is moderately correlated with performance regardless of whether it is in a team setting ($r = .16$) or nonteam setting ($r = .17$).

Independent of whether a job involves social interaction, it has been shown that of all the big five factors, conscientiousness appears to be the trait most consistently tied to performance (Barrick & Mount, 1991). Furthermore, conscientiousness is less sensitive to moderating variables (Matthews & Deary, 1998).

Personality and Military Task Performance

Personality traits have been met with renewed interest and enthusiasm with regard to their role in understanding work behavior (Mount & Barrick, 1995). Historically, interest in personality has been tied to military applications. For example, personality traits have been examined as predictors of military tasks including explosive ordnance training (Hogan & Hogan, 1989), military diving (Biersner & LaRocco, 1983), and pilot combat performance (Siem & Murray, 1994). The first personality inventory, the Woodworth Personnel Data Sheet, was used during World War I as a screening tool to identify recruits who might be prone to war time stress (Mount & Barrick, 1995). In the late 1950s, personality traits were used to predict officer effectiveness in the U.S. Air Force (Tupes, 1957 as cited in Digman, 1990). By World War II, personality inventories found wide spread usage in the military and by 1960, a significant amount of research emerged (over 40 years of factor-analytic studies) that supported the five factor taxonomy of personality (Digman, 1989; Mount & Barrick, 1995).

A main concern in the military is the proper assessment and placement of personnel (Driskell, Hogan, Salas, & Hoskin, 1994). Currently, the military relies primarily on cognitive measures, (e.g., the Armed Services Vocational Aptitude Battery;

ASVAB), to make assessment and placement decisions (Hogan & Hogan, 1989). Tools other than (or in addition to) the ASVAB may be necessary since use of the ASVAB as a sole selection tool is said to have a 20-35% failure rate (i.e., identifying those who fail to complete training in the allowed time frame; Nauta, Ward, & D'Ambrosia, 1983, as cited in Driskell et al., 1994).

Researchers have traditionally relied on cognitive measures, often leaving out personality measures that may also provide beneficial information. Sackett, Gruys, and Ellingson (1998) were able to improve predictions of job performance for soldiers by adding personality measures to existing ability test batteries. Such research suggests that there is great potential for selecting personnel based on batteries that include measures of both ability and personality. Driskell et al. (1994) state that historically the military has rarely used personality measures in training decisions. Hogan and Hogan (1989) note that cognitive ability scores are the primary tool used to select personnel for technical training in the military. Cognitive measures are viewed as the primary and often sole basis for placement decisions. This perspective slights noncognitive factors that also impact performance; "Individuals differ not only in ability but in achievement orientation, conscientiousness, and other motivational factors that are likely to affect training performance" (Hogan & Hogan, 1989, p. 32). These factors may be just as influential as cognitive abilities (Day & Silverman, 1989; Driskell et al., 1994; Hogan & Hogan, 1989).

In an effort to develop and validate a selection and placement battery for the U.S. Army, McHenry, Hough, Toquam, Hanson, and Ashworth's (1990) suggested that the military can improve performance predictions by including noncognitive factors in its current test battery. Unfortunately, there is a lack of research in this area (Rosse, Miller,

& Barnes, 1991). Barrick and Mount (1991) performed a meta-analysis on 117 studies and found that measures of personality, as related to training proficiency, are lacking in the literature. Specifically, training proficiency was included in only 12% of their sample of studies.

Personality measures provide attitudinal and motivational indicators of performance that cognitive assessment tools do not necessarily address (Barrick & Mount, 1991). Accordingly, the second objective of the present study was to use measures of personality to predict decision-making outcomes in a simulated military task. The empirical evidence gathered may have implications for: (a) the development of decision aids and selection tools, (b) identifying how a given watchstander might react to a situation, and (c) how to compensate for possible characteristics that result in suboptimal performance.

Silver (1992) studied the behavior of twentieth century infantry and carried out extensive research on battlefield behavior. Silver's work is based on over 20 years of qualitative analysis via self report memoirs, histories, case studies in a clinical capacity, and compilations of information from therapists of war veterans. Silver arrived at eight traits that impact decision-making: stability, anxiety, anger, humor, acquiescence, independence, charisma, and knowledge. These are personality traits applicable to military functions. Silver's (1992) battlefield behavior model also integrated factors such as the existing morale of the troops as a whole (a combinative function of personality traits, stress, support an leadership), situational stress (e.g., the friend-to-foe ratio), leadership, resentment, adhesion, and support. Silver's findings are based on literature and extensive clinical observation of individuals (documented in case studies) specifically related to battlefield and post-traumatic stress behavior. As conceptually significant as Silver's proposed traits are, the challenge becomes one of finding an adequate assessment tool tailored to these particular traits. There is no direct and appropriate measurement tool for these traits and therefore such traits can only be inferred from traits measured in tests such as the Costa and McCrae's (1989) personality inventory (NEO-PI).

Personality interacts with a number of significant characteristics including cognitive ability, intelligence, and self-esteem. To an extent, these factors are capable of being inferred from certain facets of the NEO-PI. However, other measures may be necessary to confirm the relationship between personality and cognitive or self-esteem factors.

Specific Personality Traits as Predictors of Simulated Military Task Performance

There is a foundation in the literature and a research paper by Silver (1992) relating job performance to personality traits. Thus far, descriptions of personality traits and their importance as potential predictors of performance on military tasks or in VEs have been introduced as a potential selection tool for identifying characteristics that aid in improving military task performance (such as the skills and knowledge necessary for standing watch under simplified military checkpoint conditions). Still, the literature appears to lack research that can tie personality traits to tasks relating to protocol that would appear to have face validity and generalizability for predicting performance on a military checkpoint task. Thus, one may have to extrapolate from Barrick and Mount's (1991) meta-analysis incorporating the big five traits and identifying their predictive power in skilled/semi-skilled occupations, (e.g., production workers, assemblers, accountants, and secretaries) and professional jobs (e.g., engineers and architects). These jobs appear to require attention to procedural detail and may therefore be somewhat generalizable to the job of military checkpoint guards. In the next section, the potential applicability of specific personality traits as predictors of success in military checkpoint guard performance will be considered.

Openness. Openness to experience has been said to be one of the more difficult factors to identify (Digman, 1989). Openness consists of six facets: fantasy, aesthetics, feelings, actions, ideas, and values. Characteristics such as creativity and cognitive ability are said to be correlated with the openness dimension measured in the NEO personality inventory. For example, research has indicated that aside from intelligence tests, intelligence can be inferred from a person's rating on the openness to experience scale in the NEO personality inventory (Barrick & Mount, 1991). Judge et al. (1999) state that people high in openness are creative, flexible, and have an 'intellectual orientation' that appears to contribute to success in many occupations. In fact, openness has often been interpreted as intellect or intellectance (Barrick & Mount, 1991; Digman & Inouye, 1986). Openness has the highest correlation of personality measures with respect to intelligence ($r = .30$); however, clearly, the factors measured in intelligence are distinct from those measured in openness (McCrae & Costa, 1987). Further, people high in openness are creative and able to think "outside the box." They are receptive to new information even if it conflicts with what one would expect. A person with high openness is an independent thinker as opposed to conforming in judgment (Costa & McCrae, 1992).

Openness is said to be a strong predictor of overall performance since participants who are curious, broad minded, cultured, and intelligent "have positive attitudes towards learning are thus more motivated and consequently will learn more" (Barrick & Mount, 1991, p. 6). Supporting this idea, Barrick and Mount also indicate that openness to experience is a strong and valid predictor ($\rho = .25$) of training proficiency, implying that high scores on this trait help identify "training ready" persons who are most willing to 'engage in learning experiences' and thus benefit from training. It is further suggested that openness may be measuring a person's ability to learn *in addition to* his or her motivation to learn (Barrick & Mount, 1991). Accordingly, it is hypothesized that in the present study persons high in openness should perform better than persons who are low in openness.

In the present study, participants performed a 45 min task repeated on two occasions. It is conceivable that a person may not describe the task as routine until they have enough familiarity with the task that it seems repetitive. Thus, it is unclear whether this effect, even if it exists, would be evident in the cumulative time period used to familiarize oneself with the task in the present study. Furthermore, if a person with high openness is not interested in the tasks, a question is raised as to whether motivational issues may cause a decrement in performance on a task. Janis (1989) states that openness, conscientiousness, and neuroticism may be associated with motivational states. Specifically, a lack of openness may be indicative of a motivational deficiency. However, it might also be true that a lack of motivation is potentially compensated for by other personality traits that may cause the person to say that although the task is not interesting, it does not discount the responsibility of the person to perform the task to the best of their ability. Given the summary of the components of openness, persons high in openness were expected to make fewer errors in the second session than persons low in openness.

Conscientiousness. Conscientiousness has been reported as one of the more consistent predictors of performance across all occupational groups (Barrick & Mount, 1991; Barrick, Mount, & Judge, 1999; Hough, Eaton, Dunnette, Kamp & McCloy, 1990; Piedmont & Weinstein, 1994; Salgado, 1998). It includes six facets: competence, order, dutifulness, achievement striving, self-discipline and deliberation. Salgado (1998) indicates evidence of generalization of conscientiousness across all occupations ($\rho = .23$). A person who is highly conscientious is generally hardworking, well-organized, persistent, and careful. In addition to being a responsible individual, Digman (1989) states that a person high in conscientiousness is not only reliable, but will perform 'with conscience.' Thus, it could be said that conscientiousness taps into internal motivational / reward structures and has even been said to be the dimension that organizes and directs behavior (e.g., Digman & Takemoto-Chock, 1981).

Conscientiousness involves dependability, which in turn, entails an appreciation and concern for the rules (Tett, 1998). Interestingly, conscientiousness has often been associated with volitional variables implying achievement orientation and perseverance. The conscientiousness construct is closely tied to the concept of a hard-working person. Reflecting on the strong tie between conscientiousness and "hard working," Barrick and Mount (1991) found that the 'work' construct has been used interchangeably with the term 'conscientiousness' (e.g., Peabody & Goldberg, 1989 as cited in Barrick & Mount, 1991). In fact, there appears to be some debate as to whether this trait should be defined in terms of responsibility or dependability or in terms of volitional aspects such as achievement orientation (Barrick & Mount, 1991). In some cases conscientiousness has been used synonymously with ‘will to achieve’ because it is highly correlated with educational achievement (Digman, 1989). Further, Costa and McCrae’s (1992) research suggests the will to achieve that is characteristic of conscientious individuals is independent of external rewards or compensation structure. In a situation where performance ratings are independent of reward, a conscientious person’s will to achieve would be unaffected. Such should be the case in the present study where course credit (an external reward) is given for participation rather than a measure of quality of performance. However, research (Barrick & Mount, 1993) also indicates that high conscientious scores are correlated with higher performance in jobs with high autonomy as opposed to jobs with low autonomy. Therefore, it may be that an inability to account and control for situational strength in military tasks (highly structured, low autonomy) will result in underestimation of the influence of conscientiousness on general task performance.

Situational strength not withstanding, individuals who are reliable, dependable, cautious, organized, planful, persistent, hard-working, and achievement oriented (all of which are found within facets of conscientiousness) tend to be better performers in most occupations (Barrick & Mount, 1991; Mount, Barrick, & Strauss, 1994). In the present study, much of checkpoint task performance involves adhering to strict procedures. Costa, McCrae, and Dye (1991) note a relationship between conscientiousness and a person’s self-discipline, desire for achievement, order, and persistence. Lack of conscientiousness reflects a person who is disorganized, lacks self discipline, and handles day-to-day tasks haphazardly. Such people may fail to give proper attention to details and suspicious cues that may serve to counter a potential a threat. A lack of conscientiousness is said to relate to counterproductive work behaviors (Judge et al., 1999). Conscientious people can identify a goal and more readily keep that goal in mind throughout the activity. One could infer that a careful, organized person is more likely to adhere to instructions and pay attention to necessary details. Thus, it is hypothesized that in the present study a person high in conscientiousness is more likely to comply with procedure and less likely to be vulnerable to social influences and therefore should perform better overall and particularly on procedural tasks and those necessitating resistance to social pressure.

Neuroticism. Neuroticism, often referred to as emotional stability, also shows evidence of validity ($\rho = .23$) as a predictor across all occupations (Salgado, 1998). High levels of neuroticism are indicative of anxiety and low stress tolerance. High levels of neuroticism are generally described as having negative effects on performance (Salgado, 1997). However, neuroticism appears to have a complex relationship with performance. In fact, Barrick and Mount (1991) found that in professional occupations (e.g., engineers, architects), persons who were neurotic were better performers. Barrick, Mount, and Judge (1999) also found neuroticism to be correlated with performance in certain occupations including police, skilled or semi skilled jobs, but not others. This is noteworthy as these particular professional occupations appear to require attention to detail and procedure –characteristics that are potentially important in the present study. Consequently, one might suggest that the anxious, worried nature of a neurotic person creates levels of alertness and motivation that might help detect a higher number of critical scenarios in the present study. However, it may not be that simple. Neuroticism does not consistently predict job performance. Barrick and Mount (1991) found little support for emotional stability measures to predict job performance across all occupations, however, they also point out that this may be a function of a restriction of range in their study.

There are two issues to consider with regard to neuroticism's effect on performance. The first concerns how neurotic individuals handle ambiguity and the second concerns how neuroticism affects the quality of response. With regard to ambiguity, hypervigilance theory (Eysenck, 1992) suggests that people with high trait anxiety will investigate their environment excessively in search of threat. However, without knowing the person's composition of other dominant traits, it is difficult to hypothesize how neuroticism will affect performance on the task as a whole. Anxiety, a component of neuroticism, could influence performance as it "narrows the focus of attention and predisposes toward the detection of threatening stimuli and the interpretation of ambiguous stimuli as dangerous" (Pew & Mavor, 1998, p. 255). Thus, a person who is highly neurotic may be anxious and impulsive and show a tendency for committing a significant number of false alarms. Stated another way, anxiety and the impulsive nature of a neurotic trainee will interfere with the trainee's ability to recall and follow procedure. This interference will be characterized by the inability of a highly neurotic person to distinguish a critical scenario from a non-critical scenario. It is hypothesized that a neurotic person will commit more false alarms than a non-neurotic person.

