NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

WORKLOAD CAPABILITY TRAINING IN AN INTERACTIVE SYNTHETIC ENVIRONMENT

by

Max E. Schlessel

September 2021

Thesis Advisor: Mollie R. McGuire
Co-Advisors: Christian R. Fitzpatrick
Imre L. Balogh

Research for this thesis was performed at the MOVES Institute.

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 0704-0188*
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC 20503.			
1. AGENCY USE ONLY ***(Leave blank)***	**2. REPORT DATE** September 2021	**3. REPORT TYPE AND DATES COVERED** Master's thesis	
4. TITLE AND SUBTITLE WORKLOAD CAPABILITY TRAINING IN AN INTERACTIVE SYNTHETIC ENVIRONMENT			**5. FUNDING NUMBERS**
6. AUTHOR(S) Max E. Schlessel			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000			**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A			**10. SPONSORING / MONITORING AGENCY REPORT NUMBER**
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release. Distribution is unlimited.			**12b. DISTRIBUTION CODE** A

13. ABSTRACT (maximum 200 words)

After a substantial investment in simulated decision-making technologies in the form of the Tactical Decision Kit (TDK), the Marine Corps is struggling to capitalize on that investment in a fleet setting as well as a schoolhouse setting. The system neither entices leaders to train their Marines nor does it allow trainers to utilize the full potential of an interactive synthetic environment (ISE), especially in a way that trains Marines by challenging them to perform tasks that are expected of them in combat situations. In today's environment, small unit leaders are subjected to heavily task-saturated and high-workload environments. Creating a scenario using current commercial off-the-shelf (COTS) hardware and software available in the TDK will have the potential to train small unit leaders, specifically infantry squad leaders; would allow training to occur at increased intervals; and enable leaders to assess how well their small unit leaders are capable of making decisions in such an environment. This thesis will test the ability of ISE scenarios to train small unit leaders and how well they can be expanded with interoperability into legacy systems that the Marine Corps currently uses.

14. SUBJECT TERMS interactive synthetic environment, tactical decision kit, Virtual Battlespace, VBS, decision making, workload analysis			**15. NUMBER OF PAGES** 89
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UU

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

WORKLOAD CAPABILITY TRAINING IN AN INTERACTIVE SYNTHETIC ENVIRONMENT

Max E. Schlessel
Captain, United States Marine Corps
BS, United States Naval Academy, 2015

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN COMPUTER SCIENCE

from the

NAVAL POSTGRADUATE SCHOOL
September 2021

Approved by: Mollie R. McGuire
Advisor

Christian R. Fitzpatrick
Co-Advisor

Imre L. Balogh
Co-Advisor

Gurminder Singh
Chair, Department of Computer Science

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

After a substantial investment in simulated decision-making technologies in the form of the Tactical Decision Kit (TDK), the Marine Corps is struggling to capitalize on that investment in a fleet setting as well as a schoolhouse setting. The system neither entices leaders to train their Marines nor does it allow trainers to utilize the full potential of an interactive synthetic environment (ISE), especially in a way that trains Marines by challenging them to perform tasks that are expected of them in combat situations. In today's environment, small unit leaders are subjected to heavily task-saturated and high-workload environments. Creating a scenario using current commercial off-the-shelf (COTS) hardware and software available in the TDK will have the potential to train small unit leaders, specifically infantry squad leaders; would allow training to occur at increased intervals; and enable leaders to assess how well their small unit leaders are capable of making decisions in such an environment. This thesis will test the ability of ISE scenarios to train small unit leaders and how well they can be expanded with interoperability into legacy systems that the Marine Corps currently uses.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

CASEVAC	casualty evacuation
CFF	call for fire
CLT	cognitive load theory
CMC	Commandant of the Marine Corps
CNR	comm net radio
COA	course of action
COC	combat operations center
COTS	commercial off the shelf
CPG	Commandant's Planning Guidance
DAGR	defense advanced GPS receiver
DIS	distributed interactive simulation
DOD	Department of Defense
DVTE	Deployable Virtual Training Environment
FDC	fire direction center
FMF	Fleet Marine Force
FOPCSIM	forward observer personal computer simulation
GRG	gridded reference graphic
HLA	high level architecture
ISE	interactive synthetic environment
ISMT	indoor simulated for marksmanship trainer
ISR	intelligence, surveillance and reconnaissance
ITDG	interactive tactical decision-making game
JFO	joint fires observer
LAN	local area network
LAR	light armored reconnaissance
LAV	light armored vehicle
LVC-TE	Live, Virtual, and Constructive – Training Environment
M&S	modeling and simulations
MAGTF	Marine Air-Ground Task Force
MCMSMO	Marine Corps Modeling and Simulation Management Office

MOUT	military operations in urbanized terrain
NCO	non-commissioned officer
NPS	Naval Postgraduate School
PM TraSys	Program Manager, Training Systems
T&R	training and readiness
T/E	table of equipment
T/O	table of organization
TDK	Tactical Decision Kit
TLX	task load index
UAS	unmanned aerial system
USMC	United States Marine Corps
VBS	Virtual Battlespace

ACKNOWLEDGMENTS

I am indebted to the ranks of Marine Infantry Sergeants, from whose expertise, fervor, and conscientiousness won the day at Belleau Wood and at Iwo Jima, from Inchon to Hue City, during long hours of fighting in Fallujah and Marjah and in many more climes and places. I would like to especially thank Staff Sergeant Daniel Myers for teaching me to lead, Sergeant Edward Rose for teaching me how to grunt, and Sergeant Connor Hawkins for teaching me how to think critically. This work is for them and their legacy, in hopes that Marines in the future will be better prepared for the next battle.

I am grateful for the opportunity to work with the Marines of 2d Light Armored Reconnaissance Battalion and to the time and effort afforded to me not only by the Marines who participated in this work directly but the willingness of Lt Col Scott Cuomo to receive my work with open-mindedness and as an opportunity to improve the future warfighter.

This work could not have been done with the assistance and patience of my thesis advisors. Mr. Fitzpatrick's optimism, Dr. McGuire's ability to help me understand cognitive capacity, and Dr. Balogh's encouragement were some of the greatest driving factors in completing this work.

Finally, I would like to thank my mother and father for their love and support throughout all my education and upbringing, which allowed me to be where I am today. I extend such gratitude to all my family, friends, and mentors, whom are many in number and whose shoulders I have long stood on to accomplish a work such as this.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

There has been increased interest in the fidelity and capabilities of training simulations throughout the Department of Defense. Many such simulations provide opportunities for Marines to train without expending live ammunition or even leaving the barracks. However, even with the ever-increasing fidelity of graphics cards, processing speed, and computing power, many Marines are reluctant to use the training systems and software even when it is readily available to them. Furthermore, there has long been a fine line between a game and a valid training tool, and commanders and unit leaders alike have been hesitant in utilizing and exploiting the advantages offered by training tools that have been programs of record for over a decade. While the benefits of training in simulations can be obvious, they include being able to train Marines without ever leaving the barracks for expensive and time-consuming field operations as well as being able to provide multiple repetitions of training in a single day. Instead of interactive synthetic environments increasing in proliferation, programs that contain environments have been divested in recent years and are coupled with a variety of issues that make it difficult and create barriers to their use. In an effort to create more mature infantry Marines through higher fidelity simulations and encourage training within interactive synthetic environments, more work needs to be done to refine the programming, architecture, and training objectives that can lead to decreased cognitive workload from using such training tools. The research will inform how easing cognitive workload can be trained within interactive synthetic environments and how the Marine Corps infantry can benefit from the use of computer simulations systems.

A. BACKGROUND

As a part of the Marine Corps Force Design 2030, the Commandant of the Marine Corps (CMC) demanded that the Marine Corps conduct full-scale, empirically based experimentation that must be deliberate and iterative, informed by both threat developments and technology advancements [1]. Serious games and interactive synthetic environments can serve as those experiments [2]. A preponderance of experimentation regarding interactive synthetic environments has been in the aviation field [3]. Live, virtual, and constructive simulations such as the Marine Corps' Infantry Immersion Trainer (IIT) have been used in the past to provide training for the infantryman. Additionally, the Marine Corps' interest and procurement of the Tactical Decision Kit (TDK) with multiple platforms such as Virtual Battlespace 3 (VBS3), HoloLens, and call for fire simulators exhibited that the Marine Corps had a vested interest in using simulations to train infantrymen.

Types of simulations the Marine Corps has readily available to its Marines include mixed live and virtual simulations such as the IIT. The IIT serves as a simulation tool to evaluate squad or platoon sized elements in various tactical scenarios. It requires a fully staffed and contractor run building to run simulations for the squad or platoon sized elements. A simulation such as VBS3 emphasizes training a fire team to company sized element, and can be executed without the need of numerous contractors or staff. However, having on-site developers able to program and bolster the scenarios can greatly impact the training outcome. Units have augmented a lack of on-site programmers by training Marines to serve as white-cell moderators for VBS3 events. Call for fire simulators have been used for decades [4] and can serve to train individual Marines on how to call for fire. A stand-alone example of a call for fire simulator is the Forward Observer Personal Computer Simulation (FOPCSIM). FOPCSIM can be used without another human controller or moderator, but simulation centers used by Marines often require a moderator to be present to assist with training.

The closest aligned research on task saturation and workload capacity training was conducted in a 2020 study by a team of French researchers that culminated in [5]. They used training scenarios in virtual reality to measure mental workload capacity. Task design was based on a flight simulator that was able to create a workload that allowed the researchers to study multitasking and the use of automation. The application used several different tasks including monitoring, tracking, maintaining a schedule window, communication, and resource management. The objectives of the study were threefold: to find whether increasing the level of difficulty for each task would increase the users' mental workload, to find if additional tasks would increase the users' mental workload, and if the subjective mental workload would have an effect on the task performances. The study concluded that "different training scenarios would be required based on the progression of mental workload over time using different task level combinations" [5].

B. PROBLEM STATEMENT

Current approaches to the development of interactive synthetic environments for the Marine Corps Infantry community do not account for the following:

- focusing on cognitive workload performance from each iteration of the training
- designing a curriculum for increasing performance with use of the training tool
- educating leaders and Marines on how to properly use the hardware, software, and effectively build training with use of interactive synthetic environments
- designing a complete package which can enhance the fidelity of the simulation to include using radio communication devices.

C. OBJECTIVES

The primary objective of this research was to better understand how federated simulations and networking can impact the cognitive workload capacity of Marine Corps infantry squad leaders.

This research will assist the Marine Corps in achieving a better modeling and simulation (M&S) product for its most junior members. It will aim to improve the opportunities for the CMC to reach his goals of creating a more mature infantry fighting force.

This research did not include integration of virtual reality into a synthetic environment.

D. RESEARCH QUESTIONS

1. Is an ISE capable of receiving and processing real time voice inputs in the form of distributed interactive simulation (DIS) voice inputs?

2. Can current ISE legacy systems be used to incorporate interoperability via high level architecture or otherwise with other simulations used by the Marine Corps such as the Supporting Arms Virtual Trainer (SAVT) or Infantry Immersion Trainer (IIT)?

3. Can training within ISE scenarios be scaled to provide greater or lesser task saturation to better edify small unit leaders as they learn to adapt to high workload environments?

E. THESIS DESIGN

This thesis establishes definitions and explores research in the areas of cognitive load theory, pupillometry, and Marine Corps doctrines and simulations in Chapter II. Chapter III offers a detailed explanation of the method of the experiment. Chapter IV presents the results from the experiment. Lastly, Chapter V describes the conclusions from this research. It will additionally describe the author's recommendations for simulations use at the tactical level within the USMC infantry.

II. LITERATURE REVIEW

A. OVERVIEW

This chapter includes an overview of the body of work relevant to understanding the stated problem. It begins with an examination of USMC simulations, to include simulations in the DOD and the advent of ISEs for training purposes and explores the doctrinal basis and requirements for USMC training. It then investigates the topics of measuring cognitive load, cognitive work theory, and interoperability. The author then provides an overview of the different technology and concepts being tested and fielded.

B. BRIEF HISTORY OF USMC SIMULATIONS

1. Military Modeling and Simulations

The importance of DOD modeling and simulations was elevated within the past two decades formally and within many of the USMC communities. The U.S. House of Representatives passed House Resolution 487 on 16 July 2007, placing particular emphasis on the M&S community as a National Critical Technology [6]. Resolution 487 defined M&S as "a unique application of computer science and mathematics that depends on the validity, verification, and reproducibility of the model or simulation" [7]. Furthermore, the resolution had an emphasis on DOD projects that used M&S in the past, with the foremost example being the expansion of scientific knowledge concerning nuclear chain reactions during the Manhattan project, a simulation which could replicate the reaction process and ultimately contributed to the end of World War II. The resolution also added that the primary continuous benefit of M&S would be to "provide vital strategic support functions to our Military" [7].

a. Uses of DOD Simulations

Military simulations are often associated with training, and while there has been an overwhelming success with pilot training using simulations, the functions and uses of simulations within the military can be expansive. A simulation is a representation of a system or process, and through a simulation a model may be utilized to achieve an end state of a realistic or intricate environment [6]. Such environments can range from flight simulators to urban buildings which can be used as backdrops from which to create complex scenarios. Within the Marine Corps, such scenarios serve two major purposes, to either train Marines and Sailors or to conduct rehearsals. Marine Corps Doctrinal Publication 5, Planning, states that Marines "should think of planning as a learning process-as mental preparation which improves our understanding of a situation. In its simplest terms, planning is thinking before doing" [8]. Simulations used by the Marine Corps and intended for rehearsals can be vital according to the Marine Corps Planning Process, which places particular emphasis on "transition events," or the events that occur when transitioning from the planning process to execution of a mission [9]. Such events are all excellent candidates for simulations, and include elements such as:

- full dress rehearsal
- reduced force rehearsal
- key leader rehearsal
- combined arms rehearsal
- rehearsal of concept drill
- communications exercises
- terrain model brief
- map brief
- transition brief [9]

The simulations used in this thesis serve purposes to include training and rehearsals, but the relevance of training simulations is important to note for the stated problem [9].

b. Historical Uses of Simulations

The advent of simulations in the DOD was in 1910, when aviators first experimented with ground-based machines that could replicate powered flight. The increased demand for pilots in World War I necessitated such simulators, and just as today it is understandable that student aviators would first practice on the ground because of safety as well as availability of aircraft. Aviators, however, were not the only troops to benefit from early simulations [6]. Just as World War I created a need for the advanced training of pilots, infantry and cavalrymen were also needed in large numbers; particularly, wooden mechanical horse simulators were used for cavalry training [6].

c. *Development of the Deployable Virtual Training Environment*

The Marine Corps began developing tactical decision-making simulation technology as a part of an initiative by the Program Manager, Training Systems (PM TraSys) Science and Technology division. In 2004, PM TraSys attempted to "achieve low-cost computer-based gaming technologies that could provide realistic scenario-based training for individual Marines, small units, and Marine air-ground task force (MAGTF) staffs" [10]. They conducted rigorous testing and cognitive task analysis as well as training effectiveness evaluations of each tactical decision-making technology that they planned to field. The training methodology outcomes of the early 2000s technologies included [10]:

- conduct planning based on the operations order provided in the scenario
- execute the plan in the scenario
- conduct an after-action review [10]

The program also aimed for total interoperability across the MAGTF. It aimed for a horizontal integration of training systems which would include the ability to train teams from different elements of the MAGTF (different elements of the MAGTF may include fire teams from different rifle companies) as well as for a vertical integration which would train teams from different echelons of the MAGTF (such echelons from company, battalion, and regimental level units) [10]. The program predicated such integrated capability specifically using high-level architecture interoperability [10].

Precursor systems included Marine Doom and Virtual Battlespace 1 (VBS1), which were both interactive synthetic environments (ISE) which could replicate first person training simulations and provided some of the outcomes desired by PM TraSys in their goal of achieving a tactical decision-making system. Marine Doom was a first major project developed by the Marine Corps Modeling and Simulation Management Office (MCMSMO) and was built out of an initiative by 1stLt Scott Barnett and Sgt Dan Snyder, who aimed to build on a commercial off the shelf (COTS) game, Doom II. They modeled a four-man fire team within the game and incorporated the weapons systems that were in the table of equipment (T/E) [10]. VBS1 was developed as a part of the DVTE system and was an adaptation of a COTS game from Bohemia Interactive Studio. VBS1 supported 32 users on a local area network (LAN) or across the internet and was an ISE that included photo-realistic terrain, user definable mission scenarios, and variable environmental conditions that could all facilitate the training of small unit tactics [11].

