Old Dominion University

ODU Digital Commons

Computational Modeling & Simulation Engineering Theses & Dissertations

Computational Modeling & Simulation Engineering

Winter 2012

A Framework for Delivering Contextually Appropriate Opportunities for Warfighter Practice

Michael Allan White
Old Dominion University

Follow this and additional works at: https://digitalcommons.odu.edu/msve_etds

Part of the Engineering Commons, and the Military and Veterans Studies Commons

Recommended Citation

White, Michael A.. "A Framework for Delivering Contextually Appropriate Opportunities for Warfighter Practice" (2012). Doctor of Philosophy (PhD), Dissertation, Computational Modeling & Simulation Engineering, Old Dominion University, DOI: 10.25777/36s7-9s61
https://digitalcommons.odu.edu/msve_etds/43

This Dissertation is brought to you for free and open access by the Computational Modeling & Simulation Engineering at ODU Digital Commons. It has been accepted for inclusion in Computational Modeling & Simulation Engineering Theses & Dissertations by an authorized administrator of ODU Digital Commons. For more information, please contact digitalcommons@odu.edu.

A FRAMEWORK FOR DELIVERING CONTEXTUALLY APPROPRIATE OPPORTUNITIES FOR WARFIGHTER PRACTICE

by

Michael Allan White
M.B.A. August 1985, Embry-Riddle Aeronautical University
B.S. April 1983, Embry-Riddle Aeronautical University

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

DOCTOR OF PHILOSPOPHY

MODELING AND SIMULATION

OLD DOMINION UNIVERSITY
December 2012

Approved by:

Frederic D. McKenzie (Director)

John A. Sokolowski (Member)

Yuzhong Shen (Member)

Kay M. Stanney (Member)

UMI Number: 3534961

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

UMI 3534961
Published by ProQuest LLC 2013. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.
All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

ABSTRACT

A FRAMEWORK FOR DELIVERING CONTEXTUALLY APPROPRIATE OPPORTUNITIES FOR WARFIGHTER PRACTICE

Michael Allan White
Old Dominion University, 2012
Director: Dr. Frederic D. McKenzie

Computer-based modeling and simulation has been a training staple in the military domain since the first aircraft simulators were adopted. More recently, virtual environments based on modeling, simulation and serious games, have introduced relatively low-cost, yet high value additions to the learning environment. As these virtual environments have proliferated, many researchers have investigated the relationship between theoretical foundations of learning, learner development and content delivery, and applied their findings in an attempt to bolster learning, yet performance deficiencies continue to exist. This study asserts that performance deficiencies exist in part because of insufficient contextually appropriate opportunities to practice.

This work is multi-disciplinary in nature. Its foundation is modeling and simulation engineering; the use of technology to deliver training. Educational psychology and human factors concepts explain the theoretical basis for modeling and simulation as an effective training delivery agent.

The study's thesis is that a framework for delivering contextually appropriate opportunities for warfighter practice can be applied to discover whether modeling, simulation and game-based virtual environments have the potential to improve individual performance for learners beyond the Novice Stage (e.g., Competent Stage) of skills acquisition. Furthermore, this conceptually appropriate practice (CAP) framework can be

used to assess the potential of low fidelity virtual environments to provide targeted practice and to improve individual performance, not only during training in high-fidelity virtual environments (near transfer) but also in the live environment (far transfer).

To evaluate the thesis, this study investigates the relationship of technology and learning science, and features an empirical evaluation of training effectiveness afforded by delivering additional training repetitions using both low-fidelity virtual environment simulator systems and high-fidelity aircraft simulators.

Copyright, 2012, by Michael Allan White, All Rights Reserved.

Dedicated to the men and women unselfishly serving in the United States' armed forces.

ACKNOWLEDGMENTS

Many people (more than I could possibly list here) were influential in "getting me across the finish line" of this educational pursuit. Several are noteworthy. RADM (Ret) Gary Jones: for his unrelenting support and encouragement. Helicopter Maritime Strike Weapons School-Atlantic: Commanding Officer, CDR Andrew Danko, and instructors LT Emile Therrien and LT Michael Henderson for unselfishly sharing their time, knowledge and professional insight. Their assistance was indispensable to facilitating this study. Naval Air Warfare Center Training Systems Division, Orlando: Mr. Jonathan Glass, Mission Rehearsal Tactical Team Trainer Program Manager and LCDR Tracy Parsons, H-60 Helicopter Program Manager for providing background reference literature and coordinating the technical support required to conduct the empirical study. I'm deeply grateful to Jack McGinn for making me aware of the program and to former ODU Graduate Program Director, Dr. Bowen Loftin, for giving me the opportunity to pursue this research. Thanks to Dr. Rick McKenzie and Dr. John Sokolowski for their patience and mentorship throughout the process and to my good friend, Dr. Kay Stanney, for urging me not to settle for mediocrity.