On the other hand, Milgram and Tenne (2000) found neuroticism to be highly correlated with procrastination in decision making for both minor decisions (where to eat dinner) and major choices (e.g., relationships), $r = .41$ and, 42 respectively. McCown and Johnson (1991) found that neuroticism is related to procrastination, lack of confidence in preparedness, and anxiety in university students during an examination period. Other researchers have indicated neuroticism is negatively correlated with job satisfaction suggesting a potential problem with motivational issues. Such a finding could imply a "will do" versus "can do" issue in which capable neurotic people fail to perform at optimal level.

With regard to the quality of response, the neurotic person appears to adopt a 'better safe than sorry' strategy that may prove useful in environments where double-checking one's intuition could mean the prevention of terrorist threat. Alternatively, a checkpoint guard who is highly neurotic may tend toward impulsive behavior. The impulsive nature may be advantageous because highly neurotic individuals may detect more critical scenarios than those who are low in neuroticism. Still, even though impulsivity may help an individual detect a critical scenario, that same trait may be a liability for handling the critical scenario. That is, the neurotic guard may be able to detect a potential threat, but not handle it well. Overall, in the present study, neuroticism is hypothesized to have a negative correlation with performance evidenced by an increased number of false alarms.

Personality joint effects. Personality traits do not operate in isolation. That is, effects of one personality trait on performance may be moderated by a second personality trait. Hypotheses related to interaction effects are not explicitly stated in this paper, however, these interactions may be present. For example, research suggests conscientiousness is a dimension that "holds impulsive behavior in check" (Hogan, 1986 as cited in McCrae & John, 1992), thus implying that a person high in conscientiousness and neuroticism may be a better performer than a person low in conscientiousness. Likewise, an interaction between openness and emotional stability might also be expected such that neurotic people with high openness will perform better than neurotic people with low levels of openness (Barrick & Mount, 1991). For example, a person who is high in openness and otherwise not interested in the task, may still perform the task to the best of their ability if they are also high in conscientiousness. Janis (1989) states that lack of conscientiousness, openness, and neuroticism may have an association with motivational deficiencies. If this is so, then perhaps such motivational deficiencies could translate into poor performance on a military checkpoint task.

Moreover, neuroticism may also be moderated by other factors such as intellect. Motivational theory (Festinger, 1957; Heider, 1958) suggests that instability is a motivating factor in performance. If intellectance is high and neuroticism is high performance may increase. In other words, a person with high openness to experience and high scores on neuroticism might perform better than a person with high scores on neuroticism but low scores on openness to experience. Accordingly, the final objective of the present study is to examine the relationship among openness, conscientiousness, and neuroticism and their effects on performance.

METHOD

Participants

Thirty-two undergraduate students from Old Dominion University with normal or corrected-to-normal acuity participated in the study. Participants included 6 men and 26 women with an age range of 19 to 38 ($M = 23$, $SD = 4.22$). They were offered either (a) four hours of extra course credit or (b) $30 as compensation for their time. Participants who were predisposed to simulator sickness were screened with the Simulator Sickness Questionnaire (Kennedy, Lane, Berbaum, & Lilienthal, 1993). No participants were excluded from participating. Three participants (one participant from Group 1 and two participants from Group 2) reported previous experience with checkpoint duty and were excluded from the analyses. An additional four participants were replaced due to system malfunctions. One other participant chose not to participate and was replaced.

Task

Participants were trained to perform duties similar to those performed at a military checkpoint, which included inspecting vehicles, verifying the occupants' identification, and monitoring driver behavior. Tasks related to these categories are shown in Table 3 and are more clearly described in the procedure section. This task set was selected under the advisement of military checkpoint personnel (e.g., persons who serve as checkpoint/gate guards, watch commanders, and intelligence officers). These tasks, although representative, were not exhaustive of tasks performed at a checkpoint. Instead, most of these tasks were chosen based on their representative nature and their ability to be implemented within a VE.

Table 3.

Task List

Step #	Task Description
1.	Inspect the vehicle (e.g., license plate)
2.	Obtain appropriate identification for both driver and passengers
3.	Distinguish between appropriate 'reason for entry' and non-appropriate entry of an item. (e.g., bringing a gun on base to return it to an armory versus bringing a gun on base for leisure purposes)
4.	Identify persons or vehicles described as a "Be on the look out for" (BOL)
5.	Follow emergency vehicle protocol by allowing vehicle immediate access only in cases where the vehicle has radioed in advance of arrival
6.	Identify suspicious behavior (e.g., gazing in the rear view mirror excessively)
7.	Identify missing items on person and vehicle (e.g., license plate, identification card)
8.	Identify contraband items including weapons and unmarked packages.
9.	Distinguish protocol for 'pull over' versus 'turn around.'
10.	Turn drivers around upon noticing missing required item (e.g., license plate, ID)
11.	Pull over persons with contraband
12.	Radio in drivers who pass through the gate without authorization
13.	Radio in persons who are denied entry

Scenarios

There were two types of scenarios in the present study, neutral and critical. A neutral scenario was characterized by a vehicle and driver who should be allowed entry. Neutral scenarios did not contain elements requiring special attention or treatment. In other words, a neutral scenario was defined by a driver who had proper identification, whose behavior did not appear suspicious, and who was driving a vehicle that had a proper license plate, base sticker and/or visitor pass. In contrast, critical scenarios required the trainee to identify and address a particular issue or item that could lead to potential denial of entry. A description of critical scenarios can be found in Appendix A.

Equipment and Implementation

The CAVE. The VE interface used in this study was the CAVE (CAVE Automatic Virtual Environment), originally developed by the Electronic Visualization Laboratory (EVL) at the University of Illinois at Chicago using three-dimensional graphical software. The CAVE has four 10 x 10 ft walls and a floor. In the present study, the simulated environment was projected onto only the front and left side wall of the CAVE. Objects in the display were texture mapped to enhance depth and realism cues. In terms of scene complexity, the polygon count for this simulation was 13,172, with a resolution of 1024x768. The frame rate was 31 frames per second (FPS) with a refresh rate of 96 Hz.

Computing systems. The present study used 3 main computers connected through a 100 mbps networked switch. First, a Silicon Graphics (SGI) ONYX II computer equipped with three-dimensional graphic software was used to produce images. The SGI ONYX II was used to display the application in the CAVE, provide audio playback, and receive the information from the positional tracking devices. The SGI ONYX II was equipped with VrTool™ , TrackD™, Jack Toolkit™, Python™, Open Inventor™ and IRIX 6.5. Another Silicon Graphics computer was used for the experiment's main console to launch the application and employ override controls during the simulation. This second computer was equipped with IRIX 6.5, Motif, and buttonfly. The third computer system was a PC laptop used for voice recognition software and to communicate to the SGI ONYX II via a network socket. This laptop was equipped with Win2000, IBM ViaVoice™, and VrSpeech™.

Visual scene rendering. A terrain model depicting a U.S. Marine combat training town in Quantico, Virginia was adapted for use in this study. This scene graph was rendered using VrTool™ and the second scene graph (rendered internally by Jack) enabled dynamic character animation calculations. Existing structures were remodeled for improved real time performance and realism of textures. Jack Toolkit, a 3D human modeling tool, had the capacity to incorporate rotations and translations of 68 joints in the virtual human resulting in the ability to manipulate virtual body parts including the arm, hand and head.

Shutter glasses. The three-dimensional images were viewed stereoscopically with LCD CrystalEyes stereo shutter glasses manufactured by Stereo Graphics Corporation. The shutter glasses weighed 85g.

Voice recognition and audio. Participants used IBM's *ViaVoice*™ voice recognition software and a headset equipped with a microphone to issue commands. This software enabled trainees to engage in conversation with the Jack virtual human or agent. The *ViaVoice*™ software required two components, a grammar and a dictionary. The grammar supported substitutions and repetitions that were able to generate complicated sentences while simultaneously offering a wide selection of the possible commands issued by the trainee. The dictionary provided *ViaVoice*™ with pronunciations for each word that needed to be recognized. This dictionary was based on an accent from the east coast of the U.S. For most of the relevant words and phrases used in the experiment, it was necessary to say the word aloud, have the computer record it, and then have the software convert it into the respective 'baseform' representation.

There were two types of audio files used in the experiment, human voice and background noise. Background sound files were used to provide scene realism. These files included airplane flybys, wind, the approach of a car on a gravel surface, intermittent walkie-talkie chatter, and in one scenario, gun fire. Fifty-one minutes of constant (just audible) wind sound was used including 6 airplane flyovers and two instances of walkie-talkie chatter between Gate 2 and base command. The flyover sounds occurred at approximately 8-min. intervals and the walkie-talkie sounds occurred at 2 min. and 23 min. into the session, respectively. The gunfire sound file was used at a predetermined time in a specific scenario. All sounds were recorded in the natural environment (e.g., flyover) using a Sony Mini Disc portable recorder with an Audio-Technica ATR55 Telemike input. All sound files were edited using noise reduction algorithms and normalization techniques and were converted into the Audio Interchange File Format Version C (.AIFFC) file format.

The second type of audio file consisted of spoken responses from the virtual human agents. For neutral scenarios (see below), the virtual humans' voices consisted of four males and four females. Voices were randomly selected for use in neutral scenarios with some repetition of individual voices. Male and female voices were used to record phrases including: "Here's my ID, Here it is, I don't have a pass, I don't know, I don't understand, Yes sir, Yes ma'am, Okay, I have an NCIS ID, Okay I'll take care of it," and "No." These phrases were intended to serve as general responses to the repertoire of anticipated and unanticipated commands or questions issued by the trainee. For critical scenarios, additional sequences of dialogue were recorded specific to a given scenario. Dialogue for critical scenarios was provided by 12 male and 7 female voices. In some instances the same voice was used in more than one critical scenario. In cases where a voice was used more than once, the second occurrence took place in a different experimental session (i.e., on a different day of the experiment). Each voice was recorded with a Sennheiser headset microphone and was edited using Cool Edit Pro 1.2A and Sound Forge 5.0. All files were recorded in mono at a 22.1 kHz sampling rate into a single extended Windows Audio Version (.wav) file divided into individual sound samples. All files were manipulated such that the voices were presented at the same level of intensity and a noise reduction algorithm was employed to eliminate unwanted noise (e.g., hiss, clicks, or pops). Voice files were eventually converted into .AIFFC format for presentation in the CAVE environment.

Sound files were conveyed via two channels of a standard 4-channel soundboard. The participant faced forward in the CAVE with the left and right speakers placed at 225 and 315 degrees respectively. The speakers were set at 5 ft above the ground. Most speech and background sound files did not exceed 85 db.

Virtual humans / agents. Virtual human models were created using Jack Toolkit, a 3D modeling environment with support for high degree of freedom human models. Jack Toolkit was selected because it could convey a large range of dynamic motion and gestures in the head, eye, arm, hand, and leg joints. Jack virtual agents respond to questions, detect where a trainee is in the environment, and orient their head position toward the head of the trainee in a manner that conveys a face-to-face interaction between trainee and virtual human.

Positional tracking. Positional tracking was provided by Ascension Technology Corporation's Flock of Birds® software, a 6 degree of freedom (6DOF) electromagnetic tracker able to track one to four sensors simultaneously. Only one sensor was used in this study. This tracker operated at a 60 Hz sampling frequency and was attached to the *CrystalEyes* shutter glasses to provide translation and rotation information of the trainee in relation to the computer generated environment. The transmitter itself was located above the center wall of the CAVE and has an operating radius of 6 ft.

Video recording. A video camera was used to record participant's performance during scenarios. The camera, with the participant facing the front wall of the CAVE, captured a rear view of the participant interacting with each driver. A pair of LCD *CrystalEyes* shutter glasses were taped over the camera lens to record a single image from the stereoscopic display. The video recording allowed the playback of selected scenarios to provide feedback upon completion of the session.

Trainee equipment. Participants were equipped with a holster and a rubber replica of a training gun in addition to a radio transmitter to communicate to the base command center (experimenter).

Test instruments. Costa and McCrae's (1992) revised NEO personality inventory (NEO-PI-r) contains 240 self-report items that provide a comprehensive measure of the 'Big Five' major personality traits, including six facet scales for each of the five factors. The NEO has demonstrated respectable reliability and validity. Internal consistency ranges from .86 to .95 for domain scales and .56 to .90 for facet scales (Costa & McCrae, 1992). The NEO-PI-R has demonstrated strong convergent validity between self, peer, and spouse reports of the test. Moreover, the NEO-PI-r scales have correlated with similar scales from other well-established personality inventories, e.g., Myers Briggs Type Indicator (MBTI), Minnesota Multiphasic Personality Inventory (MMPI), and the

California Psychological Inventory (CPI). The three scales of the NEO PI-R used in the present study were openness, conscientiousness and neuroticism. Each scale ranged from 0 to 192 with 192 representing the highest possible score on each scale.

Participants were given a background questionnaire to establish previous military experience and demographics (see Appendix B). Additionally, each participant was also given the Simulator Sickness Questionnaire (SSQ) used to identify an individual's proneness to simulation sickness (Kennedy et al., 1993).

A postsession questionnaire asked participants for descriptive information about their opinions regarding the session, problems or errors they encountered, as well as a couple of 5-point scale items addressing the degree of stress and challenge associated with the task and opinions about how they performed (see Appendix C).