The DVTE became a program of record in fiscal year 2006, although testing and design took place in the years preceding that period [11]. There were many different tactical decision-making simulations included in the original DVTE. They included real time strategy games such as Tactical Operations Marine Corps, Close Combat Marines, Combat Decision Range, MAGTF XXI, and Logistics Tactical Decision-making Scenario. The DVTE also included ISEs in addition to VBS1, which included Close Combat: First to Fight, Close Combat: Antiterrorism, and Joint Terminal Attack Controller Tactical Decision-making Simulation [10].

As the DVTE system began to materialize, the lines of effort were refined to two main branches of the DVTE suite, the first being the combined arms network (CAN) and the second being the infantry tool kit (ITK) [11]. According to LtCol Robert Armstrong, deputy director of the Training and Education Technology Division in the Marine Corps' Training and Education Command in 2002 and during the research and development of the DVTE system, "DVTE is focused on first-person interactivity with the trainee, and strives to maintain individual tactical and decision-making skills" [11]. He also stated that the two branches of the DVTE program were a divided effort to tailor the program to the training audience, and that small unit infantry units would train using the ITK while vehicle operators and infantry commanders would train using the CAN, as visualized in Figure 1.

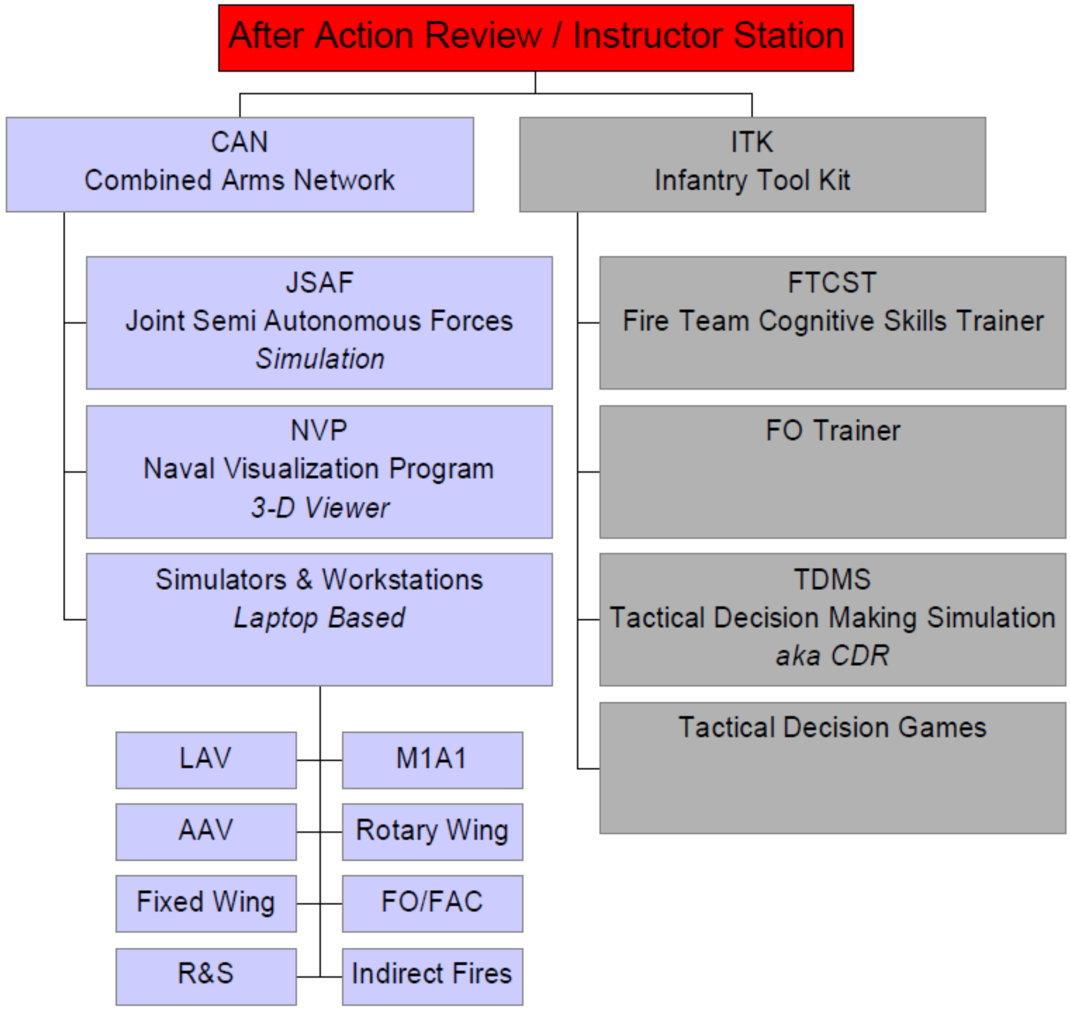


Figure 1. Original organization of the DVTE. Source: [11].

Of note, the initial version of the DVTE focused on call for fire (CFF) training as well as training within ISEs such as VBS1 which was a component of the Fire Team Cognitive Skills Trainer. The first use of the DVTE system was in December 2001 in Camp Lejeune, North Carolina. The main issues exposed in that testing phase included the CAN needing developer support in a greater capacity then what was planned, and the "necessity to migrate all CAN applications to a single, non-proprietary image generator" [11]. The DVTE system at that phase was optimistically positioned to not only conduct individual training of Marines through the ITK and CAN construct, but also more intensive mission scenarios such as Tactical Recovery of Aircraft and Personnel (TRAP), force on force training, and a helicopter assault. Even in the early iterations of the DVTE, developers assessed that it was a quality training tool that could provide the training prescribed by the Marine Corps Training and Readiness Manuals. As a standard for all Marine Corps training, satisfying the standardized training requirements was imperative for any novel training system and according to Table 1, the DVTE system was more than capable of achieving training standards across the MAGTF [11].

Table 1. Assessment of the capability of DVTE to deliver training as required in Marine Training and Readiness Manuals, 2002. Source: [11].

Training Standard	% Covered
HLMA AVG	93
Execute anti-armor CIFS mission	100
Execute MEDEVAC enroute procedures	100
Deliver ordnance using rockets and guns	100
ARTY AVG	88
Direct operations of forward observer team	60
Conduct immediate suppression	100
Conduct SEAD using artillery	80
INF AVG	91
Employ LAV-AT in support of offensive operations	100
Plan fire support	80
Make a tactical decision	100
LAR AVG	100
Employ command and control measures	100
Perform screen operations	100
Conduct tactical movement	100

d. DVTE as a Program of Record

After the initial phase of DVTE development, the suite saw minor updates over the following decades to the present day. Notably the VBS contracts through Bohemia Interactive continued. In 2006, VBS2 was introduced, and the Marine Corps overhauled the DVTE system in 2012 with an estimated dollar value of $10.33 million for three years of research, development, testing and evaluation as well as operations and maintenance [12]. For full operation capability to be achieved by FY17, the DVTE system would include after action reports (AAR) enhancements, artificial intelligence, degraded communications, and various other tools. VBS3 was then purchased at its release in 2013, and as of September 2020, the Marine Corps acquired hundreds of VBS4 licenses for the DVTE program [13].

e. The Tactical Decision Kit (TDK)

A recent and major change that the Marine Corps made to the DVTE program was the institution and development of the TDK system in 2017 [14]. The TDK system was developed by officers and NCOs from 2nd Battalion, 6th Marine Regiment in conjunction with the Office of Naval Research and were intended to "provide a means to challenge Marines to think critically, innovate smartly, and adapt rapidly in complex environments against adaptive enemies" [14]. The program was developed in addition to the DVTE system, and after nearly twenty years of DVTE use in infantry battalion, the Marine Corps Rapid Capabilities Office determined a need for a decision-making tool that offered technologies different than those in the DVTE. The TDK had three major features, which were Interactive Tactical Decision-making Game (ITDG), VBS3, and augmented reality in the form of use of HoloLens. ITDG was a software which allowed operational terms and graphics overlays on map imagery to create tactical decision-making scenarios for Marines. VBS3 was used in a similar fashion to how it was used in the DVTE, mainly to place Marines in up to platoon-sized force-on-force scenarios in which they were forced to think tactically, make decisions, and communicated" [15]. The TDK used an incremental approach to training Marines while in garrison and aimed to increase the fidelity of the simulations over time. As can be seen in Figure 2, the TDK was intended to begin with ITDG training, then use of VBS3 for force on force live missions, and lastly incorporated a mission debriefing tool that synchronized training occurrences across different platforms.

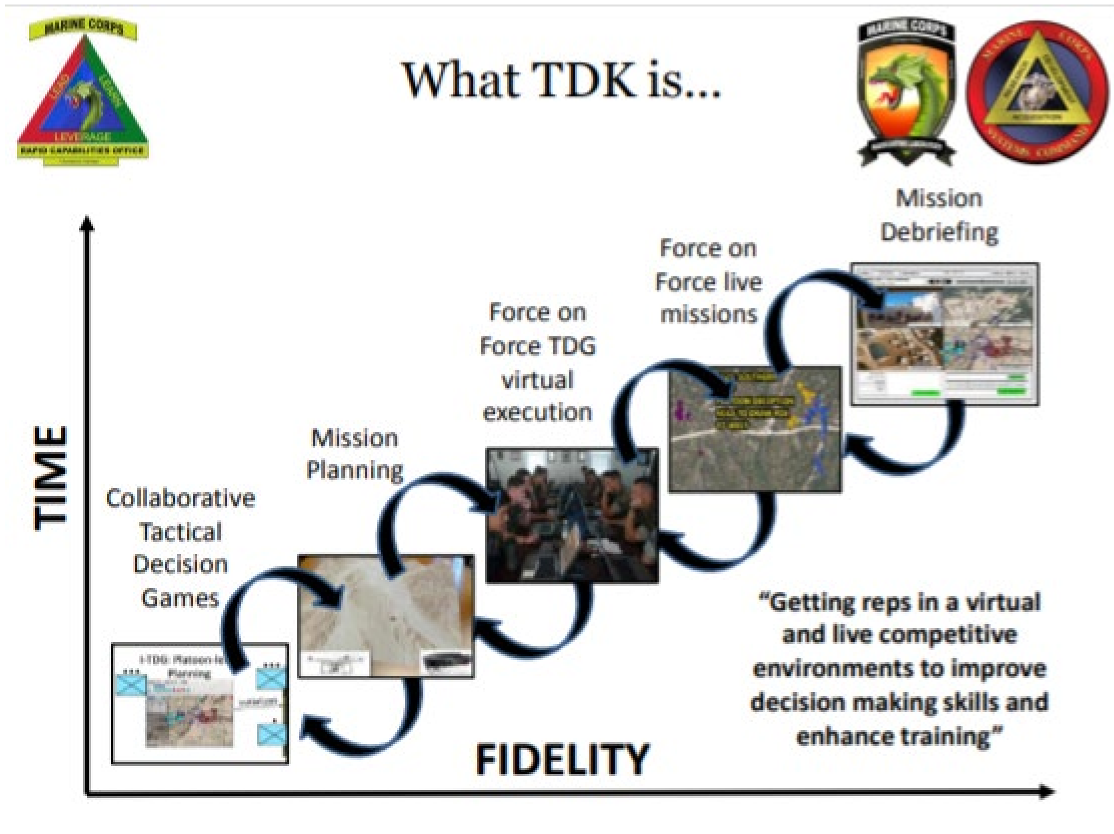


Figure 2. The fidelity-building model of the TDK system, beginning with ITDG use and transitioning into mission debriefs. Source: [15].

2. Summary

The Marine Corps began investing heavily in ISE simulations with the advent of the DVTE and VBS1 starting around the year 2001. Even with a robust layout of simulations, each iteration of the DVTE did not see significant use although a tremendous amount of research was completed on the cognitive skill development capabilities of each simulation. The Marine Corps also has continuously relied on VBS as its sole provider as an ISE since 2000, to include submitting justification waivers for each new iteration of VBS as they are needed for the approval to procure the software "using other than full and open competition" [12]. Such reliance on old systems has developed a stagnant technology that has not seen much improvement in the past two decades. While other ISEs have proven effective as training tools such as the in the case of the Iowa State research team, the Marine

Corps has almost refused to change its procurement process and approach to training over the past 20 years.

C. COGNITIVE LOAD THEORY

1. Cognitive Load Theory (CLT) Defined

A major component of the aforementioned ISE and training construct for the Marine Corps considered cognitive load theory and the impact of cognitive skill gained by training in an ISE. A main goal of the tactical decision-making process was to teach cognitive skills, and the effectiveness of such systems are measured by cognitive task analysis coupled with the training effectiveness of the simulation [10]. According to Paas, Ayres, and Pachman, "Cognitive load can be defined as a multidimensional construct representing the load that performing a particular task imposes on the learner's cognitive system" [16]. A subject's cognitive load is often times a result of their mental effort as well as performance given a certain task. Cognitive load theory offers the basic assumption that an instructional design, "that results in unused working memory capacity by lowering extraneous cognitive load may be further improved by encouraging learners to engage in conscious cognitive processing that is directly relevant to the construction and automation of schemas" [10]. In essence, the ability to lower extraneous (ineffective) cognitive load leaves more working memory available for a learner to increase their germane (effective) load.

2. Types of Cognitive Load

There are three main types of cognitive load which are intrinsic load, extraneous or ineffective load, and germane or effective load [16]. Another important assumption concerning CLT is that intrinsic and extraneous loads are additive, and therefore the total amount of load that a learner can handle must account for all three types [17]. It is therefore imperative that the three cognitive load types are accounted for and optimized in any training scenario. Figure 3 exhibits the additive nature of intrinsic and extraneous load.

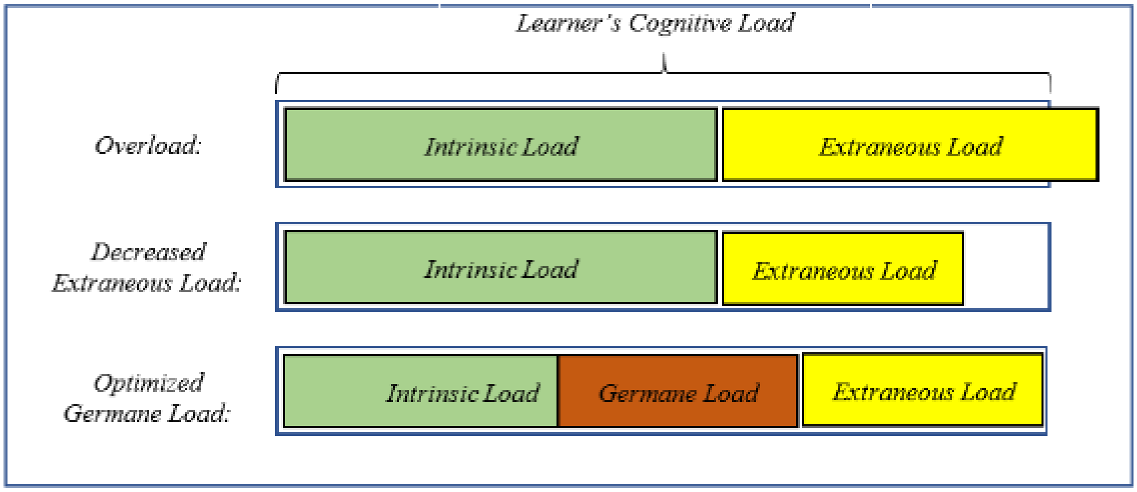


Figure 3. Examples of the additive nature of cognitive load types

a. Intrinsic Load

Intrinsic load relies heavily on the number and interactivity of the elements that must be processed by the learner [18]. The number of elements includes the concepts or procedures that require comprehension or knowledge. Meanwhile, the interactivity of elements are facets of the material that either have high or low dependencies. Interactivity directly corresponds to the degree to which the elements of a task can be learned in isolation [17].

b. Extraneous (Ineffective) Load

Extraneous load consists of factors that are imposed by instructional procedures [18]. John Sweller postured that extraneous load could have similar underlying causes as intrinsic load, specifically "that element interactivity is major source of working memory load" [17]. It is generally understood though that "extraneous load is imposed by instructional procedures" [18].

c. Germane (Effective) Load

Germane load is therefore the result of learning as a function of working memory resources used to deal with intrinsic cognitive load [18]. The ability for learners to link interactive elements to process them results more effectively in germane load. It should be noted that while it involves the element of interactivity, germane load distinguishes itself from intrinsic load as it relies on learner characteristics. Learner characteristics are the memory resources used by the learner in dealing with intrinsic load. Factors that affect learner characteristics may include age, learning style, and positive attitudes and motivation [19]. Such an emphasis is placed on the "working memory resources that the learner devotes to dealing with the interactivity of the intrinsic load" [20].

3. Measuring Cognitive Load

There are varying methods for how cognitive load is measured. Research in the field first suggested that cognitive load could predict instructional effectiveness, and therefore were supported by indirect measures of cognitive load, primarily by accounting for learner error rates and the time taken for a subject to complete a task [21]. As cognitive load theory was developed throughout the 1990s, it became more apparent that there was a need for more direct measures of cognitive load than error rates or time spent learning a task. Instead of using subjective measures of cognitive load, it was found that a direct measure of cognitive load were subjective measures [22]. Characteristics that can influence cognitive load include elements that concern the task in questions as well as elements of the subject. The task characteristics may include format, complexity, use of multimedia, time pressure, and pacing of instruction, while characteristics that affect the learner could be expertise, age, and spatial ability [10].

a. Subjective vs. Physiological Measures

In 1992, Paas recognized that subjects were able to determine their own amount of mental effort dedicated to a specific task. Albeit a subjective measure, learners could rate their intensity of mental effort as an indication of cognitive load given certain parameters. The following Likert scale was first used to help subjects' self-determination of their mental effort. It ranged from very, very low mental effort to very, very high mental effort and scored likewise from 1 to 9 [21]. Paas found in the same study that there was a consistency between subjects' self-rated mental effort and their test performance. As comparisons were made between test performance and self-evaluation, it became clear to researchers that there was in fact an ability to effectively measure cognitive load through subjective means. Figure 4 gives an example of the Paas (1992) subjective rating scale as his subjects would have seen it.

In solving or studying the preceding problem I invested

1. very, very low mental effort
2. very low mental effort
3. low mental effort
4. rather low mental effort
5. neither low nor high mental effort
6. rather high mental effort
7. high mental effort
8. very high mental effort
9. very, very high mental effort

Figure 4. Paas subjective rating scale. Source: [23]

Using such a scale, Paas

> found a match between self-rated mental effort and test performance. Learners who were presented an instructional design hypothesized to impose a low cognitive load had superior learning outcomes and rated their mental effort lower than students who were presented a design hypothesized to be high in cognitive load. [21]

Through his studies, Paas was able to assert that subjective measures of cognitive load were more effective than other measures, either the indirect measures such as errors or time, or physiological measures such as the spectral analysis of heart rate. Additionally, it was found that the physiological measures had difficulties in cognitive load measurements that subjective measures did not have. It was found that physiological measures could not detect between treatment groups, such as a group that was given a high cognitive load versus a group that was given a low cognitive load. Instead, physiological measures were only able to detect between period of mental activity or any cognitive load and periods of no cognitive load or mental activity. Another element that added a high level of fidelity to the Paas subjective rating scale was its lack of perceived intrusiveness when compared to the physiological measurements. A heart rate monitor can require a participant to take off his shirt and wear an uncomfortable or cold device, such as the author had to during an experiment in 2017 as shown in Figure 5.

Figure 5. The author being fitted with a heart rate monitor to measure cognitive load

Even with heart rate monitors becoming more advanced, they can still be uncomfortable or seemingly intrusive for wearers even when worn on the wrist. Additionally, other instruments used to measure physiological signs of cognitive load can be viewed as intrusive to the participant, such as an eye-tracking device which must be calibrated, and the participant is made aware that his pupils and face are being recorded. Lastly, saliva samples which are taken to measure salivary cortisol as related to physiological or physical stress can be very intrusive to a participant. The saliva samples are not only retained sample of the participant's biomass, but the physical act of providing the saliva can be difficult and unpleasant, especially after more than one saliva samples.

4. TLX Survey

In 1988, Sandra Hart and Lowell Staveland published a multi-year research effort aimed at defining the factors of subjective workload. The researchers found that subjective ratings could provide the most valid and sensitive indicator in workload assessment methods [22]. In their research, they proposed a rating scale of ten component scales which included:

- overall workload
- task difficulty
- time pressure
- own performance
- physical effort
- mental effort
- frustration
- stress
- fatigue
- activity type [22]

In the study the subjects self-evaluated on the above listed criteria with an overall goal of determining which scales "best reflected experimental manipulations within experiments" [22]. The researchers conducted a study with 247 subjects and worked to determine how the various criteria should have been weighted and the reliability in the evaluation techniques. The appendix contains an example of the modern TLX survey. Appendix A contains the standardized NASA-TLX survey created from Hart and Staveland's work.

D. M&S INTEROPERABILITY WITHIN MARINE CORPS SIMULATIONS

Some of the issues involved with M&S interoperability within Marine Corps simulations largely implicate VBS3 and its use within live, virtual, and constructive training environments (LVC-TE). Voice communications over a radio environment often muddle the user experience and fidelity of USMC simulations, and therefore it is imperative that the radio environment works correctly and efficiently. The two requirements for a radio environment are low latency and the ability for a large number of audio streams to be carried [24].

E. SUMMARY

In conclusion, this research incorporates the twenty years of USMC M&S as it relates to the DVTE and TDK system. It further examined the implications of cognitive load theory in the experiment design and the use of the TLX survey. Lastly, it investigates how interoperability affects training in such systems.

THIS PAGE INTENTIONALLY LEFT BLANK

III. METHOD

A. DESIGN

The design of the experiment is a two-group comparison design, with the level of cognitive load training manipulated between of the two groups. Of interest is the participant's cognitive workload difference between the two groups, which was measured in the participant's (a) performance, (b) electrodermal activity (galvanic skin response), (c) screen monitoring via eye tracking, and (d) self-evaluated performance (as described in the post-exercise questionnaire).

The hypotheses tested is:

- H_1: The Marines with increased training in an interactive synthetic environment will decrease the cognitive workload over those with less training in the interactive synthetic environment ($p_{Trained} - p_{Untrained} > 0$).

B. PARTICIPANTS AND LOCATION

A power analysis was conducted to determine the appropriate sample size. The Heinrich-Heine-Universität G*Power program was used to calculate the needed sample size along with the Cohen's d approach. The type of power analysis was an "a priori" test, meaning that the sample size was computed given the effect size, power, and the alpha. The power study was conducted based on the t-test family and consisted of a difference in means between two independent groups. In this case the first independent group was the control group, and the other being the experiment group. The power analysis with an effect size (d) of 0.80, alpha error probability = 0.05, and Power ($1 - \beta$ error probability) = 0.80. Such variables resulted in an estimated sample size of 42 with 21 participants needed in each group. Figure 6 displays the results of the power study completed with G*Power. Group A was associated with the decreased cognitive workload factor and Group B was associated with the baseline training factor. Due to operational constraints on 2nd Light Armored Reconnaissance Battalion, only 34 of the targeted 42 participants were able to participate. The reduction in sample size corresponded to having only 16 of the 21 needed participants in each group but given the time constraints and operational nature of the Marines it was adequate for the purposes of the experiment.

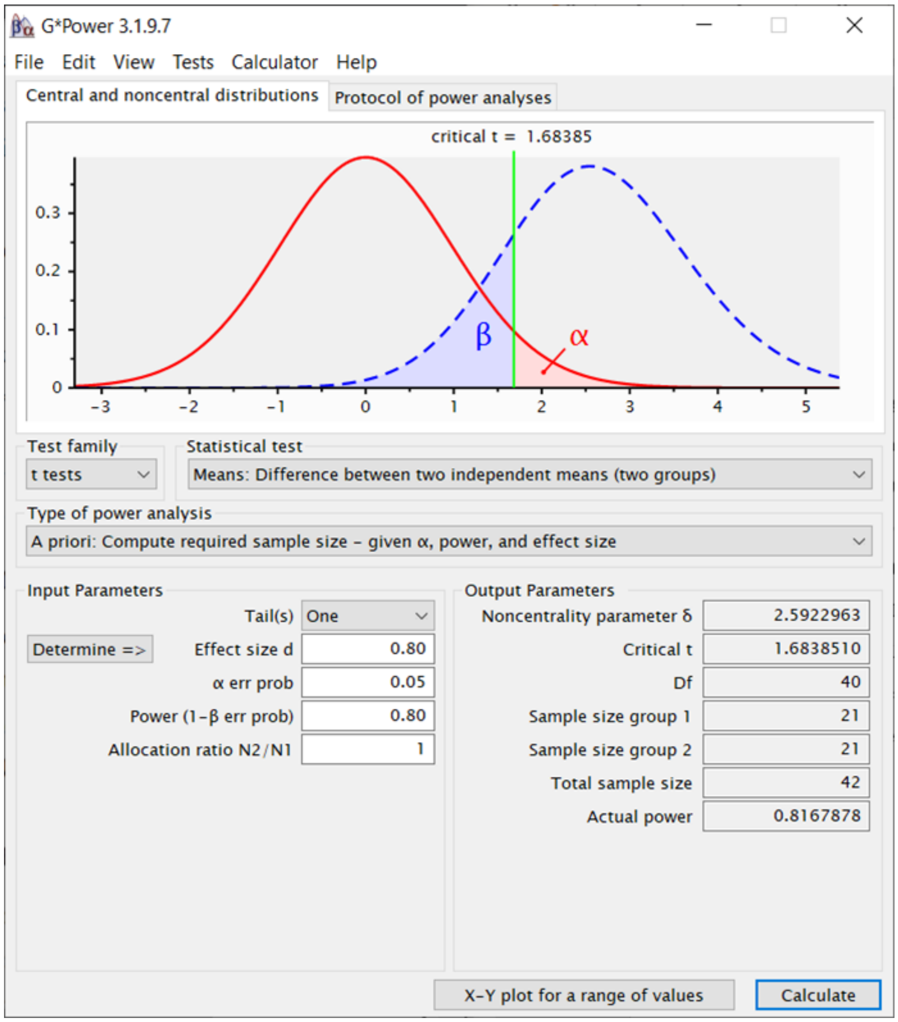


Figure 6. Screenshot of G*Power 3.1.9.7 showing data inputs required for the Power Analysis

The target population of employment of training systems such as those of an interactive synthetic environment were infantry squad leaders. In an effort to facilitate the maximum number of participants possible, the definition of infantry squad leader was expanded to include infantry Marines serving in leadership positions between the rank of Lance Corporal through Sergeant. Figure 7 provides the current standard for a Marine infantry squad leader being an 0311 Sergeant (E-5) as well as the task organization for a Marine Rifle Platoon.

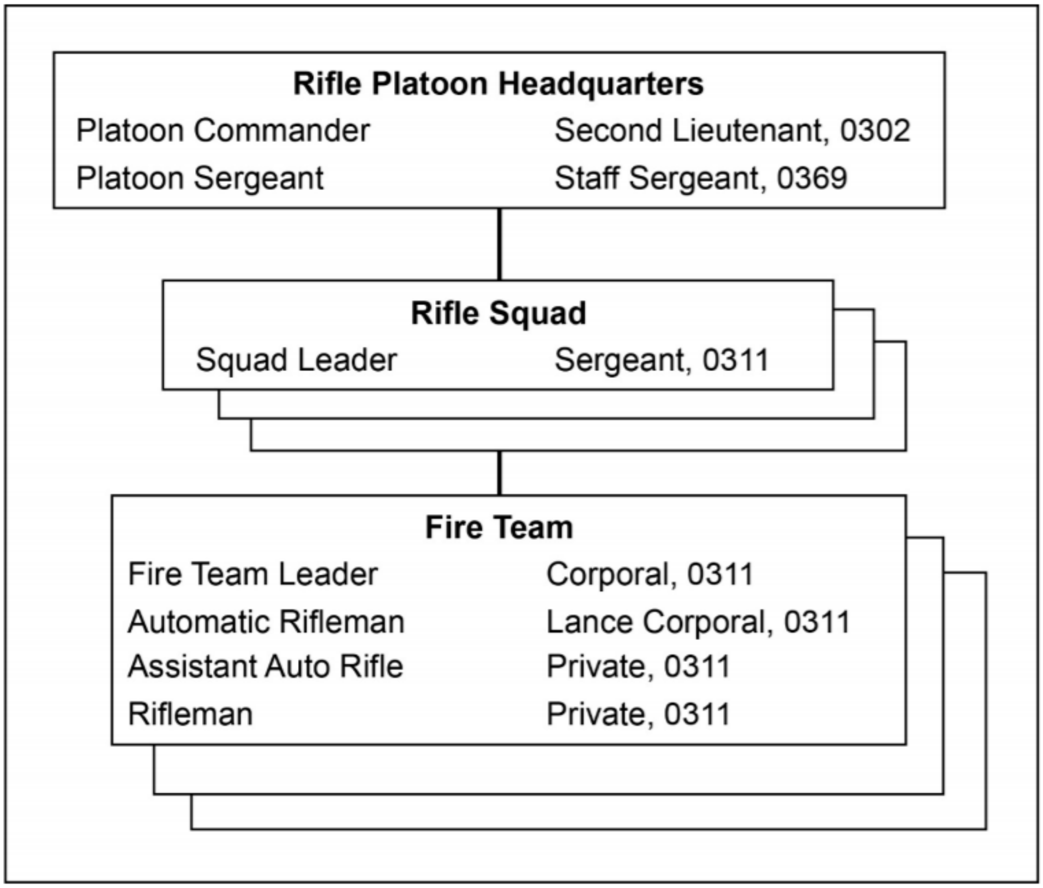


Figure 7. Organization of the Marine Rifle Platoon. Source: [25].

However, included in this study were also 0352 Anti-tank Missilemen, 0341 Mortarmen, and 0313 Light Armored Vehicle Crewmen, all of whom have similar billet descriptions to a rifle squad leader and could benefit from communication, CASEVAC, call for fire, and UAS control training in a similar manner to that of an 0311 squad leader. The exception was made to include Lance Corporals and Corporals in the moniker of squad leader as those ranks often have similar duties as fire team leaders or also serve as squad leaders although they are not yet the rank of sergeant.

While the sergeant as the rifle squad leader has been the standard organization since World War II [26], there was recent guidance and consideration given to having squads led by Staff Sergeants (E-6) by 2030 [27]. Such a statement was made by the Commandant of the Marine Corps, General Berger in his 2021 Update to Force Design 2030:

> The Division CGs, ICW M&RA and DC PP&O, will develop options for improving and sustaining the quality, maturity, and experience of small unit leader tactical skills and decision-making along with a pathway toward ensuring each squad or small unit within the infantry and reconnaissance communities is led by a Staff Sergeant. [27]

There was careful consideration given to identifying the population for the study, but with a such a limited number of Staff Sergeants available in an infantry battalion such as 2nd LAR, it would have been difficult to achieve a sample size of even half of that required by the power study (20 of the required 42 participants).

To meet the target population of Infantry Marines, participants were Active-Duty Marines at the 2nd Light Armored Reconnaissance Battalion in Camp Lejeune, North Carolina. All participants were infantry Marines from the rank of Lance Corporal to Sergeant and experienced some sort of leadership role in their time in the Marine Corps. The population was all male as there were no females who volunteered for the study.

The live execution environment was in a classroom setting in the 2nd LAR battalion area. The classroom was set up as the 2d LAR DVTE room, and included two DVTE suites, a projector, a screen, and other hardware to facilitate simulation learning. Of note, 2d LAR is the only infantry battalion in Camp Lejeune to have such a dedicated space on-site.