Experimental Design

Participants were randomly assigned to one of two conditions. In the first condition, participants performed a 40-50 min shift, received feedback on that shift, and then performed a second 40-50 min shift. Performance was compared between the two shifts. In the second condition, participants performed only a single session. The performance of these participants was compared to the performance obtained from the second session of the participants in Condition 1.

Procedure

Participants began the study by completing a survey of general background information and a personality test. They then reviewed a 3-page training manual outlining procedures to follow during their shift as a checkpoint guard (see Appendix D). Participants were then shown printed images of a base sticker and a parking pass referenced in the training manual. Following the review, each participant viewed a 7.5 min. video of a military briefing intended to provide background information on the rationale and importance of the military checkpoint as well as general instructions regarding performance as a checkpoint guard. The video identified their checkpoint location on a map, discussed the existing threat condition, the need to maintain vigilance and alertness, and demonstrated proper radio call procedures. Each participant went through a brief (4 to 7 minute) orientation to acclimate them to their role as one member of a two-person team. Their partner was a virtual human.

Following the review of the training manual and videotape, participants were taken to another room where they were fitted with their equipment for the task. They were provided with a belt and a holster containing a rubber Beretta model training weapon and a walkie-talkie. Participants removed watches, cell phones, and any jewelry that might conflict or distract the participant from the task. Next, the participants were given a log sheet depicting a time log of events that occurred on "the previous shift" and "Be on the Lookout" (BOL) information. The BOLs described events for which the participant was instructed to look for throughout the session. The log sheet (see Appendix E) also identified a dangerous person named Moe. A picture of Moe accompanied the log and participants were instructed to keep an eye out for Moe throughout their watch. Additional BOL events were presented aurally through speakers during the checkpoint task.

The participant then received a practice trial in which they were given a chance to familiarize themselves with a typical scenario and the equipment used to interact with the virtual people and objects in the simulation. The participants were given ample time to repeat this process until they became comfortable with the task. The participant then performed an experimental session lasting 40 to 50 minutes. In each scenario, the participant was asked to interact with each driver, ask for identification, and verify that the vehicle is fit for entry. The participant then rendered a decision as to whether the driver should turn around, pull over, or go ahead through the gate. The specific procedures for granting or denying access to the base are described in Appendix F. On average, each interaction took 1-2 minutes to execute.

Session 1. The first experimental session in Condition 1 contained 23 neutral scenarios and 12 critical scenarios. Upon completion of their session, participants were given feedback in the form of an after-action review (AAR). During the AAR, participants received information regarding the nature of their errors and the proper resolution of those errors. The videotape of the participant's performance was replayed in situations where it could help convey the sequence and steps leading up to an event. The participants were also given a second questionnaire to assess post experimental levels of stress and symptoms of simulator sickness. These participants returned 48 hrs later and performed a second shift.

Session 2. The second experimental session in Condition 1 contained 22 neutral scenarios and 13 critical scenarios. The difference in the number of critical scenarios between the first session (12 critical events) and the second session is accounted for by the appearance of 'Moe,' a person on the BOL list who only appears in Session 2. Prior to their second session, participants again filled out the SSQ and reviewed a log left by the previous shift (see Appendix E). The participants were given another chance to review the training materials and ask questions. At the second session, the log did not mention Moe explicitly and participants were shown the picture of Moe only if they asked to see it.

Group 2. In the second condition, participants performed only one session. Specifically, they performed the second session from Condition 1. Performance of this group served as a control to measure against the performance of those in Condition 1 who had received prior experience (on Session 1) and an AAR. The experimental procedures for this group replicated those used for the first session in Condition 1 except that these participants performed the scenarios from the second session of Condition 1.

Analyses

Performance was measured by the mean proportion of scenarios in which one or more errors were made. An alpha level of .05 was used for all statistical tests of significance and Tukey post hoc tests were used to analyze differences among the means.

Performance data were analyzed in several ways. The first analysis addressed Condition 1. Performance on the critical and neutral scenarios in Sessions 1 and 2 were compared with a 2 scenario (neutral and critical) x 2 session within-subjects ANOVA. A more detailed analysis was also performed on only the critical scenarios. Specifically, a 6 category (emergencies, social influence, memory/retention, perceptual, protocol, and situational awareness) by 2 type (identical and conceptual) x 2 session (1 *vs* 2) within-subjects ANOVA was used to assess transfer of training on those scenarios requiring the participant to take specific action.

The second analysis compared the performance of participants on the same scenarios. Specifically, performance on critical and neutral scenarios in Condition 1 and 2 were compared with a 2 x 2 mixed-factor ANOVA. Condition was analyzed as a between-subjects factor and scenario type was analyzed as a within-subjects factor. In addition, a more detailed analysis was once again performed on only the critical scenarios. Specifically, a 6 category (emergencies, social influence, memory/retention, perceptual, protocol, and situational awareness) by 2 type (identical and conceptual) x 2 condition mixed-factor ANOVA was used to assess transfer of training on those scenarios requiring the participant to take specific action. The category and type of critical scenario was analyzed as a within-subjects factors and condition was analyzed as the between-subjects factor.

Another analysis was performed to determine whether the effects of training observed for Group 1 were due to practice. Specifically, a two-tailed correlated *t*-test was performed on the error scores during the first half of session one in comparison to the second half of session one.

Additionally, analysis was performed to determine any pre-existing differences between samples. To this end, a two-tailed independent *t*-test was used compare Group 1's first session with the first and only session of Group 2.

Last, a stepwise multiple regression analysis was performed on personality scores on conscientiousness, neuroticism, and openness to determine which measures, by themselves or in combination, were effective predictors of checkpoint performance.

RESULTS

(See Appendix G for Error Classification used in analyses)

Military Background Data

A variety of background and qualitative items were collected through an initial background questionnaire (see Appendix B) and a postsession questionnaire (see Appendix C) that was administered following each session.

Military Experience. Only one participant reported previous military experience. A correlation analysis between military experience and performance was not conducted.

Familiarity with Military Checkpoints. Sixty-six percent of participants reported having passed through a military checkpoint at some time in their life (n = 14 and 7 for Groups 1 and 2 respectively). A two-tailed independent t-test on the combined first sessions of both groups also indicated no significant differences on the number of overall errors for the group who had passed through a checkpoint (M = .16, SD = .07) and the group who had not passed through a checkpoint (M = .20, SD = .15), $t(30) = .77$. These results suggest that familiarity with having passed through a checkpoint did not affect performance on the task.

Analysis of Performance Data

Results focused on two main types of analyses – analysis of performance on critical and neutral scenarios and, second, a more detailed analysis of critical scenario categories (e.g., situation awareness, emergencies). The overall analysis of critical and neutral scenarios was performed using a count of the total number of errors made on each scenario within each category (as defined in Appendix G). However, because some categories contained more scenarios than others (e.g., the situation awareness category had 1 scenario per session whereas the protocol category had 4 scenarios per session), a different measure was used in the analysis of scenario categories. Specifically, the analyses were carried out on the mean proportion of scenarios in which one or more errors were made. Analyses of the scenario categories were performed on *both* the total number of errors and the proportion of scenarios in which one or more errors were made. The results revealed no meaningful differences between the two measures. Consequently, only the analysis of the mean proportion of scenarios in which one or more errors were made will be reported. For simplicity sake, this measure will be referred to as proportion scores throughout the remainder of this paper.

Group 1: Critical and Neutral Scenarios

A 2 Scenario type (critical or neutral) x 2 Session within-subjects ANOVA was performed on the total number of errors. As hypothesized, a significant effect for scenario type indicated that participants committed significantly more errors in critical scenarios ($M = .34$, $SD = .26$) than in neutral scenarios ($M = .02$, $SD = .04$), $F(1, 15) = 60.15$. A significant session effect indicated that persons committed more errors on their first session ($M = .26$, $SD = .29$) than on their second session ($M = .10$, $SD = .17$). $F(1, 15) = 34.53$. Furthermore, as shown in Table 4, there was a significant interaction between session and scenario type. Post hoc analyses indicated that the number of errors on critical scenarios declined significantly from Session 1 to Session 2, $F(1, 15) = 27.78$. However, the decline for neutral scenarios was not significant.

Table 4.

Group 1: Mean Number of Errors for Critical and Neutral Scenarios

	Session 1	Session 2
Critical Scenarios	.50 (.22)	.18 (.20)
Neutral Scenarios	.03 (.03)	.02 (.05)

Note. Standard deviations appear in parentheses.

Group 1: Critical Scenario Categories

A 6 Category (emergencies, social status, memory, perceptual, protocol, situation awareness) x 2 Session within-subjects ANOVA performed on mean proportion scores once again, indicated a significant main effect for session, $F(1, 15) = 33.18$. Participants made significantly more errors on their first session, ($M = .50$, $SD = .43$), than on their second session ($M = .18$, $SD = .33$). A main effect for category was also found. Post hoc analyses indicated that more errors were made in scenarios requiring situation awareness, perception, and social influence tasks, than any other categories, $F(5, 27) = 4.93$. The interaction between session and category was also significant, $F(5, 75) = 2.34$. Errors in all of the categories decreased from Session 1 to Session 2, however, this decline was statistically significant for only the social influence and situation awareness scenarios. The mean proportion scores made per category for Group 1 in Sessions 1 and 2 are shown in the top portion of Table 5.

Groups 1 and 2: Critical and Neutral Scenarios

Data from Group 1's second session and Group 2's session were analyzed with a 2 Group x 2 Scenario Type mixed factorial ANOVA with group serving as the between factor and scenario type serving as the within factor. As hypothesized, Group 1

Table 5

Groups 1 and 2: Mean Proportion Scores for Category in Each Session

	Session 1	Session 2
GROUP 1		
Emergencies	.28 (.36)	.09 (.20)
Social Influence*	.56 (.51)	.13 (.34)
Memory	.38 (.34)	.13 (.22)
Perceptual	.50 (.52)	.31 (.48)
Protocol	.41 (.26)	.17 (.15)
Situation Awareness*	.88 (.34)	.25 (.45)
GROUP 2		
Emergencies		.21 (.31)
Social Influence		.56 (.51)
Memory		.41 (.42)
Perceptual		.69 (.48)
Protocol		.36 (.30)
Situation Awareness		.44 (.51)

* Significant differences between Session 1 and 2, $p < .05$.

performed significantly better ($M = .10$, $SD = .17$) than persons in Group 2 on the second session scenarios ($M = .25$, $SD = .31$), $F(1, 30) = 8.04$. A main effect for scenario type was found such that during Session 2, participants committed more errors in critical scenarios ($M = .46$, $SD = .31$) than in neutral scenarios ($M = .04$, SD = .10), $F(1, 30) = 49.43$. Furthermore, there was a significant interaction between Group and Scenario Type, $F(1, 30) = 9.59$. This interaction is shown in Table 6. A post hoc analysis indicated that Group 1 (who received two sessions and corrective feedback) made significantly fewer errors on critical scenarios than Group 2. The difference on neutral scenarios did not reach statistical significance.

Table 6

Groups 1 and 2: Total Number of Errors for Critical and Neutral Scenarios for Session 2

	Group 1, Session 2	Group 2, Session 2
Critical Scenarios	.18 (.20)	.46 (.31)
Neutral Scenarios	.02 (.05)	.04 (.10)

Note. Standard deviations appear in parentheses.

Groups 1 and 2: Critical Scenario Categories

Data from Group 1's second session and Group 2's session were compared with a 2 Group x 6 Category (emergencies, social status, memory, perceptual, protocol, situation awareness) mixed factorial ANOVA with group serving as a between-subjects variable was performed on the mean proportion scores. A main effect for group indicated that once again, Group 2 committed significantly more errors ($M = .45$, $SD = .45$) than participants in Group 1 ($M = .18$, $SD = .33$), $F\ (1, 150) = 8.56$. The results also showed a main effect for category, $F\ (5, 150) = 4.28$. Mean proportion scores per category for each group are shown on the right hand side of Table 5. Post hoc analyses indicated significantly more errors were made in the perceptual category than in the emergencies, protocol, and memory categories. Although Group 1 had better overall performance, no significant differences were found with respect to individual categories.

Groups 1 and 2: Initial Session

A two-tailed independent *t*-test was performed on overall critical error scores from Group 1's first session and Group 2 to determine whether the two groups performed comparably on their initial session. Due to the small number of errors on neutral scenarios, only critical scenarios were included in this analysis. Results indicated no significant differences between the overall errors made in Group 1's first session (*M* = .26, *SD* = .10) and Group 2's session (*M* = .25, *SD* = .18), *t*(30) = .25.

Another analysis was performed to determine whether the effects of training observed for Group 1 were due to practice. Using critical scenario data from the first session of Group 1, a two-tailed correlated *t*-test indicated that the number of errors for the first half of the session (*M* = .44, *SD* = .06) was not significantly different than the number of errors for the second half (*M* = .40, *SD* = .06), *t*(79) = .48, suggesting that no appreciable practice effects occurred within the session. A similar analysis was performed on the data from Group 2. Performance in the first half of their only session (*M* = .46, *SD* = .06) was not significantly different from that of their second half (*M* = .36, *SD* = .05), *t*(79) = 1.52.

Long-Term Retention

Data for the special scenario for long-term retention were analyzed separately. A two-tailed independent *t*-test indicated that Group 1 made fewer errors in detecting the target image (*M* = .25, *SD* = .45) than Group 2 (*M* = .63, *SD* = .50), *t*(30) = 2.24.

Personality Scores on Conscientiousness, Neuroticism and Openness

Personality scores on conscientiousness, neuroticism, and openness were analyzed using stepwise multiple regression. Results were analyzed at the .05 alpha level. For all stepwise regression analyses, variables were entered into the regression equation in order of largest correlation. Again, analyses were performed on the total number of errors as well as the analyses using mean proportion of scenarios in which one or more errors were made. A comparison of these two measures revealed no meaningful differences, thus, only the mean proportion scores are reported.