C. MATERIALS

1. Workstation, which is visualized in Figure 8.

The following hardware and software were used to create the simulation workstation:

- three Dell USMC issued DVTE Laptops
- one Gazepoint GP2 Eye Tracker

- three headsets with microphone
- one Canon 70D Camera
- router
- call for fire and CASEVAC templates from the USMC Tactical Handbook for Unit Leaders (THULS) (see Appendix D)
- one Empatica E4 wrist monitor
- VBS3
- CNR Radio
- GRG (see Appendix C)

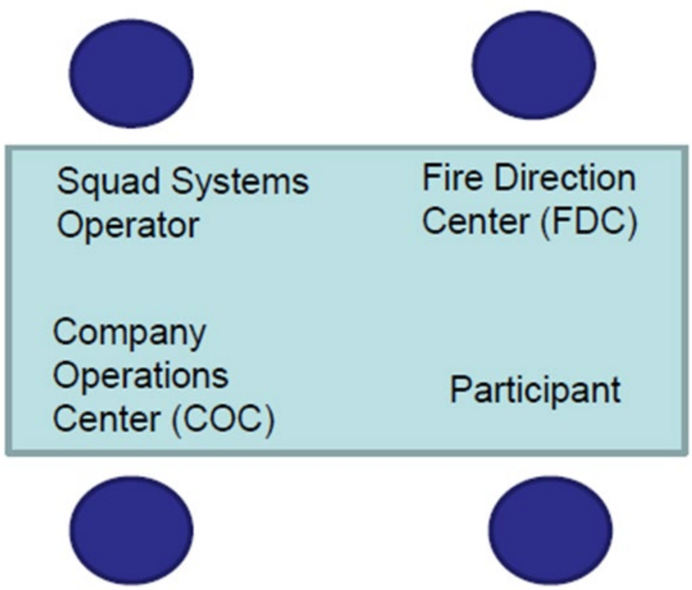


Figure 8. Schematic of the experiment workstation

D. PROCEDURE

1. Virtual Training Environment Scenario 1

A copy of the script used in the procedure can be found in Appendix B. As a program of record for the Marine Corps, Virtual Battlespace 3 (VBS3) was used as the interactive synthetic environment for the experiment. VBS3 has built in tutorials for player movement and controls as well as for weapons handling within the simulation. The first scenario was based on a combination of the provided tutorials, and some injects created for the purposes of this experiment. Participants were instructed to familiarize themselves with the controls of the ISE during this time and followed step by step instructions to experience a practical application of the controls that were to be used for the duration of the exercise. Figure 9 shows the controls mapping of the general infantry controls in VBS3.

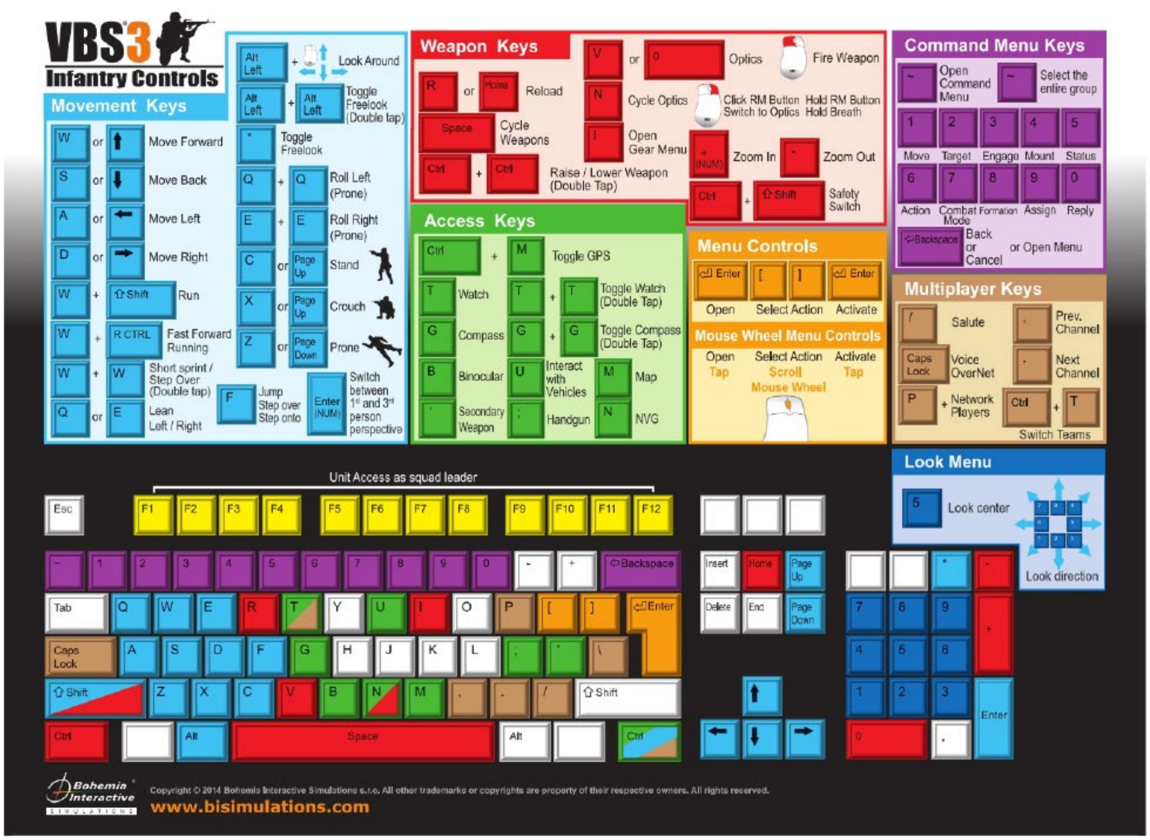


Figure 9. Mapping of the VBS3 Infantry Controls [28]

Participants were required to perform the following tasks within the tutorial:

- move forward, backward, left, and right
- roll left and right
- lean left and right
- walk, sprint, and run
- zoom in and out
- toggle from first to third person

A second tutorial gave the participants background on weapons handling within the simulation. Such instructions included:

- identifying weapons inventory
- choosing among weapons systems
- turning the safety on and off
- firing the weapon
- using the optic

Both tutorials were intended to provide familiarization with the basic controls and visuals of the scenario interface. The verbiage and the tutorials were the same for each group.

2. Virtual Training Environment Scenario 2

Scenario 2 of the exercise provided the basis of training to determine whether the ISE could decrease cognitive workload. Scenario 2 involved familiarization with a Vector-DAGR, throwing smoke, and using the UAS feed. The first control that was introduced was the Vector-DAGR, which allowed participants to laze a target on the map and receive information regarding the target such as a ten-digit MGRS grid, elevation, distance, and direction. Participants were then instructed on how to use the UAS feed within the ISE. They were able to lock a target, laze a target and identify grid information, alter the screen from white hot to black hot, and zoom in and out. Figure 10 offers a screenshot of the UAS feed given to the squad leader for the within the simulation. It offers a black and white picture with heat signatures shown as either black hot or white hot. Lastly, the scenario required participants to use the DAGR to identify the MGRS grid of their position, call a CASEVAC nine line and throw a red smoke. They were able to see the CASEVAC helicopter land as they called in the nine line. The control group was given the above training with no interaction with the squad systems operator. They were told that they had a squad systems operator at their disposal, however they did not have any interaction with the squad systems operator within the second training scenario. The experimental group on the other hand were told how to interact with the squad systems operator and were showed the capabilities of the squad systems operator and how they could assist in UAS operations.

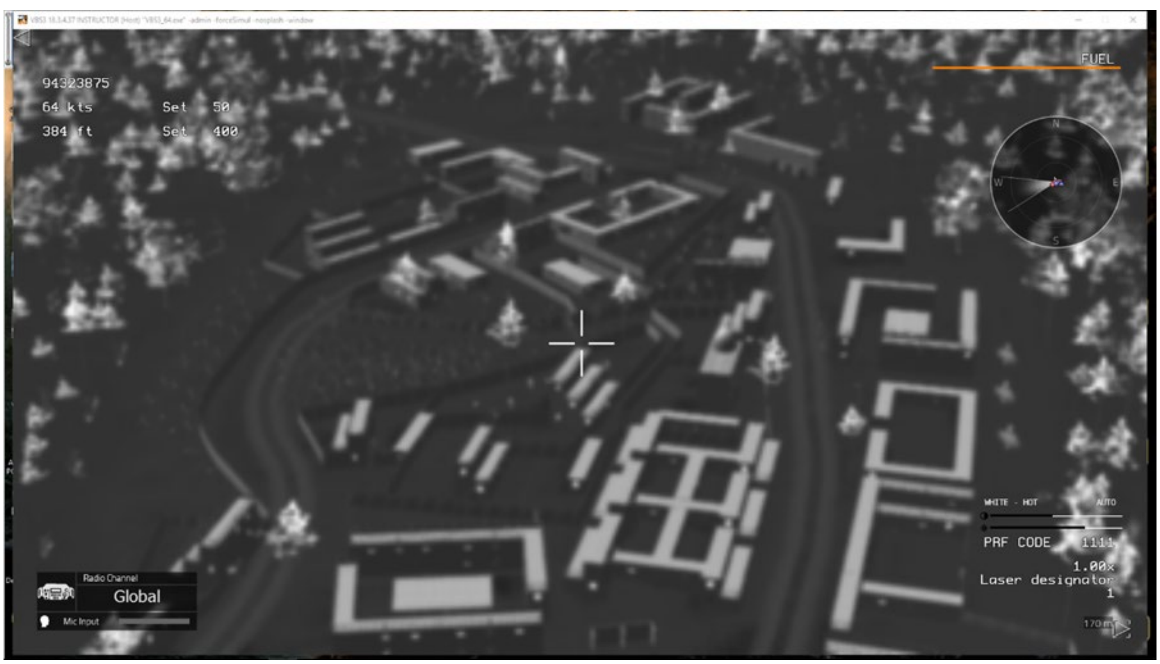


Figure 10. Screenshot of the UAS feed as viewed by the Squad Leader

3. Virtual Training Environment Scenario 3

The third and final scenario included a graphical rendering of the Mobile MOUT facility in Camp Lejeune, NC. Participants were offered a gridded reference graphic (GRG) with building numbers and military grid reference system (MGRS) grids which correlated exactly to the building numbers and grids in the scenario, as displayed in Figure 11. Participants started on the third story of Building 64 and were briefed on the enemy situation in the town.

On the left is the actual GRG of Mobile MOUT in Camp Lejeune, NC with building numbers. The right is a screenshot of the opening vantage point of the participant for scenario 3.

Figure 11. Map and image generated by VBS3

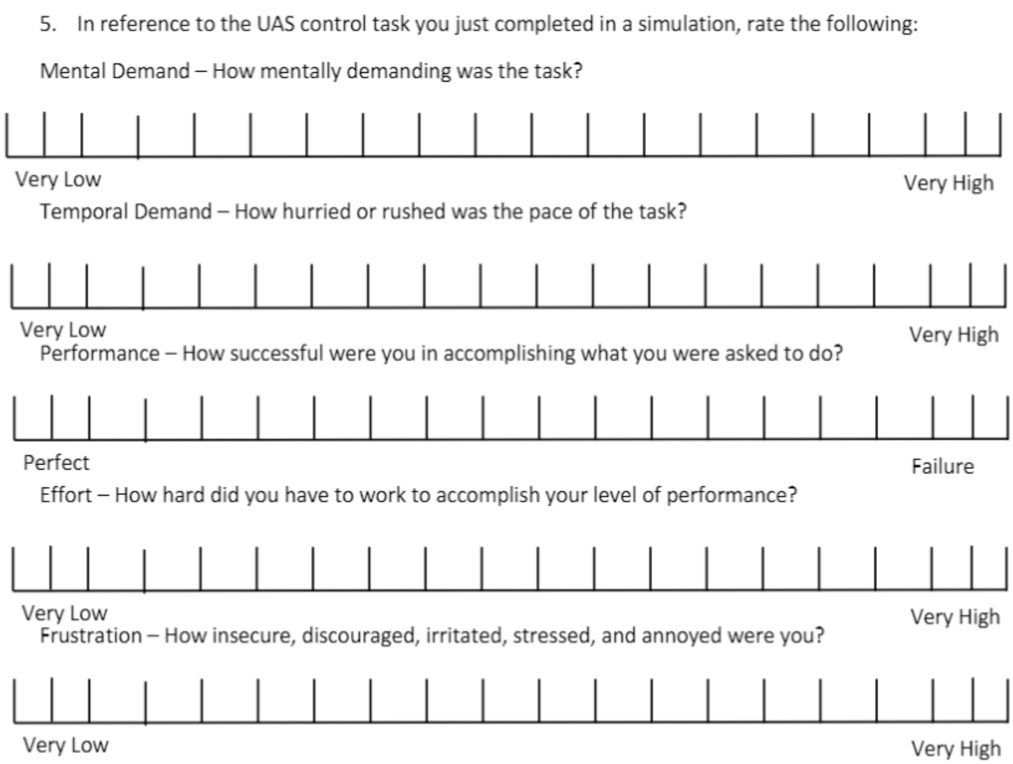
The actual enemy layout included:

- four dismounts and one mounted vehicle in a blocking position with a gate between building 17 and building 3
- three dismounts with a mortar system in the courtyard of building 11
- one mounted vehicle which spawned in the road near building 5 and would drive east across the city
- one enemy dismount in building 38 oriented towards the participant
- a platoon of dismounts (16) with 4 mounted vehicles and guard towers 300 meters southwest of building 19.

The participant spawned with a UAS operator and RQ-11 Raven overhead, as well as with 8 LAVs positioned along the road directly to his east.

2. Cognitive Workload Questionnaire

The cognitive workload questionnaire used in the study was based on the NASA-TLX questionnaire, which can be viewed in the appendix. Figure 12 is an example of the written cognitive workload questionnaire that was asked for each of the tasks of UAS control, CASEVAC, and call for fire.


5. In reference to the UAS control task you just completed in a simulation, rate the following:

Mental Demand – How mentally demanding was the task?

Very Low — Very High

Temporal Demand – How hurried or rushed was the pace of the task?

Very Low — Very High

Performance – How successful were you in accomplishing what you were asked to do?

Perfect — Failure

Effort – How hard did you have to work to accomplish your level of performance?

Very Low — Very High

Frustration – How insecure, discouraged, irritated, stressed, and annoyed were you?

Very Low — Very High

Figure 12. Cognitive workload questionnaire

IV. ANALYSIS OF RESULTS

A. HYPOTHESIS

- H_1: The Marines with increased training in an interactive synthetic environment will decrease the cognitive workload over those with less training in the interactive synthetic environment ($p_{Trained} - p_{Untrained} > 0$).

This hypothesis was chosen to determine how well the training conducted within the ISE translated to decreased cognitive workload. The aim for this hypothesis was to understand how observable factors could influence the cognitive workload gained (or lost) when using the ISE.

While there are several factors that may indicate variation in cognitive workload, the relative size of the pupil and its changes in constriction or dilation has been an established direct indicator of cognitive workload [29]. The NASA TLX survey was specifically designed to directly measure cognitive workload as a subjective measure [22]. Other biometric measures are indirectly measures of cognitive workload as indicators of increased stress. One such measure is Electrodermal Activity (EDA), which can be both a measure of stress and cognitive workload and there was evidence in [30] that it may be possible to discriminate between instances of stress and cognitive workload with EDA measurement. An indirect measure of cognitive workload but a well-established measurement of stress is heart rate variability and the inter-beat interval (IBI), which is the average amount of time between heartbeats [31]. Although a primary measure of stress, there has been research on the ability of cardiac measures such as IBI to be sensitive to changes in cognitive workload [32]. Performance also serves as a measure of cognitive workload. The following factors were chosen as indirect and direct measures of cognitive workload:

1. Difference in Left and Right Pupil Diameter – Data point was produced via Gazepoint Analysis Standard Edition (v6.5.0).
2. Difference in EDA – Data point was produced by the E4 connect program and collected via the E4 Empatica wristband.

3. Difference in IBI – Data point was produced by the E4 connect program and collected via the E4 Empatica wristband.

4. Difference in TLX Survey results – Data point was collected by the TLX survey as mentioned in Chapter III.

5. Performance of task completion – Data point was evaluated by the exercise controllers.

1. Statistical Analysis

The various aforementioned factors were used to support the hypothesis and were collected with various devices but all under the same conditions. Although each collection method was the same for each participant, a tremendous amount of data preprocessing had to be accomplished to account for outliers and other issues with the data. Individual data preprocessing comments will follow each of the sections below, however there were some overarching factors such as the ones mentioned in Table 2 as part of the survey responses.