Group 1: Both Sessions. Results indicated no significant effect for any of the personality measures (neuroticism, openness, or conscientiousness) on overall critical errors, $F(1, 28) = .31$. Thus, subsequent analyses of individual personality traits as predictors of performance within specific categories were not conducted. A second analysis was also performed to determine if persons high in openness committed fewer errors in their second session than their first session. A median split was used to separate participants with lower scores on openness from those participants with higher scores on openness. A one-tailed independent *t*-test comparing persons with high and low openness scores on the total number of errors (neutral and critical combined) was employed. Results indicated no significant differences between the number of errors committed by persons low in openness ($M = 5.00$, $SD = 1.55$) in comparison to those high in openness ($M = 5.00$, $SD = 2.20$), $t\,(30) = .11$.

Groups 1 and 2: Initial Session. Results indicated no significant effect for any of the personality measures on overall critical errors, $F(1, 30) = 3.25$. Given that the overall model was not significant, subsequent analyses of personality as a predictor of performance within particular categories were abandoned.

Groups 1 and 2: Session 2. Stepwise regression analyses were employed to determine which personality traits by themselves or in combination would predict the number of errors in each category. The results indicated significant effects for the social influence and emergency categories (see Table 7). Specifically, openness had a negative relationship with the proportion of errors in emergency scenarios. The regression equation for openness on emergency scenario errors was significant, $R^2 = .14$, *Adjusted R* $= .11$, $F(1, 30) = 4.89$. No other personality measures entered the regression equation. These results suggest that people who exhibit less openness are more likely to commit errors in emergency scenarios. The results also showed a significant effect for social influence. The regression equation was significant, $R^2 = .24$, *Adjusted R* $= .21$, $F(1, 30) = 9.26$. Again, only openness was shown to have a significant and negative relationship with the proportion of errors made on social influence scenarios. These results suggest that people who exhibit less openness are more likely to commit errors in social influence scenarios. Personality traits were unable to predict performance in the other four categories.

Table 7

Regression of Openness on Error Scores for Each Scenario Category

Scenario Category	Parameter Estimate (*B* weight)	Standard Error	R^2	*p*
Emergencies*	-.007	.003	.14	.035
Social Influence*	-.015	.005	.24	.005
Memory	-	-	-	-
Perceptual	-	-	-	-
Protocol	-	-	-	-
Situation Awareness	-.008	.006	.07	.144

Note. Those categories depicted by an asterisk are significant, $p < .05$. Those fields represented by a dash (-) represent categories for which openness did not reach significance (.150) for entry into the model.

Neuroticism and False Alarms. A correlation analysis was conducted to determine the relationship between neuroticism and the number of false alarms. Analysis of the first session of each group individually and combined indicated no significant correlation between neuroticism and false alarms.

Self-Report Performance Ratings

Participants were asked to rate their own performance on a 5-point scale with a score of 1 indicating they felt they performed "very well" and a score of 5 indicating they felt they did performed "not very well at all." A one-tailed correlated t-test indicated that participants in Group 1 felt they performed significantly better during their second session ($M = 2.13$, $SD = .72$) than during their first session ($M = 3.25$, $SD = .68$), $t(15) = 4.14$. To compare the second session of both groups, a one-tailed independent t-test was employed. Results indicated no significant differences on self-report performance ratings for Group 1's second session ($M = 2.13$, $SD = .72$) in comparison to Group 2's session ($M = 2.63$, $SD = .81$), $t\,(30) = 1.85$. A two-tailed independent t-test used to compare performance ratings for the initial session of both groups indicated Group 2 felt significantly better ($M = 2.63$, $SD = .81$) about their performance than Group 1 felt about their first session ($M = 3.25$, $SD = .68$), $t(30) = 2.37$.

Additional correlation analyses were performed to determine whether self-perceptions of performance were correlated with actual performance. Self perceptions of performance were again measured on a 5-point scale with lower scores indicating better performance. Actual performance was determined by the mean overall errors for critical and neutral scenarios combined. Correlation analyses were performed on data compiled from both sessions of Group 1 and on each session separately. The analysis revealed a significant negative correlation between actual and self-reported performance for Group 1's first session, $r = -.54$. That is, as errors decreased, persons indicated they felt they performed better. However, the same correlational analysis performed on data from the second session failed to reach significance. A similar correlation analysis performed on the data from Group 2 also failed to reach significance. Further, data from the combined initial sessions of both groups indicated no significant correlation between self-reports and performance.

Self-Report Stress Ratings. Participants were asked to indicate the extent to which they found the task stressful or challenging on a 5-point scale with 1 being the highest stress rating and 5 being the lowest stress rating. Analyses were performed to compare stress levels within Group 1, as well as between Group 1 and Group 2. A two-tailed paired *t*-test indicated that participants found their second session to be significantly less stressful and challenging ($M = 3.84$, $SD = .20$) than their first session ($M = 1.5$, $SD = .09$), $t(31) = 11.79$. In comparing the stress level of the initial sessions of both groups, a two-tailed independent t-test indicated no significant differences between the stress level reported for the first session of Group 1 and Group 2's session. A two-tailed independent t-test was also performed to compare Session 2 stress ratings for both groups and revealed no significant differences in the level of stress reported for Group 1's second session ($M = 4.06$, $SD = .85$) in comparison to Group 2's session ($M = 3.38$, $SD = 1.31$), $t(30) = 1.76$.

Preparedness to Stand Guard at Checkpoint. Although a number of participants indicated that more training would be necessary to stand guard at a real military base, 97% indicated that the training had, at least in part, better prepared them if they were to stand guard at a base. It should be noted that the one individual who did not indicate a positive effect was in Group 2 who had only experienced one session. See Appendix H for a list of related participant comments.

DISCUSSION

The focus of the present study was to evaluate a VE as a training tool for military guards and, second, to examine personality traits as predictors of performance on the task. First, these results will be discussed in terms of fulfilling the goal of using VE technology to aid in training transfer. Second, the advantages and drawbacks of the technology will be discussed in terms of their application to future developments. Finally, personality traits will be discussed as potential predictors of performance on the checkpoint task.

Training Transfer

The primary goal of the present study was to use VE technology to facilitate training transfer for military checkpoint duty. Training transfer was defined as applying knowledge acquired in one training session to a subsequent session. This transfer is based on the notion that the greater the similarity among tasks, the greater the degree of training transfer. Holding (1965) stated that if two tasks had identical stimuli, then training transfer would be maximized. Similarly, Thorndike and Woodworth's (1903) identical elements theory proposed that environments having common elements were more apt to facilitate training transfer. Transfer-through-principles theory suggests that a participant could use a previously acquired principle and apply it to a new setting without necessarily being consciously aware of the similarity between tasks and environments. Accordingly, in the present study scenarios that were identical or conceptually similar were used in both sessions. It was expected that participants would learn from their experience and feedback in the first session and commit fewer errors in a subsequent session.

Critical and Neutral Scenarios. Participants appeared able to distinguish between critical scenarios requiring special actions and the neutral scenarios. The number of errors on critical scenarios was significantly higher than on neutrals. More important, however, the number of errors made on critical scenarios decreased from Session 1 to Session 2 as expected. This result indicates that participants learned from their experience and AAR from the first session and improved upon their performance in their second session. As hypothesized, Group 1's first session and related feedback resulted in reduced errors in their second session indicating training transfer. Persons who went through two sessions committed nearly 60% fewer errors on their second session than their first session. By contrast, participants in Group 2, who did not have the opportunity to learn from a first session, made over twice as many errors on scenarios identical to those scenarios in Group 1, thereby suggesting that Group 1's performance benefit was not likely due to the specific selection of scenarios in each session. Further, scores from Group 1's first session were not significantly different from those of Group 2 in their first and only session. This finding indicates that both groups performed comparably in their initial session.

The possibility that Group 1's improvement from Session 1 to Session 2 was merely a result of increased familiarity and practice was addressed by comparing their performance on scenarios from the first and second halves of Session 1. The results showed that performance did not differ significantly between the first and second half. Collectively, the findings indicate that participants were able to use VE technology to (a) familiarize and immerse themselves in a military checkpoint task, (b) learn via the scenarios, (c) perform the task successfully, and (d) transfer their knowledge from Session 1 to novel but similar scenarios in Session 2.

Categories of Error. With regard to categories of error, scenarios were designed to evaluate skills in six categories -- emergencies, social influence, memory, perceptual skills, ability to follow protocol, and situational awareness. It was expected that if persons learned from their errors in Session 1, then their errors would decline in each of these respective categories. As indicated in Table 6, this expectation was partially supported. Because the two sessions were designed to be relatively similar, it was hypothesized that knowledge would generalize from one session to the next. Consistent with the transfer-through-principles theory, errors in all six categories clearly decreased from Session 1 to Session 2; however, this decrease was significant only for errors made in social influence and situation awareness categories. The AAR most likely contributed to this decrease in errors. Participants who received feedback on their errors and performed a second session made fewer errors on that session than participants performing the session without feedback from session one. This implies that the AAR was an important element of training.

Some scenarios and types of errors were more easily addressed in the AAR and thus, may have contributed to the decrease in errors. For example, in the case of errors on social influence scenarios, it was necessary to instruct participants that the decision for entry should be based on specific criteria such as identification cards and license plates and should not be based on status (e.g., city mayor) or situation (emergency). Participants were told it did not matter *who* the driver was or if they were in a hurry, both the driver and car had to meet proper requirements. In contrast, errors on memory or perceptual scenarios were less easily addressed in the AAR. This may be why there was less improvement on memory and perceptual scenarios than other categories such as emergencies or situation awareness. For example, showing a replay of a car with a BOL license plate does not necessarily help a trainee's level of alertness or aid in his or her ability to remember a particular sequence of letters and numbers to check against the license plate of each vehicle. Further, in terms of procedural tasks, there were four scenarios per session and some of these scenarios benefited less from the AAR. For example, the playback of the rifle in the backseat of a car in one scenario was not easy to see in the video.

The poorest performance (and also the best improvement) was obtained on the situation awareness and social influence categories. There was one situation awareness scenario per session and each of the scenarios was fundamentally different as to make performance comparisons difficult. In the first scenario, two cars pulled up and the driver in the first car informed the participant that his friend was in the car behind him and should be allowed entry. The participant was penalized if s/he allowed entry of the friend without checking for proper requirements. In the second scenario, again, two vehicles approached but the driver in the second vehicle, if not caught in time drew a gun. In the case of social influence, scenarios were more comparable. Each session contained one scenario in which a person of social status (i.e., either the mayor or the admiral's son) pleaded to enter the gate without proper requirements. Further, it was easy to address the scenario during the AAR by instructing participants to disregard social status and focus solely on gate entry requirements. The ease of addressing this scenario during the AAR most likely led to reduced errors across sessions.

Long-Term Retention. Participants were asked to look for a driver matching the picture of a man named Moe. These data were analyzed separately from the six main categories. Group 1 was shown a picture of Moe immediately before their first session. They were not shown this picture prior to starting their second session unless they specifically asked for it. Group 2 was shown the picture immediately before their only session. The scenario containing Moe appeared in the 14th scenario of the second session (about halfway through the session). Recency effects would suggest better performance for those individuals who saw Moe's picture immediately before the session in which he appeared. However, this was not the case. Results indicated that participants in Group 1, who saw Moe 48 hours earlier, were able to identify Moe more often than participants in Group 2 who had been introduced to Moe only 45 minutes beforehand. In fact, Group 1 identified Moe 75% of the time in comparison to Group 2 in which only 37% of the participants identified Moe.

There are a number of possible explanations for this finding. First, it is possible that the participants in Group 2 were still getting accustomed to the task and may have not had spare attention capacity to spend on focusing on looking for a particular face at the point at which Moe appeared. Second, it may be that participants in Group 1 had more time to reflect on their session and thought about why Moe had not appeared during their initial session. Support for this idea is evident by the fact that although not prompted, some participants remembered Moe between sessions. That is, prior to Group 1 starting their second session, nearly 25% of the participants asked to view Moe's picture again. Coupled with the fact that Group 1 participants were already familiar with the basic components of the task and could allocate mental resources to additional goals, participants in Group 1 may have been more focused on detecting Moe.

The present study focused on the potential applicability of CAVE technology for training, and represents the first in a set of studies. The second study in the series was a near replication using the same training scenarios but presenting them on a desktop system. The results from that study showed that the patterns of errors made on sessions and on the individual scenarios were similar to those obtained in the present study. Participants made more errors on critical then neutral scenarios. Also, participants who performed two sessions made fewer errors on their second session. However, a comparison of the total number of errors made in each study showed that overall, participants made more errors on the desktop system than with the CAVE. This difference between the two studies suggests that the CAVE experience may provide a more effective means of training than a desktop system for this particular task. Further follow-up studies would be necessary to isolate the CAVE's contribution to training transfer in comparison to other factors such as the training manual and AAR.

The CAVE Experience. Qualitative and performance data indicated that participants responded well to the CAVE environment (see Appendices H and I for qualitative responses). The participants in the present study were generally not familiar with VE technology, with some indicating stressful reactions to technology in general. However, any initial hesitance or anxiety with the VE technology appeared to be overcome by the end of the session. Asked whether the VE made a difference in performing the task, the general response was enthusiastic. One participant wrote:

"Absolutely, I was really into what was going on"! Another participant wrote: "you could really test what you know, the best training is on the job and this is as close as you can get."

Although participants reported some technological inconveniences (e.g., having to wear heavy goggles), as a whole, participants acclimated well to the environment, became accustomed to the mechanisms of interaction, and interacted with virtual objects rather naturally. A sense of immersion was informally observed during the experiment. Some participants reported they felt like they were embedded in the scenario and, in fact, some participants were observed physically motioning for cars to pull up to the gate and others reached out to try and hold the ID card the driver was displaying. Participants made statements including:

(a) "It made it seem more real. I felt like I was actually doing the tasks."

(b) "I felt like I was actually involved."

(c) "I think because it was more realistic it added stress to the job/task."

(d) "It feels a lot more interactive, so it grabs your attention and makes you want to do well."

(e) "I felt like I was actually on the job."

(f) "This is more alive."

In sum, participants felt like they were immersed in the role of a checkpoint guard. In one participant's opinion, the experience "provided a greater awareness of the responsibilities of guards" an experience s/he suggests is different from the desktop environment in which s/he hypothesized feeling "completely outside the environment."

Although many participants reported positive experiences and a sense of immersion, there were still numerous technological issues that detracted from the experience.

System Issues and Limitations. The natural language interface proved to be the biggest source of problems in the experiment. Two issues were evident. First, the voice recognition software used in the present study was not very reliable. Differences in various voice tones, inflections, and accents among the participants all affected recognition success. Participants often had to repeat commands. Unfortunately, this required the experimenter to spend additional time training participants to speak clearly and enunciate with the proper inflection to improve recognition success. Second, it was apparent that participants, especially when well immersed in the environment became more conversational and as expected, used a wider vocabulary than was preprogrammed. The additional words and phrases were interpreted by the software erroneously and sometimes caused commands to be carried out unintentionally. In particular, erroneous interpretation of extraneous conversation resulted in drivers running through the checkpoint before the participant could render a decision. Consequently, it was often necessary for the experimenter to intervene and manually implement the appropriate driver responses and actions.

A significant number of persons indicated they were consciously aware that they were dealing with software and worded commands accordingly. A few participants reported that they did not handle important issues because they were aware of technological limitations. For example, some participants indicated a desire to issue a particular command but were unaware of how to word it so that the virtual human would recognize it. Consequently, the participant refrained from verbalizing commands when s/he felt unable to successfully communicate with the virtual human. Still, the participant was able to issue commands directed at the main problem but was occasionally unable to issue commands relating to finer details of a problem. For example, the participant may have noticed an object such as a box but was unable to ascertain what was in the box or what the box was for.

Human behavior and human language are quite complex. In any given situation, it was not possible to foresee all the many combinations of comments and levels of inquiry that a participant might use. Knowing this ahead of time, several default phrases such as "I don't know," "I don't understand," and "I'll take care of that" were prerecorded. However, when the driver was asked such questions as "What's in the box?," the resulting "I don't know" response was construed as suspicious behavior leading to an inappropriate decision on that particular scenario. In other words, for those participants who did ask questions and issued unanticipated commands, occasionally the limited response set did not suffice.

Another source of problems concerned the fidelity of the virtual object models. For example, the back trunk area of the jeep model contained a square-shaped wheel well that was occasionally perceived as a package or box in the trunk. Some participants perceived the top layer of the front seat as a flat package and asked questions about it. Another participant misconstrued a rifle as a fishing pole. Participants who made decisions based on these misperceptions were not penalized. For example, in at least 8 separate sessions (19% of the total sessions), participants construed wheel wells as boxes or seat cushions for flat parcels. These instances were not coded as errors if the participant followed through according to proper protocol (i.e., contacting base on the radio and detaining the vehicle).

The car models were not the only challenges with respect to object models. The Jack virtual human used in the study was originally developed to model anthropometric/ergonomic requirements and as such, it was not designed for many of the subtleties of human expression or behavior. The virtual human was modified to open its mouth during dialogue but the mouth movement did not coincide with the speech and detracted from the realism of the interaction.

Some participants also reported that the virtual humans appeared suspicious because their eyes seemed a bit unnatural or glazed. Adding to suspicious behavior, the female models had a rather masculine outward appearance leading to some confusion in scenarios where the female dialogue was integrated. Participants indicated that they ultimately attributed this issue to technology rather than their initial notions that the behavior was suspicious.

Other software issues occurred as a function of shared computing resources leading to slow cars or cars that were not displayed. Thus, it became necessary to reload certain scenarios manually. This occurred rarely and was usually caught before the participant noticed the problem.

The physical equipment was also inconvenient for some participants. This was especially true for participants who had smaller ears and who could not easily support the weight of the goggles. As shown in Appendix I, this particular problem was reported on 7 separate sessions or on 14.6% of the total sessions.

Another significant issue was the projected image. On at least 12 separate sessions (25% of the total sessions), participants indicated that the image was not being displayed properly on the walls of the CAVE. On 6 separate sessions (12.5% of the total sessions), participants described the display as blurry or indicated that the image appeared "jumpy," "blinky," or "flashing."

Technological issues aside, participants reported that their VE sessions contributed to their knowledge of a military checkpoint task. Some indicated that they would feel comfortable assuming the role of a gate guard if they were asked to do so (see Appendix G). Others indicated they were more comfortable but would require additional training. Generally speaking, it is evident that the technology served to familiarize participants with the overall task, immerse them in the environment, acclimate them to various scenarios that are typical in the day-to-day life of a gate guard, and provide them with a training tool in which they could learn from their results and reinforce them by applying them to subsequent training sessions, and ultimately the real world.

Personality

The second major objective of the study was to determine whether several personality characteristics would predict performance on this task. Specifically, openness, conscientiousness, and neuroticism were identified as the most promising predictors of task performance. Each one of these traits will be discussed in turn.

Openness. The first trait hypothesized to impact performance was openness. Openness, often said to reflect intelligence or cognitive ability (Barrick & Mount, 1991, Digman & Inouye, 1986) was hypothesized to be a strong predictor of overall performance. Persons high in openness are creative, curious, flexible, have an intellectual orientation, and have a positive attitude toward learning that contributes to their success across a diverse set of occupations. In this study, it was hypothesized that these attributes would manifest themselves in a reduced number of overall errors. Results indicated that persons high in openness committed fewer errors than persons low in openness, however, this difference was not significant. It was further hypothesized that the flexible, creative nature of persons with high openness would help them to not only look for what belongs in a scenario, but also to see what is missing in a scenario (e.g., a license plate). The hypothesis that persons high in openness would do better on perceptual tasks than persons low in openness, was not supported. Because there was only one perceptual scenario per session, it is possible that there were not enough instances for missing items to stand out in contrast to the presence of inappropriate features (e.g., wrong identification card). That is, after performing this task on a daily basis for months, a participant may be more accustomed to searching for the presence of absent features as part of their routine.

Openness was found to have a significant negative correlation with the number of errors in social influence scenarios. Openness is reflected in persons who are independent thinkers as opposed to those who conform. In social influence scenarios, the participant is pressured to give into the demands of a person with noteworthy social status and allow the driver to enter without proper requirements. The participant must be able to resist the pressure of social influence and follow protocol. A person who is high in openness would be less conforming and consequently it is conceivable that openness has a positive correlation with performance (fewer errors) as supported by the results of the present study. Similarly, openness also had a positive influence on performance in emergency scenarios. Recall that emergency scenarios involved an ambulance or injured occupants in a vehicle and the participant was pressured to allow the driver to enter the gate regardless of proper requirements. Thus, the emergency scenarios were similar to the social influence scenarios in that they both required the participant to consider deviating from protocol and may have tapped into the same openness characteristics. The hypothesis that persons high in openness would commit fewer errors than persons low in openness failed to reach significance.

Conscientiousness. Conscientiousness is closely related to the construct of volition. A person high in conscientiousness is characterized as hardworking, well organized, persistent, careful, motivated, responsible, and dependable. As such, it was hypothesized that persons high in conscientiousness would have fewer errors than persons low in conscientiousness. Further, it was hypothesized that careful, organized persons are more likely to adhere to instructions and pay attention to details, comply with procedure, and consequently be less vulnerable to errors in social influence scenarios, however, this effect was not observed. Conscientiousness has been reported as one of the most consistent predictors of performance across all occupational groups (Barrick & Mount, 1991), thus, it is surprising that it was not correlated with performance in the present study.

There are a number of possible explanations for why the hypotheses for conscientiousness were not supported. First, Mount and Barrick (1995) suggest that conscientiousness appears to predict particular aspects of performance in contrast to overall performance success. Further, Mount and Barrick also stated that conscientiousness is most strongly related to those criteria that are determined by motivational effort or *will do* elements (ρ ranges from .42 to .45) in contrast to ability or *can do* elements (ρ ranges from .25 to .26). This is an important qualification because it suggests the effects of conscientiousness would be difficult to observe if participant recruitment, by it's very nature, attracted those participants who were already of a will do inclination. This may be the case in the present study considering students received extra credit or financial compensation in exchange for participation, and, therefore, scores were representative of participants who made the effort to gain course credit as opposed to those who did not participate for reasons including irresponsibility or lack of motivation. In other words, if conscientiousness predicts motivational measures (*will do*) moreso than ability or *can do* measures, the sample in the present study may not have adequately represented the continuum of human motivation. It is then possible that conscientiousness would not be an effective discriminator between poor and good performers. Means and standard deviations for all three personality traits measured on the current sample and compared with those of the general population are shown in Table 8. As seen in the table, the mean score for conscientiousness in the present sample (114.3) is notably higher than the mean score for conscientiousness in the general college population (44.1). Consequently, it could be argued that the present study did not have a representative sample across the range of conscientiousness.

Another explanation for why conscientiousness did not relate to performance in this study may lie with Barrick and Mount's (1993) research concerning the role of autonomy as a moderator of personality on performance. Barrick and Mount's results revealed that high conscientiousness scores are correlated with better performance in jobs with high autonomy. Autonomy is defined as the extent to which a task provides

Table 8.

Sample Personality Scores and Norms for the General Population

	The Present Study	General College Population*	General Adult Population*
Neuroticism	84.1 (23.2)	86.1 (21.1)	67.2 (20.3)
Conscientiousness	114.3 (19.9)	44.1 (8.8)	53.1 (9.2)
Openness	123.3 (14.9)	121.9 (19.9)	105.6 (18.3)

*Note. Norms for the general population are based on male responses for the NEO-PI (Costa & McCrae, 1989).

substantial freedom, independence, and discretion in determining the procedures used in carrying out a task (Hackman & Oldham, 1980). Autonomy is a function of the number of behaviors able to be expressed in a situation. The checkpoint task can be characterized as a job with relatively low autonomy because much of the procedure is dictated by training protocol. Participants in the present study had a fairly limited set of behaviors from which to choose. Scenarios presented the participant with variability in situational factors, however, this variability was still low relative to other types of jobs (e.g., managerial jobs). For example, the participant encountered different critical scenarios where drivers did not have the proper requirements to enter a base, however, the procedure clearly dictated that drivers who do not have the specific set of required items are to be denied entry onto the base. Because conscientiousness has a greater effect on performance in high autonomy jobs, the relationship between conscientiousness and performance may not have been evident in this low autonomy checkpoint task.

Neuroticism. Neuroticism was also hypothesized to be a predictor of performance. Persons high in neuroticism generally have anxiety and low stress tolerance. It was hypothesized that neuroticism would have a negative relationship with performance and that the impulsive nature of the neurotic person would lead to significantly more false alarms. However, neither of these hypotheses were supported.

There are a few explanations for why neuroticism was unrelated to performance in the present study. As previously stated, neuroticism has a potentially complicated relationship with performance. For example, the impulsivity of a neurotic person could aid in identifying threats, but may also cause anxiety that interferes with the ability to recall and follow procedure. The complicated relationship between neuroticism and performance may also be due, in part, to potential moderating variables (e.g., level of conscientiousness), however, future research would need to determine such a relationship. Persons high in neuroticism may perform well if there are other traits that compensate for its negative effects. A person high in neuroticism and intellectance (openness) may perform more adequately than a person who does not have a high level of intellectance to compensate for high levels of neuroticism. These types of interactions are hard to investigate in a study with a small sample size. Last, it is necessary to explore which jobs would distinguish among individuals with different levels of neuroticism. Barrick and Mount (1991) found that neuroticism predicts job performance in some occupations (e.g., police and skilled or semi-skilled occupations) but not others. It is possible that the checkpoint task is one in which neuroticism does not have much impact or is masked by other factors. However, other limitations of the study must be resolved before drawing such a conclusion.

Summary of personality measures. Although the hypothesis regarding openness and performance was supported, overall, personality had little relationship with performance. One possibility may be that personality has a more discernable relationship to performance in some jobs and not others. Terborg (1977) suggested that the utility of a personality trait in predicting performance on a task is a function of task complexity. More complex tasks allow greater opportunity for personality to affect performance. Personality effects in the present study may have been attenuated because the checkpoint task was not very complex. Similarly, Barrick and Mount (1993) examined whether job autonomy and situational strength moderate the relationship between personality and performance in managerial jobs. A weak situation is one in which there is room for a participant to exercise judgment regarding which behaviors to use. Weak situations offer a larger set of options for how to go about achieving a goal. Strong situations have procedures, sequences, and methods that are laid out in specific detail and offer less opportunity for the participant to exhibit individualized strategies and behaviors. There are fewer options for accomplishing the goal. For example, an assembly line worker operating under close supervision does not have as many opportunities to express individual personality in his or her work as a CEO (Barrick & Mount, 1993). The military checkpoint task used in the present study is a structured task. There are proper procedures and sequences of action that are performed on a repetitive basis through training (i.e., strong situation strength). Thus, this situation could mask the effects of personality.

A second explanation may lie with the nature of the assessment instrument. Personality assessment tools were not originally created for the purpose of predicting performance on military tasks and may require further adaptation in order to be more successfully used as selection tools. Personality inventories were originally developed for clinical purposes and as such, they may be more applicable for clinical purposes rather than as performance indicators. It is not to say that all traits are not operating and influencing behavior, however, it appears that the existing personality inventories may not be suited for this purpose (e.g., Rosse et al., 1991). Moreover, these tools may have only limited utility in highly structured job contexts. Further, Hollenbeck and Whitener (1988) argued in their literature review that personality is only valid when used in combination with ability tests. Together, these issues may explain why the personality measures used in this study were weak predictors of performance.