Table 2. Survey response questions regarding experience and time in service of participants

	Years in the Infantry	Years in a Leadership Position	Perceived CFF Expertise (on a 1–5 scale
Experimental Group	2.76years (approx. 2 years and 8 months)	2.032 (approximately 2 years)	2.389 (approximately 2 years and 5 months)
Control Group	2.35 (approximately 2 years and 4 months	1.14 (approximately 1 year and 1 month)	2.438 (approximately 2 years and 5 months)

While the experimental group reported an average of 2 years in a leadership position, the control group reported an average of just over 13 months in a leadership billet. Although the groups were randomly selected, the control group seemed to have far less self-reported leadership experience than the experimental group. Although such a factor could normally be very detrimental to a study, in this case the leadership experience might be as an 0313 Light Armored Reconnaissance Marine and would differ greatly from the experience of an 0311 Rifleman serving as an LAR scout or as an 0341 Mortarman. The reason for such a difference is that a Marine with an MOS of 0341 Mortarman would have tremendously greater amount of experience and expertise when calling for fire than an 0311 Rifleman or 0313 LAR Marine. Additionally, the training of such Marines varied from billet to billet, and while all Marines were still within the infantry community, their individual aptitudes as leaders varied greatly depending on their training focus. Marines such as 0313 LAR Marines were typically more focused on vehicle maintenance, operations, and tactics, than LAR scouts or mortarmen. However, such a difference remains anecdotal and would again vary greatly depending on the training of the individual Marine. All Marines who participated in the experiment were from 2d LAR battalion but from different companies across the unit. Therefore, with all factors regarding MOS, time in service, and time in leadership considered, the statistical analysis could still be considered.

Data preprocessing across both the Empatica E4 and Gazepoint results required statistical computer-generated and manual careful sorting of the data. Manual sorting, dissection, and removal of results was necessary as there were a tremendous amount of data points. During an average 50-minute experiment, the Gazepoint dataset contained 187,000 entries across 51 different statistics, for a total of over 9.5 million datapoints per research participant. The Gazepoint software delivered statistics including data relating to point of gaze, the position and state of the mouse, pupil data, and data relating to designated areas of interest on the screen. Similarly, the Empatica E4 data included up to 9,000 individual datapoints for seven parameters, including temperature, blood volume pulse, accelerometer response, interbeat interval (IBI), heartrate, and time tags. Data was parsed manually by checking for obvious errors or egregiously incorrect data. After several participants conducted the experiment, for example, it became obvious that the Gazepoint eye tracker was not properly tracking participants' gazes in the bottom 20–25% of the computer screen. Therefore, it was appropriately taken into consideration that areas of interest in that portion of the screen may return skewed results and were omitted. Another manual parsing of data that needed to be done was the division of data among the three different training scenarios. As seen in Figure 13, each dataset for every participant was carefully deconstructed using the available video and audio recording timestamps to match individual biometric events to a certain task, such as a CFF, UAS control, or CASEVAC task. Timestamps were recorded by both the Gazepoint and Empatica software, and therefore made it easy to match a task to a potential biometric event. Such division was also necessary for the three very different scenarios that each participant was subject to during the experiment. The first two scenarios were training scenarios, and the final scenario was the testing scenario. Manual division of data across each scenario was accomplished using recordings to separate the recorded data across each of the scenario, and such parsing was crucial to understanding how much cognitive load was achieved or added during each phase of the experiment.

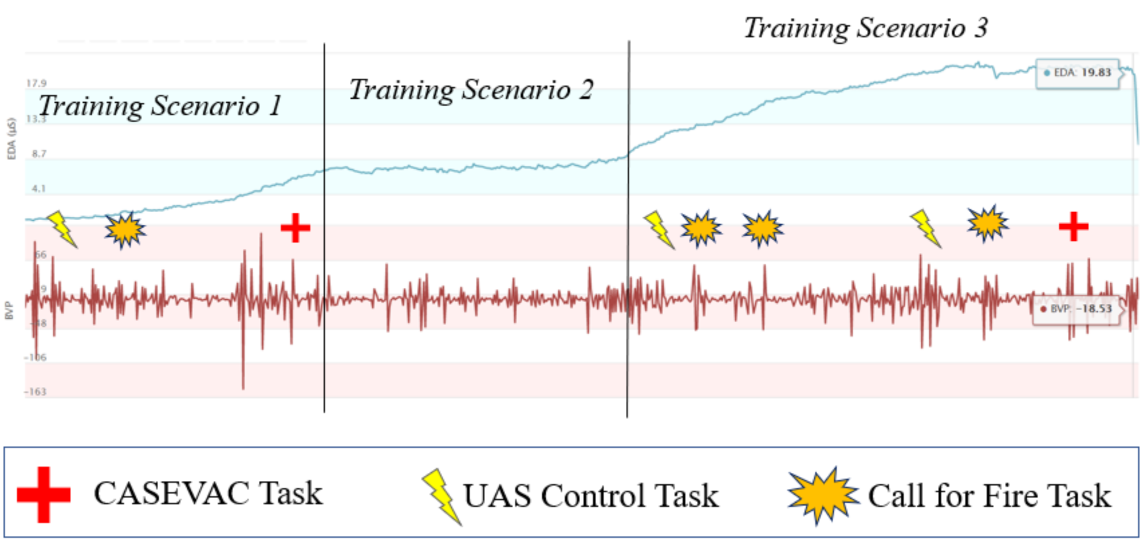


Figure 13. Example of data preprocessing using recording timestamps to determine biometric related events

Another major factor in data preprocessing was the number of outliers across the different datasets recorded. Interquartile range criterion, I = [q0.25−1.5·IQR; q0.75+1.5·IQR], was used to determine outliers across the datasets. Outliers were found across datapoints from both the Empatica E4 and Gazepoint devices. Figure 14 shows an example of outliers for an average datapoint.

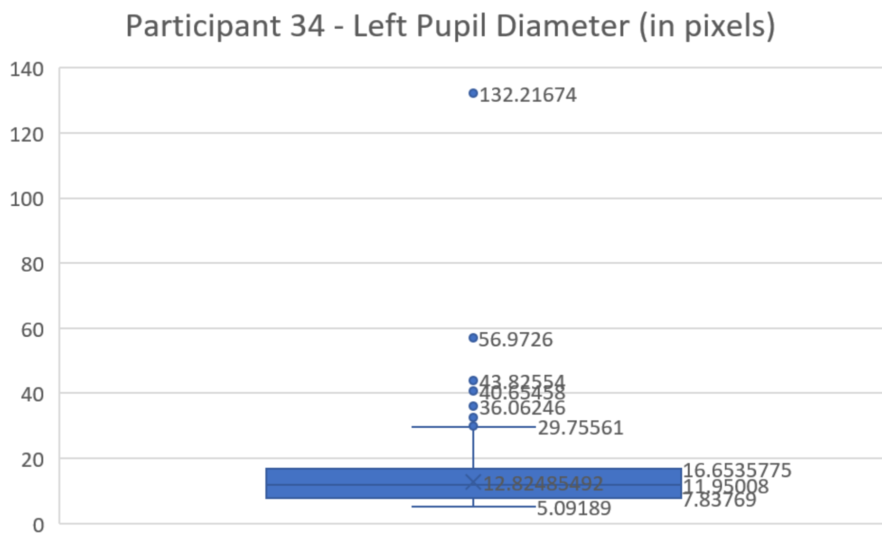


Figure 14. Display of outliers for participant 34 left pupil diameter dataset

In every example of the Gazepoint pupil diameter datasets, each participant had outlier values greater than the longest whisker value of the plot. Only one participant had outlier values less than the lower whisker of the plot. No participant had outliers that consisted of more than .01% of the total dataset, which was approximately 10 outliers per 90,000 data entries. It was concluded that such outliers were minor issues with the Gazepoint system.

a. Difference in Left and Right Pupil Diameter

Because the sample size did not meet the required criteria of the power analysis (only 34 participants conducted the study while 42 were required by the power analysis) and the distributions of the left and right pupil diameter data were not normally distributed, a two-sample t-test could not be used in determining the significance of the difference in pupil diameter for each group. A Wilcoxon signed-rank test was used instead. The result is displayed in Figure 15. An increased pupil diameter difference is an indication of increased cognitive workload. While the experimental group does have higher mean differences in pupil diameter, the result of the Wilcoxon signed-rank test was that the difference was not significant for either the left eye between Group A (control) (M = 25.99, SD = 5.44), to Group B (experimental) (M = 27.87, SD = 5.71), Z = -0.458, p = .322 or the right eye between Group A (control) ($M = 23.75$, $SD = 6.75$), to Group B (experimental) ($M = 26.52$, $SD = 5.49$), $Z = -0.153$, $p = .440$.

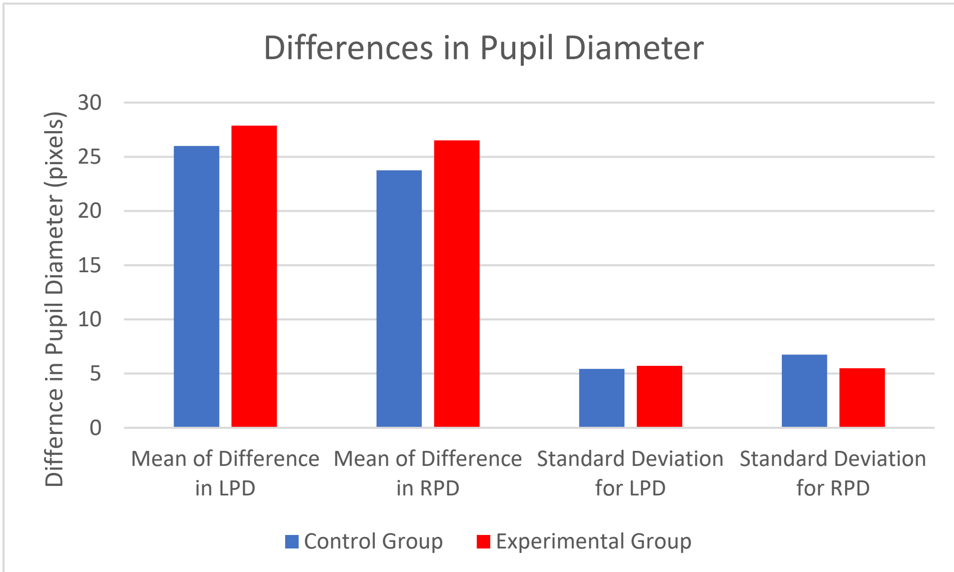


Figure 15. Mean and standard deviation of differences in pupil diameter

b. Difference in Electrodermal Activity (EDA)

Difference in EDA required a significant amount of data preprocessing, as it appeared that the Empatica E4 had a more difficult time registering the EDA correctly and that is reflected by the large standard deviations relative to the means for each group. There were also zero values in each participant's dataset which had to be removed. Figure 16 displays the results of the average difference in EDA for each group, with higher differences being a possible indicator of greater cognitive load. The result of a Wilcoxon signed-rank test was that the difference was not significant between Group A (control) ($M = 6.060$, $SD = 6.834$), and Group B (experimental) ($M = 4.694$, $SD = 5.301$), $Z = -0.517$, $p = .301$.

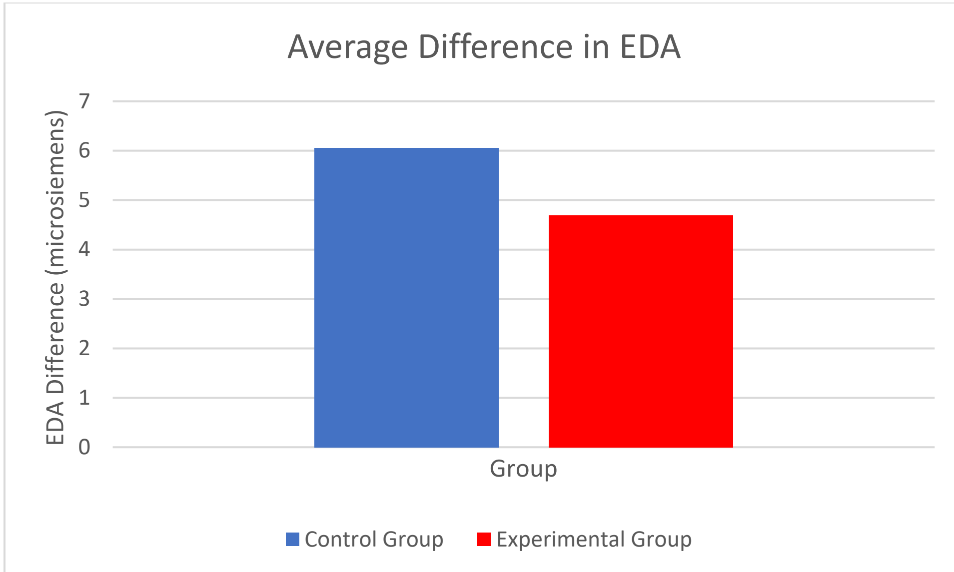


Figure 16. Average difference in EDA

c. Difference in Interbeat Interval (IBI)

The IBI was recorded with a greater level of accuracy with the Empatica E4 than the EDA as evidenced by the smaller standard deviation compared to the mean values in each group. A greater difference in average IBI would indicate different levels of stress, and could have been associated with different levels of cognitive load. Figure 17 displays the results of the average differences. The result of a Wilcoxon signed-rank test was that the difference was not significant between Group A (control) (M = 0.305, SD = 0.092), and Group B (experimental) (M = 0.381, SD = 0.129), Z = -1.3444, p = .090.

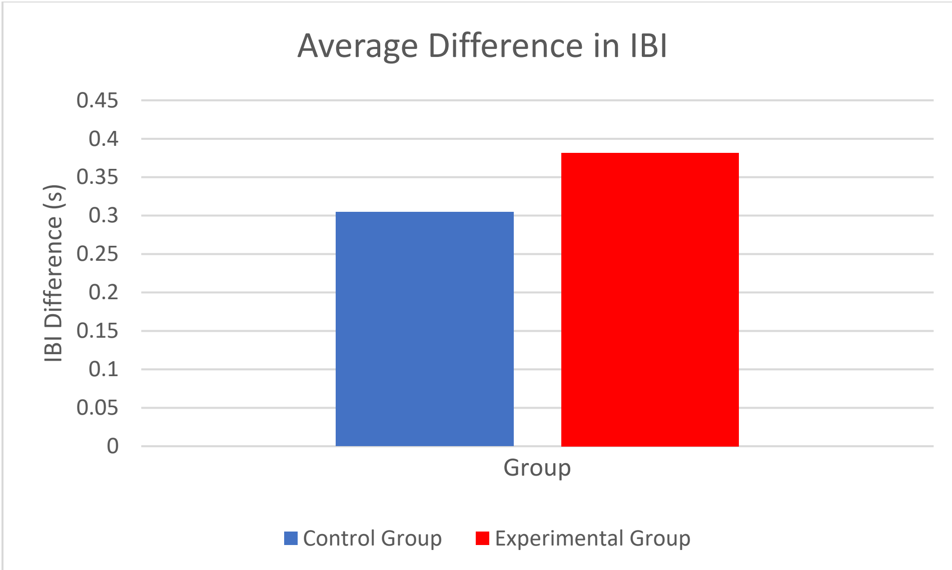


Figure 17. Average difference in IBI

d. Difference in Survey Results

As discussed in Chapter II, perceived workload is important in assessing cognitive load. Figure 18 displays the average survey response by the participants for each group.