Further, personality is a construct that is difficult to assess objectively or to quantify. Instead, researchers often rely on self-report measures, such as the NEO. Unfortunately, self-ratings may underestimate the importance of personality as performance predictors. Mount, Barrick and Strauss' (1994) results indicate that relying on self-report personality inventories alone will result in an understatement of the validity of personality constructs in performance prediction.

Additionally, the inability to detect strong personality effects in the present study may be a function of other known and unknown moderating factors that were not assessed. Day and Silverman (1989) note numerous factors that could cause the relationship between personality traits and job performance to be underestimated. One example is the failure to account for differences in role requirements between occupations. The nature of the military guard's occupation is very different than that of a speech therapist. It is conceivable that the traits that are important in task performance for a military guard are different than the set of traits critical to performance as a speech therapist. Day and Silverman (1989) suggest that different sets of variables may be applicable to different types of occupations. It is possible that a different set of traits may have been more effective as predictors of military task performance. It may also be necessary to distinguish between training performance and on the job performance because some personality traits may have more utility in training contexts than on the actual job. For example, openness is a strong predictor of training proficiency (Barrick & Mount, 1991) and was related to performance in the present study.

The results of this study should not cast doubt on the theoretical relationship of personality and performance. As with any construct, it is challenging to develop quantitative tools to measure traits. Personality inventories have been used for performance selection as well as clinical assessment, two categories that reflect very different purposes. Hollenbeck and Whitener (1988) caution against abandoning personality measures as performance predictors, implying that such a relationship between personality and job performance may exist hidden beneath theoretical shortfalls and methodological artifacts that contribute to underestimates of relationships between performance and personality traits. These factors include lack of statistical power (as in the present study), contamination of measures due to reliance on self-report/observational methods that are prone to artifacts (e.g., social desirability) as opposed to perception/judgment methods that obtain less contaminated indices. Digman (1989, 1990) suggests that the use of aggregation techniques may strengthen relationships of personality correlation coefficients from .30 (the conventionally accepted upper limit to be expected of personality measures) to higher more impressive levels (for a discussion of aggregation, refer to Digman, 1990).

Consistent with Hunter and Hunter (1984), the theoretical foundation for interactive effects of ability and personality in predicting performance led Hollenbeck and Whitener (1988) to suggest using ability tests in conjunction with personality measures. Doing so provides a more positive picture of the utility of personality as performance predictors and selection aids (Hollenbeck & Whitener, 1988; Rosse, Miller, & Barnes, 1991; Sackett, Gruys, & Ellingson, 1998). For instance, as noted above, openness may be measuring a person's ability to learn in addition to motivation (Barrick & Mount, 1991). Few studies have looked at personality in addition to ability, however, as Rosse et al. (1991) note, studies that have incorporated both measures found that personality is a valuable contributor to performance prediction and may be related to certain criterion measures of performance more than others. For example, McHenry et al. (1990) found that certain personality measures were better predictors of performance than the standard ASVAB tool used in military placement decisions. Such research suggests that the consideration of ability in addition to personality has support in the literature and deserves further attention (Rosse et al., 1991).

It should be noted that although this study offers promising results in terms of VEs and their potential as training tools for military checkpoint tasks, there were still some problems with the present study that impacted the results. First, the sample size was small and was a limiting factor regarding the effects of personality. Using personality traits as predictors often requires more power than was available in the present study. Time limitations did not allow the use of ability tests in conjunction with personality tests. Practical limitations did not allow participants to perform in a real life checkpoint scenario in order to test transfer to the actual job. The small sample size did not afford the opportunity to create additional experimental groups that could help strengthen the finding that performance benefit was not merely a function of practice.

Motivation may have also been constrained. In this study, participants were recruited who were interested in extra credit for a course. Any participant who showed up for this study voluntarily could be assumed to have similar levels of motivation and thus, it would be consistent across all participants and its impact would not be readily apparent in the results.

As Day and Silverman (1989) stated, personality variables may be significant predictors of job performance when matched with the appropriate occupation and organizational characteristics. Furthermore, the influence of personality variables as predictors may be better supported with the progression of tools once designed for clinical purposes that are altered for nonclinical performance purposes. As Hollenbeck and Whitener (1988) state "it would be unfortunate if the study of personality, so long a focus of attention for psychologists, was prematurely and permanently abandoned by personnel psychologists" (p. 89). Conceivably, personality measures, when matched with relevant scales in particular occupations and organizational characteristics hold an important role in selection research (Day & Silverman, 1989). Rosse et al. (1991) concur and suggest that personality traits may be more predictive of certain job categories than others. Future research would be best focused on identifying which personality traits should be used for which situations, and for what purpose (Guion & Gottier, 1965 as cited by Day & Silverman, 1989). Doing so would ensure a better selection tool for military guards, resulting in reduced training costs, reduced attrition, and increased safety of both the military checkpoint personnel as well as identification of threat risk to the public. It is imperative that we be able to answer the questions of 'Who can complete this task?,' 'Who will perform well at this task?' as well as differentiate between a person who may perform better during training (training proficiency) versus performance on the job which may draw on different motivation factors. Complications aside, personality measures deserve greater consideration in personnel selection (Rosse et al., 1991). In sum, it is conceivable that personality plays a role in performance, however, to identify the specific relationship(s) personality traits have in certain performance domains, it will be necessary to identify personality tools appropriate for such assessments, and to identify and control for moderating effects found in the environment and task.

CONCLUSIONS

Implications and Applications

The primary focus of this experiment was to evaluate the VE as a training tool for a military task. The use of VEs is beneficial for numerous reasons, among them, to provide a standardized mechanism for training large numbers of personnel with minimal space, increased safety, and within a reasonable budget. One noteworthy reason to train military guards in a VE is that doing so allows them to experience events that, although hard to train in the real world, are still real threats. As noted earlier, the current training of military guards is accomplished with on-the-job training. Because most of the time activities at a military checkpoint are mundane and critical events are few and far between, the nature of the environment poses a challenge for using on-the-job training as the primary means to prepare for rare and real threats. The present study shows promise for the use of VEs for training. Moreover, it is not unreasonable to conclude that such an application might generalize to other important tasks and similar occupations. For example, some of the categories of skills necessary in a military guard are also found in other occupations including airport security, police officers, and security guards at government buildings. Virtual environments are powerful training tools and will become even more useful as voice recognition software becomes more sophisticated and as future research improve upon the modeling of human behavior subtleties.

Riva (1997) once stated that VR is a "solution looking for a problem." Arguably, if a cost-efficient, safer means of military checkpoint training is the challenge, CAVE technology offers elements that may constitute the solution. In this case, VE provides a solution that could ultimately train personnel on one of the most important and perilous tasks at hand by providing the skills necessary to preserve and protect freedom, property and human life.

REFERENCES

Alluisi, E. (1991). The development of technology for collective training: SIMNET, a case history. *Human Factors, 33*, 343-362.

Anderson, C. A., & Bushman, B. J. (2001). Effects of violent video games on aggressive behavior, aggressive cognition, aggressive affect, physiological arousal, and prosocial behavior: A meta-analytic review of the Scientific Literature. *Psychological Science, 12*, 353-359.

Andersson, R. L. (1993). A real experiment in virtual environments: A virtual batting cage. *Presence, 2*, 16-33.

Arthur, W., Bennett, W., Stanush, P. L., & McNelly, T. L. (1998). Factors that influence skill decay and retention: A quantitative review and analysis. *Human Performance, 11,* 79-86.

Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, 44,* 1-26.

Barrick, M. R., & Mount, M. K. (1993). Autonomy as a moderator of the relationships between the big five personality dimensions and job performance. *Journal of Applied Psychology, 78,* 111-118.

Barrick, M. R., Mount, M. K., & Judge, T.A. (1999, April). *The FFM personality dimensions and job performance. A meta-analysis of meta-analyses.* Paper presented at the annual conference of the Society for Industrial and Organizational Psychology, Atlanta, GA.

Battles get close-up: Troops fight Iraqi guards about 50 miles from Baghdad. (2003, April 1). *The Virginian-Pilot,* p. A1.

Biersner, R.J, & LaRocco, J.M. (1983). Personality characteristics of U.S. Navy divers. *Journal of Occupational Psychology*, *56*, 329.

Blade, R. A., & Padgett, M. L. (2002). Virtual environments standards and terminology. In K. M. Stanney (Ed.), *Handbook of virtual environments: Design, implementation, and applications* (pp. 15-27). Mahwah, NJ: Lawrence Erlbaum Associates.

Cable News Network (CNN). (2003, June 6). *Gunmen kill U.S. soldier at Iraq checkpoint.* Retrieved July 7, 2003 from, http://www.cnn.com/2003/WORLD/meast/06/09/sprj.irq.soldier.killed/

Caird, J. K. (1994). Issues in the application of virtual environment systems to military training and selection (Tech. Note 9/94). Canada: Canadian Forces Personnel Applied Research Unit.

Costa, P. T., & McCrae, R. R. (1989). *The NEO PI/FFI manual supplement.* Odessa, FL: Psychological Assessment Resources.

Costa, P. T., & McCrae, R. R. (1992). *NEO-PI professional manual.* Odessa, FL: Psychological Assessment Resources.

Costa, P. T., McCrae, R. R., & Dye, D. A. (1991). Facet scales for agreeableness and conscientiousness: A revision of the NEO Personality Inventory. *Personality and Individual Differences, 12,* 887-898.

Cruz-Neira, C., Sandin, D. J., DeFanti, T. A., Kenyon, R., & Hart, J. C. (1992). The CAVE audio visual experience automatic virtual environment. *Communications of the ACM, 35*(6), 64-72.

Cutler, T. J., & Herdt, J. (1999). *The bluejacket's manual* (22nd ed.). Department of the Navy: United States Naval Institute Press.

Day, D. V., & Silverman, S. B. (1989). Personality and job performance: Evidence of incremental validity. *Personnel Psychology, 42,* 25-36.

Digman, J. M. (1989). Five robust trait dimensions: Development, stability, and utilty. *Journal of Personality, 57*, 195-214.

Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology, 41,* 417-440.

Digman, J. M., & Inouye, J. (1986). Further specification of the five robust factors of personality. *Journal of Personality and Social Psychology, 50,* 116-123.

Digman, J. M., & Takemoto-Chock, N. K. (1981). Factors in the natural language of personality: Re-analysis, comparison, and interpretation of six major studies. *Multivariate Behavioral Research, 16,* 149-170.

Driskell, J. E., Hogan, J., Salas, E., & Hoskin, B. (1994). Cognitive and personality predictors of training performance. *Military Psychology, 6*, 31-46.

Ellis, H. C. (1969). Transfer and retention. In M.H. Marx, *Learning: Vol. 1. Learning Processes* (pp. 380-478). New York: MacMillan Press.

Ewing, J. (1998). Marines go to battle stations that are workstations. *Government Computer News, 17(31)*, 62.

Eysenck, M. W. (1992). *Anxiety: The cognitive perspective*. Hillsdale, NJ: Erlbaum.

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.

Garcia, V. (2002). *U.S. foreign military training: A shift in focus.* Retrieved September 12, 2003 from, http://www.cdi.org/terrorism/miltraining.cfm

Goldberg, L. R. (1990). An alternative "description of personality:" The Big Five factor structure. *Journal of Personality and Social Psychology, 59,* 1216-1229.

Hackman, J.R., & Oldham, G.R. (1980). *Work redesign.* Reading, MA: Addison-Wesley.

Hall, C. R., Stiles, R. J., & Horwitz, C .D. (1998). Virtual reality for training: Evaluating knowledge retention. *Proceedings of the Virtual Reality Annual International Symposium (VRAIS),* USA, 184-189.

Heacock, B. (1999). Virtual bootcamp (Department of Defense's technology-based training). *Training, 36*(5), 136-138.

Heider, F. (1958). *The psychology of interpersonal relations.* New York: Wiley.

Heinze, C., Goss, S., Josefsson, T., Bennett, K., Waugh, S., Lloyd, I., Murray, G., & Oldfield, J. (2002). Interchanging agents and humans in military simulation. *AI Magazine, 23*(2), 37-47.

Hills, A. (2002). Can we fight in cities? *The Urban Operations Journal.* Retrieved March 10, 2003, from http://www.urbanoperations.com/hills1.htm

Hogan, J., & Hogan, R. (1989). Noncognitive predictors of performance during explosive ordnance disposal training. *Military Psychology, 1*, 117-133.

Holding, D. H. (1965). *Principles of Training.* New York: Pergamon Press Ltd.

Hollenbeck, J. R., Brief, A. P., Whitener, E. M., & Pauli, K. E. (1988). An empirical note on the interaction of personality and aptitude in personnel selection. *Journal of Management, 14*, 441-451.

Hollenbeck, J. R., & Whitener, E. M. (1988). Reclaiming personality traits for personality selection: Self-esteem as an illustrative case. *Journal of Management, 14*, 81-91.

Hough, L. M. (1992). The "Big Five" personality variables – construct confusion: Description versus prediction. *Human Performance, 5,* 139-155.

Hough, L. M., Eaton, N. K., Dunnette, M. D., Kamp, J. D., & McCloy, R. A. (1990). Criterion-related validities of personality constructs and the effect of response distortion on those validities. *Journal of Applied Psychology, 75*, 581-595.

Hunter, J. E., & Hunter, R. F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin, 96,* 73-98.

Janis, I. L. (1989). *Leadership in policymaking & crisis management.* New York: Collier MacMillan Publishers.

Judge, T. A., Higgins, C. A., Thoresen, C. J., & Barrick, M. R. (1999). The Big Five personality traits: General mental ability, and career success across the life span. *Personnel Psychology*, *52*, 621-652.

Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *International Journal of Aviation Psychology, 3*, 203-220.

Kolasinski, E. M. (1995). *Simulator sickness in virtual environments* (Technical Report 1027). Arlington, VA: United States Army Research Institute for the Behavioral and Social Sciences.

Lampton, D. R., Knerr, B. W., Goldberg, S. L., Bliss, J. P., Moshell, M. J., & Blau, B. S. (1995). *The virtual environment performance assessment battery: Development and evaluation* (Technical Report 1029). Arlington, VA: United States Army Research Institute for the Behavioral and Social Sciences.

Lintern, G. (2001). An informational perspective on skill transfer in human-machine systems. In R. W. Swezey & D. H. Andrews (Eds.), *Readings in training and simulation: A 30 year perspective.* Santa Monica: CA: Human Factors and Ergonomics Society.

Loftin, R. B., & Kenney, P. J. (1995). Training the hubble space telescope flight team. *IEEE Computer Graphics and Applications, 15*(5), 31-37.

Loftin, R. B., Savely, R. T., Benedetti, R., Culbert, C., Pusch, L., Jones, R., et al. (1997). Virtual environment technology in training: Results form the Hubble space telescope mission in 1993. In R. J. Seidel & P. R. Chantelier (Eds.), *Virtual reality: Training's future* (pp. 93-104). New York: Plenum.

Macedonia, M. (2002, March). Games soldiers play: Computer games and virtual reality are radically altering the way the military prepares for war. *IEEE Spectrum, 39*(3), 32. Retrieved March 10, 2003, from http://www.spectrum.ieee.org/WEBONLY/publicfeature/mar02/mili.html

Magee, L. E. (1997). Virtual reality simulator (VRS) for training ship handling skills. In R. J. Seidel & P. R. Chantelier (Eds.), *Virtual Reality: Training's Future* (pp. 19-30). New York: Plenum.

Maier, N. R. F. (1965). *Psychology in industry* (3rd ed.). Boston: Houghton-Mifflin.

Matthews, G., & Deary, I. J. (1998). *Personality Traits.* Cambridge, MA: Cambridge University Press.

Mazuryk, T., & Gervautz, M. (1996, February). *Virtual reality: History, applications, technology and future* (TR-186-2-96-06). Vienna, Austria: Vienna University of Technology Institute of Computer Graphics. Retrieved September 12, 2003, from ftp://ftp.cg.tuwien.ac.at/pub/TR/96/TR-186-2-96-06Paper.pdf

McCown, W., & Johnson, J. (1991). Personality and chronic procrastination by university students during an academic examination period. *Personality and Individual Differences, 12,* 413-415.

McCrae, R.R., & Costa, P.T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology, 52*(1), 81-90.

McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality, 60,* 175-215.

McHenry, J. J., Hough, L. M., Toquam, J. L., Hanson, M. A., & Ashworth, S. (1990). Project A validity results: Relationship between predictor and criterion domains. *Personnel Psychology, 43,* 335-354.

Milgram, N., & Tenne, R. (2000). Personality correlates of decisional and task avoidant procrastination. *European Journal of Personality, 14,* 141-156.

Miller, K. E. (2001). Video Games, TV and Aggressive Behavior in Kids. *American Family Physician, 64,* 863.

Mount, M. K., & Barrick, M. R. (1995). The Big Five personality dimensions: Implications for research and practice in human resources management. *Research in Personnel and Human Resources Management, 13,* 153-200.

Mount, M. K., Barrick, M. R., & Stewart, G. L. (1998). Five-factor model of personality and performance in jobs involving interpersonal interactions. *Human Performance, 11*, 145-165.

Mount, M. K., & Barrick, M. R., & Strauss, J. P. (1994). Validity of observer ratings of the Big Five personality factors. *Journal of Applied Psychology, 79*, 272-280.

Pew, R. W., & Mavor, A. S. (Eds.). (1998). *Modeling human and organizational behavior: Applications to military simulations.* Washington, DC: National Academy Press.

Phillips, J. M., & Gully, S. M. (1997). Role of orientation, ability, need for achievement, and locus of control in the self-efficacy and goal-setting process. *Journal of Applied Psychology, 82*, 792-802.

Piedmont, R. L., & Weinstein, H. P. (1994). Predicting supervisor ratings of job performance using the NEO personality inventory. *Journal of Psychology, 128*, 255-257.

Psotka, J. (1995). Immersive training systems: Virtual reality and education and training. *Instructional Science, 23*, 404-431.

Reintzell, J. F. (1997). When training saves lives. *Training and Development, 51,* 41.

Riva, G. (Ed). (1997). *Virtual reality in neuro-physiology.* Amsterdam: IOS Press.

Rolfe, J. M., & Staples, K. S. (Eds.). (1986). *Flight simulation.* Cambridge, UK: Cambridge University Press.

Rose, D., & Foreman, N. (1999). Virtual reality. *The Psychologist, 12*, 550-554.

Rosse, J.G., Miller, H. E., & Barnes, L. K. (1991). Combining personality and cognitive ability predictors for hiring service-oriented employees. *Journal of Business and Psychology, 5*, 431-445.

Sackett, P.R., Gruys, M. L., & Ellingson, J. E. (1998). Ability-personality interactions when predicting job performance. *Journal of Applied Psychology, 83*, 545-556.

Salgado, J. F. (1997). The five factor model of personality and job performance in the European community. *Journal of Applied Psychology, 82*, 30-43.

Salgado, J. F. (1998). Big Five personality dimensions and job performance in Army and civil occupations: A European perspective. *Human Performance, 11*, 271-288.

Siem, F. M., & Murray, M. W. (1994). Personality factors affecting pilot combat performance: A preliminary investigation. *Aviation, Space and Environmental Medicine, 65,* A45-A48.

Silver, S.M. (1992). *The psychology of men in combat - a computable model.* Unpublished manuscript.

Smith, R. D. (1995, October/November). Military training via wargaming simulation. *IEEE Potentials.* Retrieved October 1, 2001, from http://www.simulationfirst.com/papers

Stanney, K. M., Mourant, R. R., & Kennedy, R. S. (1998). Human factors issues in virtual environments: A review of the literature. *Presence, 7,* 327-351.

Stansfield, S., Shawver, D., Rogers, D., & Hightower, R. (1995). Mission visualization for planning and training. *IEEE Computer Graphics and Applications, 15*(5), 12-15.

Swezey, R. W., & Andrews, D. H. (2001). *Readings in training and simulation: A 30 year perspective.* Santa Monica, CA: Human Factors and Ergonomics Society.

Terborg, J. R. (1977). Validation and extension of an individual differences model of work performance. *Organizational Behavior and Human Performance, 18,* 188-216.

Tett, R. P. (1998, July). Is conscientiousness ALWAYS positively related to job performance? *Society for Industrial and Organizational Newsletter (TIP), 36*(1), Psychology. Retrieved March 10, 2003, from http://www.siop.org/tip/backissues/TIPJuly98/tett.htm

Thomas, E. (2001, October 1). The road to Sept. 11. *Newsweek.* Retrieved November 7, 2001, from http://www.msnbc.com/news/632825.asp

Ulate, S. O. (2002). *The impact of emotional arousal on learning in virtual environments.* (Masters Thesis, Naval Postgraduate School, 2002). Monterey California. Retrieved March 10, 2003, from http://www.movesinstitute.org/Theses/Ulatethesis.pdf

APPENDIX A

Scenario Descriptions and Day Assignment

Scenario		Day		Category
#	*Description*	1	2	
1	Left military ID at work.	X	X	Social
2	Passenger has no ID.	X	X	Social
3	Missing license plate	X		Perceptual
4	Hunting rifle is on front seat. When prompted driver will indicate intent to return rifle to armory.	X		Protocol
5	Inappropriate ID (school ID on day 1 & credit card ID on day 2)	X	X	Protocol
6	BOL person Moe is introduced on paper in Session 1 but doesn't appear until Session 2.		X	Long-Term Retention (Special Case)
7	BOL license plate appears. This BOL is issued on radio on day 1 and on paper on day 2.	X	X	Memory
8	Vehicle matching the description issued on a BOL radio alert appears.		X	Memory
9	Injured passenger appears with a bandaged injury.		X	Emergencies
10	Ambulance appears preceded by proper radio alert.		X	Emergencies
11	Ambulance appears but without proper radio alert.	X		Emergencies
12	Admiral's son appears without ID.		X	Social Status
13	Car is missing base stickers.		X	Perceptual
14	Injured passenger appears with no obvious injury.	X		Emergencies
15	Person in first car asks for automatic entry of his friend in a second car.		X	Situational Awareness
16	Driver exhibits suspicious behavior by repetitively looking into his rear view mirror.	X		Situational Awareness
17	The Mayor w/ a drivers' license but no pass	X		Social Status
18	Hunting rifle is on front seat. Upon prompt, driver will say that there is a navy shooting competition.		X	Protocol
19	A vehicle matching the description in the pre-briefing incident log appears.	X		Memory

APPENDIX B

Background Questionnaire

Participant # _____

Background Questionnaire

Please fill out the following items by either circle a response or filling out the open-ended items as indicated. If you have any questions about the items, please feel free to ask the experimenter.

1. Gender? M/ F

2. Age? ___

3. Do you have any previous military experience? Yes / No

If yes, explain.____

4. Are you from a military family? Yes / No

If yes, have you lived on a base? _____

5. Have you ever passed through a military checkpoint? Yes / No

6. What do you know about military checkpoints? (use separate sheet if necessary)

____

____

7. What are the specific duties of a guard who works at a military checkpoint? (use separate sheet if necessary) _______________________________________

____

____

8. What are the requirements for a civilian to enter a military base? (Please be as specific as possible). _________________________________

____

____

9. What are the requirements for a military person to enter a military base? (Please be as specific as possible). _________________________________

____

____

APPENDIX C

Post-Session Questionnaire

Participant# ____

Post-Session Questionnaire

Please answer these questions as honestly as possible as they help us to assess the usefulness of this technology as a training tool. If you run out of space, please feel free to use a separate page to finish your responses or provide additional comments.

1. Did you find this task particularly stressful or challenging?	Very stressful o o o o o Not very stressful at all 1 2 3 4 5
2. How do you feel you performed?	Very well o o o o o Not very well at all 1 2 3 4 5
3. Did you ever actually remove the gun from its holster?	☐ Yes ☐ No
4. Were you ever tempted to remove the gun from the holster? If yes, explain.	☐ Yes ☐ No
5. Do you feel that the military background video was helpful? Please explain.	Very helpful o o o o o Not helpful at all 1 2 3 4 5
6. Did you find the training manual adequate and helpful? Explain.	Very helpful o o o o o Not helpful at all 1 2 3 4 5
7. Did you encounter any problems (e.g., technological difficulties) during your session? Please explain.	
8. Were there any errors you made or incidents that you felt unprepared for? Explain.	
9. Did being in the virtual environment make a difference in doing the task versus on another platform such as a desktop pc? Explain.	

10. If you were assigned to stand guard at a real military base tomorrow, how well do you feel that this training experience prepared you for that type of job? Please be as detailed as possible. (Use a separate piece of paper if necessary.)	
11. Please ask your experimenter for the SSQ form.	

APPENDIX D

Training Brief

Training Brief

You will now be introduced to the main task for this experiment. Your job is to perform the duties of a security guard at a military checkpoint. You will see a brief training video to familiarize yourself with your assignment and be given specific information about how to perform your duties in a virtual environment. You will receive a standardized briefing so that each trainee receives the same instructions. You are free to take notes, however, you will not be able to use them during the actual training session.

Are you familiar with military checkpoints?

Each military base has a checkpoint set up for major entry points to the base. Vehicles and their occupants are screened for security reasons. This base contains a hospital in addition to military personnel buildings and training areas. You are going to work at one of two entrances to a military base. You will be in charge of the checkpoint located at Gate 1. At this checkpoint, you will serve as a head guard tasked with examining each vehicle and occupant for specific information.

You will be working with a virtual partner, however, you will perform the majority of the task. Your partner's job is to cover you in situations where you are overwhelmed or feel that a vehicle or occupant is acting suspiciously. If you ever feel you are in a situation that could get out of control, you are to issue the command, "Cover that car," to direct your partner to provide cover for you. Please issue the command in a loud, clear voice.

Prior to the start of your shift, you will be given a briefing of what happened during the shift before yours. These briefings will contain all the necessary information about previous events that you need to be aware of.

During your shift, cars will approach your gate and attempt entry through your checkpoint. Your job is to check the car for appropriate stickers, passes, license plates, and ask the occupants for their identification cards (IDs). During your shift you must also identify BOL's that indicate a vehicle or person you must "Be on the Lookout" for suspected illegal or potentially threatening activity.

You must follow proper procedure at all times.

Equipment

Walkie-talkie. You will be issued a walkie-talkie. You are to use the walkie-talkie to contact the base at any time you deny entry to a vehicle. Proper procedure requires you to identify your point of contact, identify yourself, give a vehicle description (color of car, number of occupants, etc.), and the reason for denying entry. For example: "Base, this is Gate 1, be advised that a gray car with 2 passengers has been denied entry for improper ID."

Side arm. You have also been issued a side arm. You are to use the gun only if you suspect that a person or vehicle may cause you, your partner, or the base harm.

Checkpoint Protocol

Distinction Between Civilian Requirements and Military Requirements

	ID Type	**License Plate**	**Base Stickers**	**Visitors Pass**
Civilians	Driver License (White card)	Yes	No	*Yes
Military	Military ID (Green card)	Yes	Yes	No

* All civilians require a pass unless accompanied by military personnel.