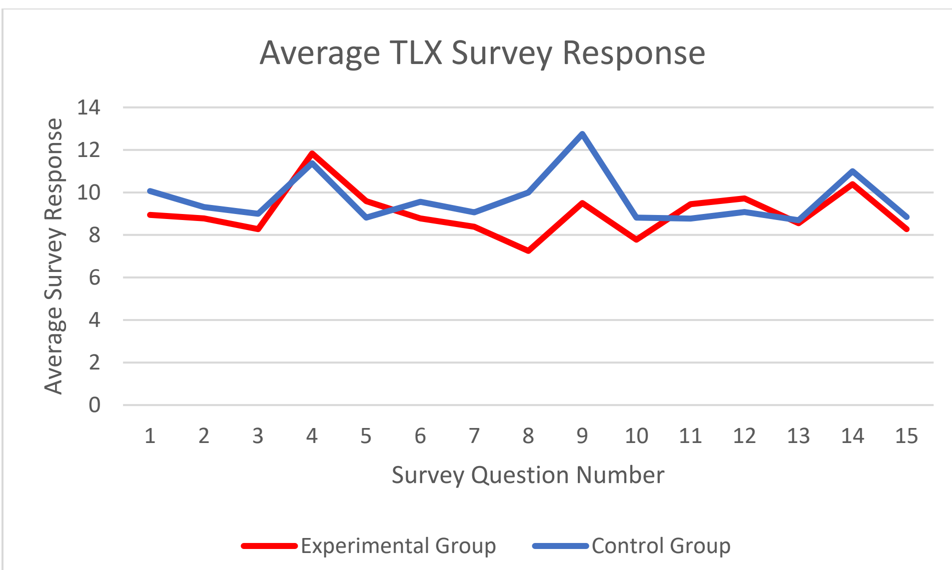


Figure 18. Average TLX survey response by each group

The survey responses are tightly coupled throughout most of the survey questions. However, there was greater variation in questions 8 and 9 than the rest of the survey responses. Questions 1–5 dealt with CFF tasks, questions 6–10 regarded UAS control tasks, and questions 11–15 concerned CASEVAC tasks. Question 8 was labeled "Performance" and was under a section that asked about the UAS control tasks. It asked, "How successful were you in accomplishing what you were asked to do?" Figure 19 offers a detail of questions 8 and 9:

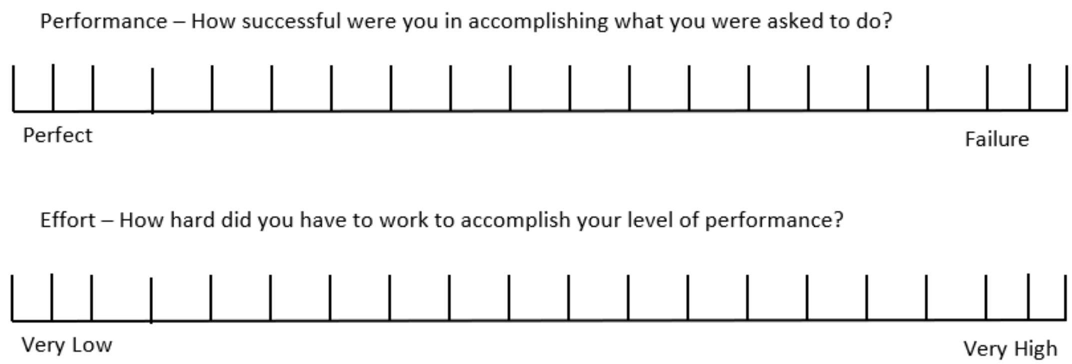


Figure 19. A detailed look at questions 8 and 9 of the TLX survey. The questions were asked in a section regarding the UAS control task.

A lower score on question 8 indicated perfect performance while a higher score indicated failure. In the case of question 8, the control group reported that they thought they were more successful in performing the UAS tasks than in the experimental group. Similarly, question 9 also asked about UAS and was labeled "effort" and asked, "How hard did you have to work to accomplish your level of performance?" A higher score indicated higher effort while a lower score indicated lower amount of effort to achieve the performance. Since the main difference between the two groups was the amount and type of UAS control training, the survey result is the most supportive evidence to the described hypothesis. After data preprocessing, it was concluded that questions 8 and 9 of the survey exhibited the greatest amount of evidence for rejecting the null hypothesis. Due to the limited sample size and distribution, a Wilcoxon signed-rank test was again chosen instead of a two-sample t-test. For survey question 8 (UAS task performance assessment), there was no significant difference between Group A (control) (M = 10, SD = 4.47) and Group

B (experimental) (*M* = 7.25, *SD* = 2.38), *Z* = -1.632, *p* = .051. While the p-value in for question 8 was less than .05 and not considered significant, it had a value less than that of any of the objective biometric measurements.

On the other hand, question 9 (UAS task effort assessment) did indicate a significant result. There was a significant difference between Group A (control) (*M* = 12.75, *SD* = 3.19) and Group B (experimental) (*M* = 9.5, *SD* = 4.09), *Z* = -2.228, *p* = .012. This is the most fitting significant statistic to the study, as the study aimed to decrease cognitive workload over the course of the two training sessions. In the third session, the participants from the experimental group, who were given the additional UAS training, said they took a significantly less amount of effort than the control group in executing the UAS tasks. This could mean that they in fact did have a decreased cognitive workload and such a statistic gave the greatest evidence to reject the null hypothesis.

There are undoubtedly other factors than the training style or implementation of the simulation that could have had such an effect on the perceived effort of the participants to execute their UAS tasks. There was a self-reported difference in time in a leadership position. But other variables such as experience with UAS or even time on a radio could have played a factor. However, experience with radio communications could also be associated with expertise with CFF, to which participants self-reported an almost equivalent amount of expertise between the two groups. A question that was not asked but could have been included in the survey was "What is your experience with UAS?" That question was not asked because it was assumed that there was little experience with UAS within the unit, and in speaking with participants only one of the 34 participants claimed to have any UAS controller experience.

e. Performance

The performance metrics showed no major difference between the two groups. Performance was based on task completion of the three major required tasks in the scenario. Figure 20 details that all participants were able to execute a successful CFF. All but two of the participants were able to correctly request a CASEVAC.

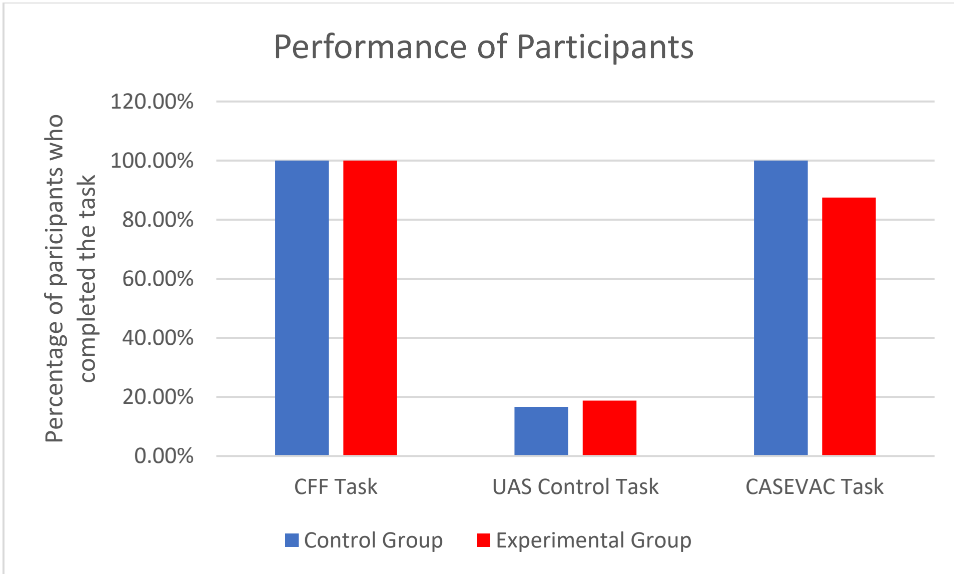


Figure 20. Comparison of percentage of participants in each group able to achieve success on their tasks

Contrary to the overwhelming success of the CFF and CASEVAC tasks, the participants were not as successful in the UAS control task. Both the experimental and control groups only had three participants employ the squad systems operator and request to change the position of the UAS to better complete their reconnoiter mission.

The result may have suggested that there was an insignificant difference in the quality of cognitive workload decrease between the two groups. Concurrently, it also may imply that more training and research is needed in regard to employment of the squad systems operator by USMC infantry squad leaders.

B. SUMMARY

The hypothesis was aimed at assessing the ability of an ISE to decrease the cognitive workload of those using it to train. Subjective and objective measures were used as evidence for rejecting or failing to reject the null hypothesis. The objective measurements were difference in pupil diameter in the left and right eyes, difference in IBI, and the difference in EDA. The subjective measures taken into account were the TLX survey results. One of the measurements, question nine from the TLX survey regarding the effort used to complete the UAS tasks had a significant difference between the control and experimental groups. None of the other questions or objective measurements concluded with significant differences between the two groups.

The significance of the effort question in UAS task performance could be explained by other variables:

- The training received by the participants was completely independent of the ISE performance.
- The participants had prior UAS experience that led them to use less cognitive workload in executing UAS tasks.
- The participants had greater leadership experience and therefore were better suited to handle UAS tasking.
- The participants perceived that they performed better and used less effort to execute UAS tasks because there was a greater focus on UAS tasking during the ISE training scenario.

With such variables and others in mind, the research could still point to use of the ISE as the critical factor as to why the experimental group self-assessed their effort to be eased and performance greater for the UAS control tasks.

V. DISCUSSION, LIMITATIONS, RECOMMENDATIONS, FUTURE WORK, AND CONCLUSIONS

A. DISCUSSION

1. Use of ISEs to Improve Cognitive Workload Performance

The overarching goal of this research was to determine if an ISE could decrease the cognitive workload of a USMC infantry squad leader. The Paas subjective rating scale found in [21] as well as the NASA TLX research conducted in [22] placed high importance on the subjective measures of cognitive workload performance. Such measures can have a high amount of bias as they are self-reported, but this research found that they produced the most significant difference between the experimental group and the control group. They could also be used as the best evidence to reject the null hypothesis and support the stated hypothesis.

While none of the objective measures exhibited a significant difference between the two groups, there was still tremendous knowledge value added from the performance indicators and other survey results reported by the participants. The comments from the participants collected during this study can have great impact on future work concerning ISE design and training.

2. Simulations Treated as Training or Games

13 of the 34 participants mentioned training in responding to the survey questions. Much of the time and effort building ISEs is dedicated to making the graphics and interface look and feel as realistic as possible. The differences between VBS3 and VBS4 are subtle other than the increased graphics.

During a visit to NPS by Sergeant Major of the Marine Corps Troy Black in March 2021, the question was asked, "How do we get Marines to want to play [VBS]?" His sentiment was not unique and many other warfighting professionals have demanded the answer to that question. Many Marines may return to their barracks rooms after a day of work to play video games such as Call of Duty: Modern Warfare. They willingly play such games in garrison and while deployed and have dedicated countless hours to gaming. Figure 21 shows the similarity between the two interfaces of VBS3 and Call of Duty.

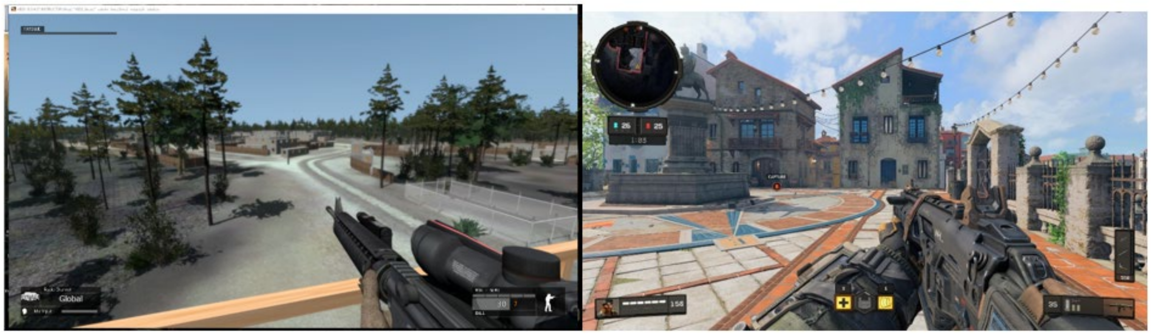


Figure 21. A VBS3 screenshot on the left and a Call of Duty: Black Ops 4 screenshot on the right. The two interfaces are nearly indistinguishable as first-person shooter viewpoints except for the more detailed graphics in Call of Duty

Instead of focusing on how to make the ISE more realistic, the focus must be on the training and what can be achieved within the ISE. Marines expressed that they liked the training during the study and would like to do it more often. Participants reported that insinuating that their training could be likened and reduced to a video game was taken as an insult by the enlisted infantry community.

The knowledge value added of a first-person shooter is very little and from comments by the participants the value seems to be more about radio and communication training. Therefore, ISEs do not have to be a first-person shooter or replicate one to be an effective training tool. They should not focus on first-person shooter interfaces as they cannot do so very effectively or compete with COTS industry standard games. Flight simulators are ubiquitous as simulations because there is an interface between the pilot and the real world. On the other hand, there is no interface between the infantry Marine and the world, other than either the rifle scope, radio, or a tablet. In such an example, ISEs should not focus on first-person shooter type interfaces as they do not effectively train for marksmanship and the USMC already has a marksmanship simulator (indoor simulated for marksmanship trainer or ISMT). Instead, it can be concluded that the radio is the best interface for infantry squad leader training within an ISE and that it can be used to increase the cognitive workload capability of infantry Marines especially in regard to communications training. Such training was used within the ISE for this research, and the participant never had to use the ISE as a first-person shooter. Instead, the ISE was used as a platform for communicating with the FDC, the COC, and the squad systems operator.

3. Accessibility of Training Simulations

The term “digital native” describes a person who has grown up in the digital age of technology with devices such as the internet, tablets, smartphones, and was first coined in a 1996 paper by John Perry Barlow [33]. Most Marines from the rank of Lance Corporal to Sergeant are considered digital natives by virtue of being born with the technologies such as internet and smartphones. This was also assumed in the design of the experiment used in this study. Along with the technologies mentioned, another factor that makes a digital native is that the person was in close contact with video games in addition to computers and the internet.

There is a difference between digital natives and those who can play computer games. A major issue with ISEs such as VBS3 is that they have a learning curve that can be frustrating due to a presumption that the user has some experience with video games and “ASWD” controls. The preponderance of survey results, 14 out of the 34 participants, complained specifically about the control scheme within VBS3. Participants became easily confused and frustrated even after a training scenario dedicated to the user interface and simple tasks such as moving the avatar and manipulating the weapon. While such frustrations may have come from poor user interface, ISEs could benefit from less controls and keys to remember in an effort to decrease the extraneous cognitive load placed on the trainee.

4. Training of Squad Leaders and the Squad Systems Operator

In 2018, then-Commandant of the USMC, General Robert B. Neller introduced several decisions to improve the technology, mobility, and firepower capabilities of the ground combat element, and one such capability was the addition of the squad systems operator to the rifle squad [34]. The role was never formalized but most likely introduced to alleviate some of the cognitive workload faced by the squad leader in tasks such as those tested in this research. Squad leaders in General Neller's new operating concept had to be able to balance their traditional roles of the tactics, techniques, and procedures of leading a rifle squad with the new responsibilities of using UAS, tablets, and other new technology. To assist with the new technologies, the squad systems operator would, "receive formal training on a variety of technologies" [34]. General Neller explained that each squad would have its own UAS (to be operated by the squad systems operator) and each company would also have counter-UAS capabilities [34].

Although yet to be implemented, such a concept paired well with this research. It was an insightful response to the added burden of so many technologies to the infantry gear set. There was little training or formal emphasis on the employment of the squad systems operator. Additionally, there are no formalized reports on how to communicate reconnaissance requests to a UAS pilot, unlike the formalized reports of CFF and CASEVAC.

A peculiarity of this research is based upon the performance statistic that only six participants of the 34 ever communicated or utilized the squad systems operator during the study. Each of them was forced to communicate with the COC or FDC by virtue of completing the CFF or CASEVAC tasks. The UAS control task allowed for different options for completion, one of which was that the squad leader in the study did not have to use the squad systems operator at all. Instead, the squad leader was able to view the UAS image feed but did not ask to move the UAS or ask the squad systems operator for assistance in gaining situational awareness. This may suggest that more training is needed for squad leaders to employ the squad system operator to achieve the full potential of the billet and the Marine serving in it.

B. LIMITATIONS

There were both technical as well as exercise execution limitations that impacted the results and implementation of the study. Such limitations were outside of the control of the experiment design or that of the researchers.

1. Technical Limitations

a. Gazepoint

As previously mentioned, the Gazepoint hardware had some issues correctly tracking the participant when the participant looked towards the bottom of the screen. Even with a successful initial calibration, the Gazepoint analysis would still not correctly track the participant's eyes towards the bottom of the screen. Since the main area of interest was towards the bottom of the screen, the analysis from that area of interest could not be included in the results.