Checkpoint Procedure:

- To get on base, both drivers **and** passengers must meet certain requirements.
- Military drivers need a base sticker and a green military ID.
- Civilians need a visitor's pass and a white driver's license.
- If a visitor's pass is needed, only one pass per car is necessary.
- Civilians accompanied by military do not need a pass. Civilians always need proper ID.
- People with NCIS (National Criminal Investigative Service) IDs should be allowed immediate entry.
- Ask people who show inappropriate ID for a valid ID.
- Emergency vehicles with radio alert get in immediately (without ID check).
- Emergency vehicles without radio alert are to be treated as a normal car procedure wise (ask for ID etc.)
- People in uniform may or may not represent American military personnel.
- Contraband items are not allowed. Returns to the armory are **not** considered contraband. You may inquire about an item you think is contraband.
- The only people you need to pull over are: (1) people or vehicles who match a BOL, (2) people carrying contraband, or (3) people who demonstrate suspicious behavior.
- Report any and all instances where a driver does not obey your command and instead runs the gate.
- Maintain situational awareness for all people, all vehicles and environmental factors at all time.
- Ask your virtual partner to cover you in the event of suspicious behavior.
- Report all people who run through your gate without your consent.

Summary of Important Commands

- ID please
- Go ahead
- Turn around
- Pull over
- I need to see a valid ID.
- Where is your passengers ID.
- Cover me

Limitations and Troubleshooting

This training exercise takes place in a virtual environment. It is a complex and sophisticated computer program, but it is still a computer program. There are limitations to what can and cannot be done in this environment.

Each interaction you have with a vehicle ends with one of three commands ("Go ahead," "Turn around," or "Pull over"). When you issue one of these commands, the system responds best if you say the phrase exactly. Try not to use additional words.

If you encounter a problem, inform the driver of the problem. If the driver does not resolve the problem, issue a command. If the driver still does not respond appropriately, you may assume one of two things:

1) the driver does not understand you. Thus you can repeat your command in a louder voice, enunciating each word, or you can rephrase it. (or)
2) after having tried again, you might assume that the driver is not complying with your inquiry. If the driver still does not respond appropriately or does not meet the requirements after having repeated your request, you must make a decision and tell the driver to either "Go ahead," "Turn around," or "Pull over."

Other issues to be aware of:

- If there are two ID's in a vehicle, the driver will hold both IDs.
- White ID cards may sometimes appear gray. Please note that this is not intentional. Treat gray cards just the same a white.
- If a driver or passenger does not show proper ID, you may ask again for a valid ID.
- If you find a contraband item, you may inquire about that item and make a decision based on the vehicle occupant's response.
- In any instance where you say 'cover me,' please assume your virtual partner will take necessary action although this action may not be overtly observable.
- Blue cars may appear bluish gray.

APPENDIX E

Incident Logs from Previous Shift

Incident Log for Session 1

CHECKPOINT INCIDENT LOG – MORNING SHIFT – S1	
8am	Moe sighted in the area. He is wanted for questioning for terrorist activity and a BOL has been issued for him. Pull him over on sight. Command has given us a picture of him.
9am	
10am	Blue car with two male drivers attempted to illegally pass through gate. All black cars with two male passengers should be pulled over for investigation and vehicle search.
11am	
12pm	
1pm	
2pm	
3pm	
4pm	
5pm	
6pm	

*Incident Log for Session 2**

CHECKPOINT INCIDENT LOG – MORNING SHIFT – S2_G2	
8am	Moe sighted in the area. He is wanted for questioning for terrorist activity and a BOL has been issued for him. Pull him over on sight. Command has given us a picture of him.
9am	
10am	
11am	Vehicle with license plate ZK4 attempted to run Gate 2 and managed to evade security officials. Pull over and detain on sight.
12pm	
1pm	
2pm	
3pm	
4pm	
5pm	
6pm	

*Note. This log represents the log given to participants in Condition 2 who did experience a first session. The log for participants in Condition 1 differed in only one respect: The words "Command has given us a picture of him" were omitted.

APPENDIX F

Procedure for Clearing Vehicle for Entry

1. Assessing whether a license plate is appropriately in place and does not match a license plate that experimenters have asked the trainee to be on watch for during prebriefing or via radio alert.
2. Identifying the presence or absence of base stickers or, in the case of a civilian passenger, the presence of an appropriate visitor pass.
3. Identifying the presence of suspicious objects that require the trainee to ask questions.
4. Determining whether any identified questionable items need to be for entry into the base. e.g., military weaponry
5. Carrying out proper protocol for emergency vehicle entry when there is and is no a radio advisory.
6. Identifying and detaining the cars that match the description of a vehicle that experimenters have told trainees are on the "be on the lookout" (BOL) list.
7. Establishing proper radio communication for all vehicles identified as critical events.

In addition to assessing the vehicle, the trainee is tasked with assessing the occupants of the vehicle in order to ensure:

1. Ensure that both the driver and any passengers have the appropriate identification.
2. Identify and detaining any people who match descriptions of people found on the "be on the lookout" (BOL) list.
3. Verify that any civilian occupants are represented by a visitor's pass.

4. Ensure that all military vehicles have proper base stickers.
5. That any inappropriate or suspicious driver behavior is acknowledged and handled by our trainee.
6. Establish proper radio communication for all people identified as having suspicious behavior or lacking the necessary requirements for entry.

In addition to assessment of vehicle and occupants, the trainee will be tasked with maintaining good situation awareness and following appropriate protocol for reporting critical incidents including inappropriate or missing identification cards, license plates, base stickers, visitor passes, or the assessment of possible contraband items.

APPENDIX G

Error Classification

During the analysis of results, an error will be recorded for failures to follow protocol in the following instances:

Error Classification for Checkpoint Task

1. Failed to radio in gate runners (special case of fail to make radio call).
2. Failed to ID contraband
3. Incorrectly performed emergency vehicle protocol.
4. Failed to identify suspicious behavior. (e.g., rear view mirror)
5. Failed to make radio call (see gate runners as subcase).
6. Failed to obtain identification for passenger.
7. Failed to distinguish protocol for 'pull over' versus 'turn around.'
8. Asked driver for non-required item.
9. Failed to 'turn around' driver upon noticing missing item.
10. Failed to notice or properly identify missing item.
11. Failed to identify BOL.
12. Misidentified a vehicle or person as a BOL.
13. Other
14. Failed to perform correct protocol for contraband.

Eventual errors for follow-up studies:

- Requests ID every time (with exception of emergency vehicle)
- Indicates reason for denied entry during radio call.
- Makes delayed radio call.

APPENDIX H

Participant Responses

Participants were asked:

"If you were assigned to stand guard at a military base tomorrow, how well do you feel that this training experience prepared you for that type of job? Please be as detailed as possible."

Below are verbatim responses to the above item:

	Participant Responses
GROUP 1	
1	Session 1: It gave me a preview of what could happen. I would be prepared if I experienced the training multiple times. Session 2: I would be prepared only if I didn't have to use my gun.
2	Session 1: The video was very helpful. It could prepare you a little bit. It would be easier outside virtual reality. Session 2: It would have prepared me pretty well.
3	Session 1: It was helpful, but I do not feel that I would be prepared to actually stand guard at a real base. Session 2: It was somewhat helpful but I would want more training before I actually did it.
4	Session 1: I think it was very helpful and it made me learn some of the rules I did not know about base guard. Session 2: I think I'll do well but I'm not sure how I can carry a real gun.
5	Session 1: The experiment trained me well, but I would still make minor mistakes in how to radio the base. Session 2: I would be confident that I will know how to handle myself and do a good job. I would probably need someone with me though, in case I had any problems or questions.
6	Session 1: It would have briefly trained me. I don't feel like I would be a good guard. Session 2: I would be more prepared than before, but I still wouldn't feel comfortable doing it.
7	Session 1: I would feel more comfortable. I would also have a little feeling on what to expect for that job. Session 2: Very well...
8	Session 1: Well, I'd know how to ask for cover when more than one car approaches. I feel I could guard a real military base. Session 2: Sure. It's not hard to check ID's and pull people over.

9	Session 1: It gave me sufficient training so that I would feel comfortable guarding as long as I had a partner to back me up. Session 2: Yes I feel I would be very prepared because I have felt like I have already been on the job and this also gives me a chance to see what my weak areas are that I need to look over.
10	Session 1: I would be better trained for the job from the experience received in this experiment. I started to get more comfortable with the job as time progressed. Session 2: I feel that I could efficiently command my gate. I was much more comfortable with my duties and performance this session.
11	Session 1: I think that I would do ok. More traffic would be present in real life. Session 2: Yes, this gave you the basics. The essentials to do the job.
12	Session 1: This experiment provided me with basic information about the checkpoint. Session 2: Have some ideas about security check.
13	Session 1: I think it has trained me very well. I would be more cautious than if I had not been trained. Session 2: Yes and No. The training experience prepared me to look and ask for basic things. It prepared me for the different reactions of people.
14	Session 1: I feel I would need to be trained more in different situations that may occur. A few more training sessions. An on-the-job training session would be helpful too. Session 2: I feel more prepared after this session. I believe VR is an excellent way of training. You can put the trainee in many different situations to see how they handle it and how they should have handled it. I felt that this experiment was very effective for training.
15	Session 1: I think I would need more training to not be in charge for a few days. However, checking ID's and plates and stickers would be okay. Session 2: Better than last time. Have more awareness.
16	Session 1: It is definitely good training experience, however, I don't think it is sufficient. Guards in training should be able to do it with an actual guard who is significantly acquainted with the job. Session 2: I think it prepared me pretty well. I would be ready to stand guard at the military base. I feel this way because I have done this for the second time.
GROUP 2	
1	Yes, because I know what to look for and what procedures to follow.
2	I think I would need a lot more training because I'm not even very good at the virtual one let alone a real situation. I think this helps let me know what to look for but I probably wouldn't know how to react in a real live situation.
3	I would be some what prepared. Training should be a little longer and the virtual environment should interact more. I feel that I have more knowledge than before but I don't feel that I have every command down to perform it in real life.

4	Ok. More training would be better. More suspicious vehicles and how to react towards them.
5	Gave a good idea of suspicion, i.e., people, vehicles. Good start. May be a bit hard to really interpret "suspicious" behavior.
6	I feel more prepared than if I had done it before this training experiment.
7	I would definitely need more training on when it would be okay to allow certain individuals to enter the base. For the most part, I think I did okay for my first day on the job, but again, I definitely think more training is needed. The guard is extremely important and must know exactly how to handle situations.
8	It didn't fully prepare to go on task tomorrow so I wouldn't not feel ready at all. I would need more experience.
9	This experience has provided a greater awareness of the responsibilities of the guards. This was a good start but to feel comfortable, I would need to undergo intensive training. I need to learn how to react in a more efficient manner.
10	I would do a good job, but I don't think I'd be prepared for a real situation. ex. If the guy with the rifle shot me.
11	Well, exceptionally well.
12	I feel prepared but I feel I need a little more practice.
13	I think I would do great because this session allowed me to be more aware and look for specific information.
14	I think I would do a pretty good job. The main do's and don't were taught for the experiment.
15	I think it would get me prepared to a certain extent. But real life experience throw more at you then that.
16	I think it would prepare me to some degree, but I would probably be too scared to do it in real life.

APPENDIX I

Technological Difficulties

Participants were asked if they encountered any problems or technological difficulties during their session(s).

On 46% (22 of 48 reports) of the occasions, participants responded having experienced no technical difficulties. In instances where participants responded having experienced at least partial technological issues, these responses are shown below.

1. Having to repeat myself and blurriness.
2. Just went a little blurry
3. I misconstrued parts of the car for boxes in the backseat
4. A man in an ambulance was being confrontational. A man who pulled up with a van was acting suspicious.
5. There was a lot of flashing. I couldn't tell if it was my own eyes or the program.
6. Just making sure that the person understood my command so you have to repeat a lot.
7. Sometimes the scene remains still when no virtual car approaching.
8. The glasses repeatedly slipped off my ears.
9. The left eye of the glasses was blinking and blacking out sometimes.
10. Blurred vision when trying to see into car. Goggles kept bleeping.
11. Yes, the goggles were twitching and the head piece sometimes was about to fall off my head.
12. Vision problems (blurriness) and technology not responding to my commands.
13. Sometimes the goggles were blurry.
14. No, the only technological difficulties was the simulation not picking up my voice.
15. My goggles kept falling and it confused me as to whether my goggles were moving or the driver was nodding/looking away. The picture would become somewhat distorted.
16. I did not encounter big problems except the [headset] was too big for my ears.
17. The goggles were too large and heavy. The microphone was also too large.
18. The headset wasn't sized for small heads and ears.
19. Trying to remember all the potential suspects to look for
20. The screen was slightly jumpy
21. It skipped a few times.
22. Glasses were too large for my small ears.
23. The lag was a bit annoying.
24. Something wrong with the glasses during one scenario, the environment looks one-dimensional.
25. One minor blur.
26. The left side of goggles still acting up and causing some frustration.
27. Yes, the wires kept getting in the way.

VITA

JEAN M. CATANZARO

Jean received her B.A. degree from the University of South Florida (USF) with honors in Tampa, Florida in December of 1995. While at USF, she participated in numerous organizations, became inducted into honor societies including Psi Chi, Phi Kappa Phi and Golden Key and was elected most outstanding senator through her participation in the student senate. Her honors thesis was on the topic of workplace violence. Jean completed her masters degree in December of 1998 at East Carolina University in Greenville, North Carolina with her master's thesis topic: Motivators of Internet Usage: How a Keyboard Charmed the World. She received her Ph.D. in Industrial Organizational Psychology with an emphasis on Human Factors at Old Dominion University in Norfolk, Virginia in December of 2003. Her department can be reached at: Department of Psychology, Mills Godwin Building, Rm. 250, Old Dominion University, Norfolk, Virginia, 23529. Throughout her academic career, Jean has obtained a variety of work experiences including projects and research in personnel selection, workplace violence, training applications, societal implications of the internet, vigilance studies, virtual environments, and issues related to the usability of products.