Users also voiced frustration in the calibration process. While a majority of users were able to achieve an acceptable calibration upon first or second try (an acceptable calibration for the Gazepoint software is achieving a match of at least four out of five points of gaze to actual pupil position with both eyes), many participants found it difficult to calibrate the Gazepoint hardware and therefore became frustrated. This may have compounded further frustrations and skewed results within the experiment since the Gazepoint calibration occurred prior to any training scenario within the study. Several participants needed more than five calibration attempts to correctly calibrate the Gazepoint device.

b. E4 Empatica

The E4 Empatica device performed extremely well with few limitations. It accurately recorded a number of necessary variables and did so with a seemingly high degree of fidelity and few outliers. It did not have to be recharged often and could last through an entire day of testing. It reliably synchronized to the computer to record results and always recorded immediately.

The drawback to the E4 Empatica was that an internet connection was required to upload results and they had to be uploaded to the cloud. A participant's results could not be uploaded directly to a computer. While it was not necessary to see the results in real time, it was helpful to look at results following a participant for several reasons:

- To ensure the participant's data was recorded correctly

- To ensure that the device was functioning properly
- To view the device's charge level
- To ensure transparency during the informed consent process and display a participant's results if they desired to see them following the study.

While in the 2nd LAR Battalion area, the cellular service was adequate and the researcher's cell phone was used to create a mobile hotspot to execute the above listed tasks. However, in a USMC field environment or without Wi-Fi, the Empatica device would have been limited and could have had major issues in losing participant data over the course of the study.

c. VBS

The VBS software provided some of the greatest challenges and limitations during the study. A primary challenge was using VBS4 and VBS3. VBS4 version 20.1.3 was used in creating some parts of the study. However, it was found by the author that VBS4 was very limited. While it boasted better graphics and interoperability than VBS3, it lacked in many technical areas.

The computers using VBS4 had 16GB of RAM or more and should have been able to handle the graphics workload of such a program. However, during scenario execution or upon exiting a scenario, VBS4 would crash and had to be closed with the Windows Task Manager. It also experienced crashing whenever code was written into the scenario. A common code inject for a VBS scenario is to allow or prevent an object from being damaged. The description of the function from the VBS3 and VBS4 manual was the same: "Allow or prevent an object being damaged (or injured, or killed)" [28]. Additionally, the syntax for the function was listed as "object allowDamage allow," with the third argument being a Boolean value of either true or false. If the third argument in the function was set to "true," then the object provided in the first argument would be allowed to receive damage. Concurrently, if the third argument was set to "false," then the designated object in first argument would not receive damage. Such was code was necessary for the squad leader avatar in the scenario, and the allowDamage function was used to prevent the squad leader from receiving damage, “dying,” and ending the scenario before the study constraints were concluded. Each time the code was executed in VBS4, the scenario would start with an error message across the top of the screen, indicating that the allowDamage function somehow caused an error within the execution of the scenario script. The scenario would then crash.

Another major frustration in using VBS4 was the change in how maps were saved in VBS3 to VBS4. The file extension used and therefore the file type was different between the two versions. To compensate for the file change, VBS4 comes with a map convertor tool. However, any attempt to convert maps from VBS3 to VBS4 was met with errors and incorrectly converted maps. The VBS4 map feature is very advanced with a complete mapping of the landscape of the world offered upon initial start of the application. However, if the inability to fully convert maps from VBS3 to VBS4 persists, it will be labor intensive to fully recreate all of the VBS3 maps used by the USMC in VBS4.

A final major issue with the version of VBS4 being used was that scenarios would not save properly. After creating a scenario in the scenario editor, it would be saved, and the scenario and program would exit properly. However, when reopening that scenario at a later time, one or more of the objects in the scenario would no longer be there. The objects would, on the other hand, reappear once the scenario was being executed.

VBS3 did not have nearly as many problems as VBS4, but the software was a major complaint of survey responses. Four of the 34 participants complained about issues with the ISE such as “lag” or “glitches.” VBS3 would freeze and crash upon exiting almost every scenario, but it was able to be quickly relaunched for the conduct of the experiment. VBS3 issues could also not be solely attributed to the software but also the hardware and DVTE computers being used. They may not have had the correct graphics cards or RAM for optimal performance.

Another issue with VBS3 was the intercommunication system that was offered. It would either not work or come across as static. This was an issue in both VBS3 and VBS4, however there were multiple variables such as firewalls and machine hardware that could have caused such issues. Instead of using the communication software internal to VBS, another software was used that was also installed on the USMC DVTE laptops. Calytrix Comm Net Radio (CNR), was instead used and never had any issues while operating concurrently to VBS3 on the DVTE laptops. It allowed seamless communication among the participant and the other roles within the study.

2. Study Execution

The study was coordinated with the 2nd LAR Battalion staff and due to the operational demands of the battalion not all participants needed were able to be provided by the various companies in the battalion. Additionally, with one company on a restriction of movement due to COVID-19 protocols another company deployed at the time of the experiment execution, the number of available participants who held an infantry MOS and were the rank of E3-E5 was greatly limited. Therefore, only 34 of the desired 50 participants were able to complete the study, or an amount of 68%. The lack of participants influenced the study as mentioned previously in the type and quality of analysis that could be conducted. It does nevertheless enable an opportunity for future research.

C. RECOMMENDATIONS

(1) The USMC should include more instruction on ISEs in the Infantry T&R Manual

The Infantry T&R Manual is used within the USMC as the “primary tool for planning, conducting and evaluating training, and assessing training readiness” [35]. The manual allows for simulations to replace live training, and even encourages commanders to use simulations. It states:

> Training simulation capabilities offer an opportunity to build and sustain proficiency while achieving and/or maintaining certain economies. Commanders should take into consideration simulation tools as a matter of course when designing training. [35]

Simulators listed within the T&R Manual include:

- MAGTF Tactical Warfare Simulation
- Combined Arms Command and Control Training Upgrade System

- Instrumentation and Tactical Engagement Simulation System
- IIT
- DVTE
- TDK
- Combat Convoy Simulator
- ISMT
- Supporting Arms Virtual Trainer
- Heavy Anti-Armor Simulator System

The current version of the T&R Manual was released on 07 May 2020 and unfortunately includes the TDK system as a suitable simulation for training when the TDK is no longer available to most units. Additionally, the Manual was updated to include the DVTE system, but it also lists outdated software no longer included in the DVTE system to include VBS2. Important individual squad leader tasks such as "Lead a Patrol" are do not list TDK or DVTE as a suitable simulation for training.

The recommendation is the Infantry T&R Manual becomes more precise in its language and direction of how simulations are used. Instead of listing the DVTE system as a suitable trainer, it needs to direct Marines towards what simulation to use (most likely VBS3 or VBS4 in the future) and which scenario they can use to train. A second recommendation is to take a refined look at which tasks can be trained with which simulation. These recommendations would formalize ISE training within USMC doctrine and only help commanders conduct training with such systems.

(2) The USMC should create a curriculum for the DVTE

Another recommendation from this research is that the USMC needs to create a training curriculum for the DVTE system and specifically for use of VBS3. Systems like the TDK were divested because Marines failed to use them and failed to know how to use them. There is no standardized training course in the USMC for the DVTE either in person or on the distance learning platform, MarineNet.

The USMC should begin with a distance learning course designed on MarineNet to teach Marines the functionality, use, and implementation of VBS3 within the DVTE system. A training course even on a distance learning platform will help bolster the training opportunities and enrich the available training demanded by the T&R Manual. The USMC should also develop two in person courses to be made available at each major base. One would be for NCOs to learn how to operate and conduct training with the DVTE for their fire teams and squads. The other course would be for staff NCOs and officers to conduct training for their platoons and squads.

(3) There needs to be further integration of Joint Fires Observer (JFO) and UAS training

General Berger placed an emphasis on distributed operations and the expeditionary advanced base operations concept in [27]. Command and control are not only important to operations, but they will have new prominence in distributed operations. The roles of the squad systems operator and the JFO will be tightly coupled if not possibly one in the same to achieve a greater amount of control and operational picture from the squad level during distributed operations. Some of the best performance results in this study were from a participant who was trained as both a JFO and UAS operator. One thing this might suggest is that the USMC must do a better job offering JFO and UAS training to Marines, especially those filling the role of the squad systems operator.

The role of the squad systems operator will be crucial in the distributed fight, and the USMC needs to do a better job of training squad leaders to use the Marine in the squad systems operator billet and train the squad systems operator. A curriculum for the operator should include a JFO course to better understand fires, procedural reporting, and communication and the UAS operator’s course to be able to pilot squad and platoon sized UAS.

As part of the training curriculum, a standardized reporting method, such as CFF, CASEVAC, or close air support reporting formats should be instituted when controlling

UAS. Such a report would help to facilitate the communication between the squad leader and squad systems operator, as was not achieved by 28 of the 34 participants in this research.

D. FUTURE WORK

1. Experimental Redesign

Future work in this area could focus on redesign of the experiment in two main methods: either choosing a different ISE other than VBS or by reducing the variability within the experiment.

A redesign with a completely different ISE could help to solidify the results and offer a refreshed take on the USMC commitment to VBS. It would be heavily involved research and may require either the coding of a new ISE, revival of a previously used ISE such as something like Close Combat Marine or searching for a COTS solution that could meet the training requirements needed in the USMC ISE.

Another method of redesign would be to reduce the variability of the study to make it more easily repeatable and reduce the number of variables. Instead of focusing on CASEVAC, CFF, and UAS control tasks, reduce the task to one task at a time and tie it directly to the T&R Manual. Such a redesign would help limit the time necessary to conduct the experiment and could afford the opportunity to allow more participants in a shorter amount of time.

2. Virtual Reality integration

A possible area for future work particularly within the computer science field would be to integrate virtual reality within the experiment. It could add to the fidelity of the ISE and attempt to increase the user experience of the squad leader participating in the study. It could also assist in creating the interface that is needed by infantry Marines within training simulations.

The drawback of virtual reality integration is that it would need to be done in conjunction with Bohemia Interactive if VBS were to be used as the ISE in the study. Such a collaboration could prove very beneficial for both Bohemia Interactive and the USMC.

It would be possible to also design a virtual reality ISE without the help of Bohemia Interactive or use of their proprietary VBS software. Such a project would be a substantial undertaking or could use another COTS system.

3. Validation with Live Training

While the USMC continues to implement LVC-TE across its training and is included in the T&R Manual, live training is necessary to validate the ability of any ISE. Such an undertaking would require more time, participants, ammunition, and reservation of training areas to facilitate. It would take more time as participants would have to be trained within the ISE and then would have to conduct training in a live environment, either in a force-on-force scenario during which other Marines act as opposition forces for the experiment participant, or during a live fire exercise in which ammunition would be used to conduct a simulated attack.

The training would also require more participants such as those acting as opposition forces or for the members of the squad leader’s squad. Exercise control would have to be a larger component of the experiment as well, and it would require more than three people to conduct the study.

An intermediary solution could be to use the IIT as the live validation test. Although a simulated environment, the IIT could offer validation for cognitive workload training in an ISE. The IIT combines some components of force-on-force training with the use of actors as opposition forces and components of live fire with the use of blank ammunition and detonation simulators.

E. SUMMARY

This research aimed to identify ISEs as a valuable training tool that could replicate scenarios to an extent that would decrease cognitive workload for infantry squad leaders in the USMC. Squad leaders were chosen as the target demographic as they could potentially have the most to gain from the features of modern ISEs and the training associated with them.

In Chapter II, the literature review explored the background of cognitive workload research and the importance of subjective and objective measures of cognitive workload. The background research led to the development of the experiment and the design implementation of several key features within the simulation. The experiment design served as an attempt to train squad leaders over the course of various scenarios to decrease their cognitive load. A decrease in cognitive load would help them undertake additional responsibilities such as communicating with a COC, an FDC, or a squad systems operator.

The only statistically significant value determined during the analysis of results was a difference in a self-assessed question reported on the TLX survey. That question related to the amount of effort used by the participant to perform in the manner that they did for their individual UAS control task. Many other data points were collected and analyzed, but none proved statistically significant to include differences in EDA, IBI, and pupil diameter.

The results are inconclusive due to a lack of sample size but future research and recommendations from this study are substantial. The survey results and performance of the Marines who participated in the study provided many avenues for future research to include exploring the possibility of implementing virtual reality or validating the ISE with live fire or force-on-force training. It also provided several recommendations to the USMC as an organization to better take advantage of the DVTE and ISEs of VBS to train its leaders.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX A. NASA TLX SURVEY

NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

Name	Task	Date

Mental Demand — How mentally demanding was the task?

Very Low — Very High

Physical Demand — How physically demanding was the task?

Very Low — Very High

Temporal Demand — How hurried or rushed was the pace of the task?

Very Low — Very High

Performance — How successful were you in accomplishing what you were asked to do?

Perfect — Failure

Effort — How hard did you have to work to accomplish your level of performance?

Very Low — Very High

Frustration — How insecure, discouraged, irritated, stressed, and annoyed wereyou?

Very Low — Very High

The NASA Task Load Index. Source: [23].

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX B. EXPERIMENT SCRIPT

Ask applicant to review consent form.

COC/EXCON: “To begin, we need to calibrate the eye-tracking software. As you read in the consent form, this eye tracker will be used to track your gaze on the screen and to measure the diameter of your pupils which can be an indicator of cognitive workload.”

Conduct Gazepoint calibration

“We are going to begin with some simple training exercises that will help you learn the basic controls of the game such as moving around the map and engaging with your weapon. Follow the instructions on the upper right-hand corner of the screen and it will display your tasks.”

Begin VBS3 tutorial 1.

“Now we are going to enter a second tutorial that will assist with marksmanship training.”

Begin VBS3 tutorial 2.

“Now we are going to enter a training scenario in which you will practice call for fire, UAS control, and CASEVAC. You have in front of you the standard templates for CASEVAC and CFF, feel free to write on them and reference them throughout. Please place the heartrate monitor on your wrist and hold down the button.”

Assist participant with placing heartrate monitor on wrist

“In this scenario you will be able to communicate with three resources. You will be playing the role of squad leader, and your callsign for this exercise will be ‘1A’ (one alpha). I will be playing the role of the Company COC, any CASEVAC requests can be routed through me. My callsign for this exercise will be ‘COC.’ You also have a squad systems operator. His callsign is ‘1B’ (one bravo). He has the ability to control the UAS that is flying overhead. You can task him with moving the UAS to a grid or in a certain direction, or changing the elevation. He can also assist in providing situational awareness from the UAS feed picture. Lastly, the FDC will process your CFF requests. Their callsign is ‘FDC.’

You will be able to talk to any of these people with your headset radio. Go ahead and wear your headset now."

All members don headsets.

"The '~' button allows you to transmit through your radio. We will now conduct a radio check. 1A, 1A, this is COC, radio check, over."

Participant: "COC this is 1A, roger, over."

COC/EXCON: "Roger, out. We will now begin the training scenario. Feel free to ask questions at any time, this scenario again allows you to become more familiar with the required tasks."

Begin scenario 1

"To begin, select your laser weapons designator from the weapons screen using the mouse scroll wheel. If you click on a target in the distance, you can see that the designator displays information such as direction, distance, 10 digit grid, and elevation of the target. We will now use that information to practice CFF."

CFF

"Next, choose the UAS control pad. You are able to use this to view the UAS feed and have some options. If you click on a target with the UAS you are able to lase it and pull its information such as 10-digit grid. You can use the "l" button to lock onto that target, and use "l" again to unlock. The "n" button changes the picture from black hot to white hot which may help in seeing some objects. Lastly, the "+" and "-" buttons next to the number pad allow you to zoom in and zoom out. Remember to use your squad systems operator if you choose to move the position of the UAS."

FOR EXPERIMENTAL GROUP PARTICIPANTS:

"Now we will practice UAS controlling. Your squad systems operator has a map and is an expert in UAS piloting. As an example, request him to move overhead to your position."

Participant: "1B, this is 1A, can you move the UAS over to my position?"

SSO: "1A, 1B, enroute to your position time now." … "UAS is one km out from your position, it will be overhead in 45 seconds." … "UAS is overhead your position at 500 feet elevation in a 200 ft radius holding pattern."

Participant: "Roger, copy all."

COC/EXCON: "Excellent, you can now see your own avatar on the screen at this time. If you exit the UAS feed, we will practice CFF. To get your own grid, press "k" to see your DAGR. Assume that a member of your squad just took gunfire and has a sucking chest wound. Assume no CBRN contamination. Proceed with a CASEVAC request to your location."

Conduct CASEVAC scenario.

"Now we will attempt another scenario in which you will be asked to perform some of the same tasks that you just went over. For this next scenario, you will be in a creation of the Mobile MOUT range on Camp Lejeune. You will begin in building 64. Your squad systems operator is in building 63, and he has control of the UAS. You are currently dismounted with a platoon of LAVs in a herringbone formation holding security along the north-south running route to your east. The intersection to your east at the 387 northing can be used as a CASEVAC LZ, named LZ EAGLE.

Your mission is to reconnoiter the city in order to allow for follow on clearing operations.

(Repeat) Your mission is to reconnoiter the city in order to allow for follow on clearing operations.

LAVs are not permitted into the city as the roads have been deemed unsuitable for LAV traffic. The ROE states that all activity in the city is enemy and cleared for engagement.

The squad systems operator with UAS, FDC, and COC are all available to process requests and reports. They all have the same maps and GRG as you and will be able to reference the building numbers labeled on your map.

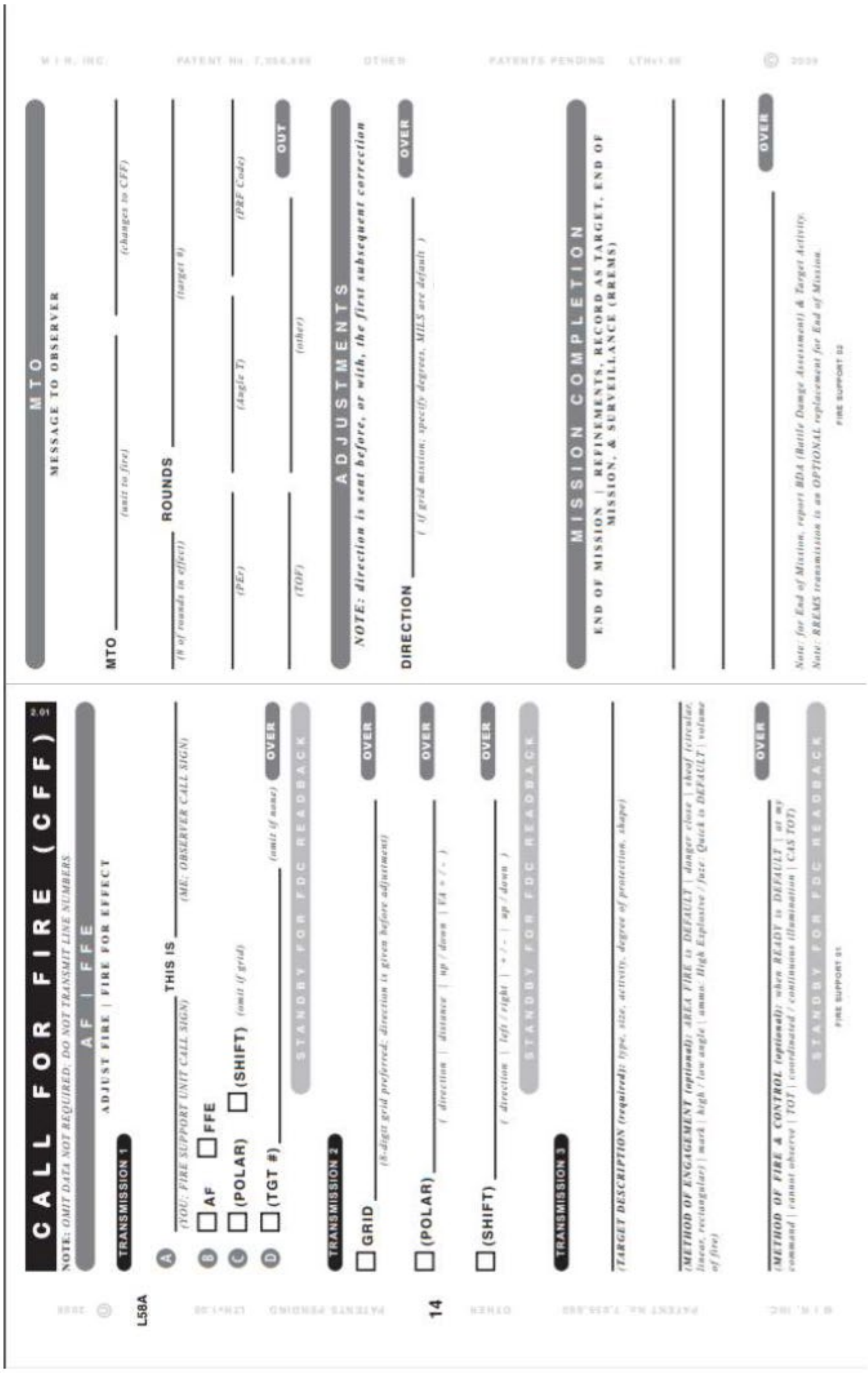
Please let me know if you have any questions regarding your task, the scenario, or the simulation."

Conduct Scenario #3

APPENDIX C. GRG USED IN EXPERIMENT

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX D. CFF AND CASEVAC TEMPLATES



L58A

2

CASEVAC REQUEST / MEDEVAC INFO

1.00

FROM FAC TO DASC / DASC(A) / TACC(A) HIGHER

VHF (S/C/PT) (PRI): ____________ (ALT): ____________

UHF (PRI): ____________ (ALT): ____________

SAT PHONE: ____________

CASEVAC REQUEST / NATO 10-LINE

NOTE: *Report the first three lines of the report to begin the process; follow-up with the remainder*

LINE 1
GRID COORD OF PICK-UP SITE: ____________

LINE 2
RADIO FREQ: ____________ CALL SIGN: ____________

LINE 3 NUMBER OF PATIENTS BY PRECEDENCE:

_____ A URGENT _____ B URGENT / SURGICAL

_____ C PRIORITY _____ D ROUTINE _____ E CONVENIENCE

LINE 4 SPECIAL EQUIPMENT REQUIRED:

☐ A NONE ☐ B HOIST ☐ C EXTRACT EQUIP ☐ D VENTILATOR

LINE 5 NUMBER OF PATIENTS BY TYPE:

_____ L LITTER _____ A AMBULATORY

LINE 6 SECURITY AT PICK-UP SITE:

☐ N NO ENEMY TROOPS ☐ E ENEMY TROOPS IN AREA *(CAUTION RECOMMENDED)*

☐ P POSSIBLE ENEMY TROOPS ☐ X ENEMY TROOPS IN AREA *(ARMED ESCORT RECOMMENDED)*

REPORTING 01

LINE 7 METHOD OF MARKING PICK-UP SITE:

☐ A PANELS ☐ B PYRO ☐ C SMOKE ☐ D NONE ☐ E OTHER

LINE 8 PATIENT NATIONALITY & STATUS:

☐ A US MILITARY ☐ B US CITIZEN ☐ C NON US MILITARY

☐ D NON US CITIZEN ☐ E ENEMY PRISONER OF WAR (EPW)

LINE 9 CBRN CONTAMINATION:

☐ A RADIOLOGICAL ☐ B BIOLOGICAL ☐ C CHEMICAL ☐ D NONE

LINE 10 PATIENT INFORMATION:

____

____

____

FIRST INITIAL, MIDDLE INITIAL, LAST NAME, LAST 4 SSN, BLOOD TYPE

CASEVAC / MEDEVAC CONSIDERATIONS

MEDEVAC: medical capability on board (e.g. a corpsman)
CASEVAC: *TRANSPORTATION ONLY;* no medical capability

1. Determine the location of the Platoon CCP and the Company CCP
2. Determine location of Battalion Ambulance Exchange Point
3. Designate MEDEVAC vehicle and ensure that it is empty
4. Walk through both Ground or Air or combination of both
5. Equipment considerations
   a. Pole-less litters
   b. Smoke
   c. IR Strobe
   d. Air Panel
6. Additional considerations
   a. Who is your MEDEVAC vehicle?
   b. Who is security escort for med vehicle?
   c. Will there be a vehicle exchange at the CCP? Escort responsibilities at CCP to higher echelon?
   d. Ground vs. Air
      ◊ Time
      ◊ Distance

REPORTING 02

LIST OF REFERENCES

[1] D. Berger, “Force Design 2030.” Headquarters Marine Corps, 2020. [Online]. Available: https://www.hqmc.marines.mil/Portals/142/Docs/CMC38%20Force%20Design%202030%20Report%20Phase%20I%20and%20II.pdf?ver=2020-03-26-121328-460

[2] T. B. Sheridan, “Mental workload in decision and control,” *IEEE Conf. Decis. Control Symp. Adapt. Process.*, vol. 18, pp. 977–982, 1979, doi: 10.1109/CDC.1979.270095.

[3] J. P. Machuca, M. E. Miller, and J. M. Colombi, “A Cognitive Task Analysis-based Evaluation of Remotely Piloted Aircraft Situation Awareness Transfer Mechanisms,” *2012 IEEE Int. Multi-Discip. Conf. Cogn. Methods Situat. Aware. Decis. Support*, pp. 179–182, Mar. 2012, doi: 10.1109/CogSIMA.2012.6188376.

[4] M. Strom and J. McDonough, “The Forward Observer Personal Computer Simulator (FOPCSim) 2.” M.S. thesis, Dept of Computer Science, NPS, 2005. [Online]. Available: http://hdl.handle.net/10945/1970

[5] T. Luong, F. Argelaguet, N. Martin, and A. Lecuyer, “Introducing Mental Workload Assessment for the Design of Virtual Reality Training Scenarios,” *IEEE Conf. Virtual Real. 3D User Interfaces VR*, pp. 662–671, 2020.

[6] J. Forbes, “A Primer on Modeling and Simulation.” National Training and Simulation Association, 2011. [Online]. Available: https://www.trainingsystems.org/-/media/sites/ntsa/homepage/miscellaneous/ms-primer.ashx?la=en

[7] 110th Congress, “House Resolution 487.” U.S. House of Representatives. [Online]. Available: https://www.congress.gov/bill/110th-congress/house-resolution/487/text

[8] “Planning, Marine Corps Doctrinal Publication 5.” Headquarters Marine Corps, 2018. [Online]. Available: https://www.marines.mil/portals/1/Publications/MCDP%205.pdf?ver=2018-10-04-151514-513

[9] “Marine Corps Planning Process, Marine Corps Warfighting Publication 5–10.” Headquarters Marine Corps, 2020. [Online]. Available: https://www.usmcu.edu/Portals/218/CDET/content/other/MCWP%205-10.pdf

[10] Staff, PM TraSys S&T Division, “Marine Corps Family of Tactical Decisionmaking Simulations,” *Mar. Corps Gaz.*, 2004.

[11] M. P. Bailey and L. R. Armstrong, “The Deployable Virtual Training Environment,” presented at the Interservice/Industry Training, Simulation and Education Conference, 2002.

[12] “Justification and Approval to Procure Using Other Than Full and Open Competition.” United States Marine Corps Systems Command, 2012.

[13] “U.S. Marine Corps strengthens training toolkit; acquires lots of virtual battlespace product licenses.” Geospatial World, Sep. 01, 2020. [Online]. Available: https://www.geospatialworld.net/news/u-s-marine-corps-acquires-hundreds-of-vbs4-licenses-for-dvte-program/

[14] “MARADMIN 219/17: TACTICAL DECISION KIT DISTRIBUTION AND IMPLEMENTATION.” Headquarters Marine Corps, May 04, 2017. [Online]. Available: https://www.marines.mil/News/Messages/MARADMINS/Article/1176937/tactical-decision-kit-distribution-and-implementation/

[15] “Tactical Decision Kit (TDK) System Overview.” Marine Corps Rapid Capabilities Office, Jun. 15, 2017.

[16] F. Paas, P. Ayres, and M. Pachman, “Assessment of Cognitive Load in Multimedia Learning Theory, Methods and Applications,” in *Recent Innovations in Educational Technology that Facilitate Student Learning*, Information Age Publishing, 2008, pp. 11–26.

[17] J. J. G. Merrienboer and J. Sweller, “Cognitive load theory in health professional education: design principles and strategies,” *Med. Educ.*, vol. 44, no. 1, pp. 85–93, 2009, doi: 10.1111/j.1365-2923.2009.03498.x.

[18] N. Debue and C. van de Leemput, “What does germane load mean? An empirical contribution to the cognitive load theory,” *Front. Psychol.*, 2014.

[19] M. Vandewaetere and G. Clarebout, “Cognitive Load of Learner Control: Extraneous or Germane Load?,” *Hindawi Publ. Corp.*, vol. 2013, Mar. 2013, doi: 10.1155/2013/902809.

[20] J. Sweller, “Element Interactivity and Intrinsic, Extraneous, and Germane Cognitive Load,” *Educ. Psychol. Rev.*, vol. 22, no. 2, pp. 123–138, 2010, doi: 10.1007/s10648-010-9128-5.

[21] J. Sweller, P. Ayres, and S. Kalyuga, *Cognitive Load Theory. Explorations in the Learning Sciences, Instructional Systems, and Performance Technologies.*, vol. 1. New York, NY: Springer, 2011.

[22] S. G. Hart and L. E. Staveland, “Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research.” Elsevier Science Publishers, 1988. [Online]. Available: https://humansystems.arc.nasa.gov/groups/TLX/downloads/Hart_Staveland_ORIGINAL_1.pdf

[23] F. Paas, “Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach,” *J. Educ. Psychol.*, no. 84, pp. 429–434, 1992.

[24] D. Nemeth, “Benchmarking the RTI for Use in a Simulated Radio Environment.” Advanced Simulation Technology, Inc.

[25] “Marine Infantry Platoon, Marine Corps Interim Publication 3–10A.3i.” Headquarters Marine Corps, 2019. [Online]. Available: https://www.marines.mil/portals/1/Publications/MCIP%203-10A.3i.pdf?ver=2019-06-20-095433-550

[26] “Guidebook for Marines.” The Leatherneck Association, Inc., 1948.

[27] D. Berger, “Force Design 2030 Annual Update.” Headquarters Marine Corps, 2021. [Online]. Available: https://www.marines.mil/Portals/1/Docs/2021%20Force%20Design%20Annual%20Update.pdf

[28] “VBS Developer Documentation.” Bohemia Interactive, 2013. [Online]. Available: https://community.bistudio.com/wiki/allowDamage

[29] V. Jorrig, D. Vera, and K. Jutta, “The Index of Cognitive Activity as a Measure of Cognitive Processing Load in Dual Task Settings,” *Front. Psychol.*, vol. 9, 2018, doi: 10.3389/fpsyg.2018.02276.

[30] C. Setz, B. Arnrich, R. La Marca, G. Troster, U. Ehlert, and J. Schumm, “Discriminating Stress From Cognitive Load Using a Wearable EDA Device,” *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 2, pp. 410–417, Mar. 2010, doi: 10.1109/TITB.2009.2036164.

[31] M. Ahmed, S. Begum, and M. Islam, “Heart Rate and Inter-beat Interval Computation to Diagnose Stress Using ECG Sensor Signal.” Mälardalen University, 2010.

[32] A. Hughes, G. Hancock, S. Marlow, K. Stowers, and E. Salas, “Cardiac Measures of Cognitive Workload: A Meta-Analysis,” *Hum. Factors J. Hum. Factors Ergon. Soc.*, vol. 61, no. 3, pp. 393–414, 2019.

[33] J. Barlow, “A Declaration of the Independence of Cyberspace.” Electronic Frontier Foundation, Feb. 08, 1996. [Online]. Available: https://www.eff.org/cyberspace-independence

[34] R. Alvis, "Marines Announce Changes to Ground Combat Element Aimed at Improving Lethality and Agility." Headquarters Marine Corps, May 09, 2018.

[35] "NAVMC 3500.44D Infantry Training and Readiness Manual." Headquarters Marine Corps, May 07, 2020. [Online]. Available: https://www.marines.mil/Portals/1/Publications/NAVMC%203500.44D%20Infantry%20T-R%20Manual.pdf?ver=2020-05-21-104009-273

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California