Of course, adult educational pursuits require a significant investment in time and energy, and accommodations on the home front; especially when the student has a full-time "day job." My wife, Kathy gave me the space and time needed to complete this effort, while our children and grandchildren were a constant source of support and inspiration; cheering me on when I sometimes felt like quitting.

S.D.G. – M.W. 2012

NOMENCLATURE

Computer Simulation – The discipline of designing a model of an actual or theoretical physical system, executing the model on a digital computer, and analyzing the execution output. (Fishwick, 1995)

Constructive Simulation – A constructive simulation is a computer program in which simulated people operate simulated systems. Real people stimulate (make inputs) to such simulations, but are not involved in determining the outcomes. (MSCO, 2010)

Fidelity – The identification of key parameters for a system and the degree to which the aggregate of those parameters match a baseline system. The components of fidelity include functional, physical, psychological, tactile, visual, and wallpaper. (MSCO, 2012)

Game – A physical or mental competition in which the participants, called players, seek to achieve some objective within a given set of rules. (MSCO, 2010)

High Fidelity Virtual Environment – For this investigation, a high physical fidelity, full motion simulator containing authentic emulated controls, systems and sensor displays connected to a synthetic representation of the real world.

Live Simulation – Real people operating real systems. Military training events using real equipment are live simulations. They are considered simulations because they are not conducted against a live enemy. (MSCO, 2010)

Low Fidelity Virtual Environment – For this investigation, a low physical fidelity, laptop PC-based simulator containing a synthetic representation of the real world mission space.

Model – A physical, mathematical, or otherwise logical representation of a system, entity, phenomenon, or process. (MSCO, 2010)

Simulation – A method for implementing a model over time. (MSCO, 2010)

VE or Virtual Environment - Interactive, virtual image displays enhanced by special processing and by nonvisual display modalities, such as auditory and haptic, to convince users that they are immersed in a synthetic space. (Ellis, 1994)

Virtual Simulation – Real people operating simulated systems. Virtual simulations inject human-in-the-loop in a central role by exercising motor control skills (i.e., flying an airplane), decision skills (i.e., committing fire control resources to action), or communication skills. (MSCO, 2010)

TABLE OF CONTENTS

LIST OF TABLES

Table Page

Table Page

LIST OF FIGURES

Figure Page

Figure Page

CHAPTER 1

INTRODUCTION

> On the value of contextually appropriate targeted practice:
>
> "Fifteen minutes on the driving range can make a four-hour round of golf much more productive and enjoyable." – Gary R. Jones, RADM (Ret), Former Commander, Naval Education and Training Command

Computer-based modeling and simulation has been a training staple in the military domain since the first aircraft simulators were adopted. More recently, serious games and virtual environments have introduced relatively low-cost, yet high value additions to the learning environment. As models, simulations and serious games-based virtual environments have proliferated, researchers have investigated the relationship between theoretical foundations of learning, learner development and content delivery, and their application to learning, yet performance deficiencies continue to exist. Is this phenomenon a result of error in or improper application of theory, shortfalls in instructional design / instructional systems design, inappropriate content delivery, or is it as Bowers et al. (2009) suggest in their unpublished manuscript – the lack of opportunity to practice?

A few researchers have investigated the relationship between instructional design and instructional systems design (e.g., Hill & Hanaffin, 2001; Clark, 2004), and others have compared specific training system attributes and how they can be applied to achieve certain desired learning outcomes (Engel et al., 2009; Kraiger, Ford & Salas, 1993). Cohn et al. (2007) addressed the problem of assessing training effectiveness from a systems engineering perspective, presenting an iterative lifecycle approach to instructional system design to shape a system's ability to meet training objectives. However, a gap remains in the literature vis-à-vis introduction of modeling, simulation and serious games-based virtual environments into the learning environment to provide additional contextually appropriate opportunities to practice.

The old saying "practice makes perfect" is correct, to a point. What if learners practice the wrong thing, or practice in the wrong way? Will the resultant performance reflect those errors? What if content delivery systems do not provide contextually appropriate opportunities to practice? How accurate do these systems need to be to support meaningful practice?

Investigating whether practicing the wrong thing or practicing in the wrong way will result in subsequent performance errors is beyond the scope of this study. However, previous studies have investigated some of the other factors. For example, Salas and Cannon-Bowers (2001) observe that context matters, trainees need an opportunity to perform, and delays between training and actual use on the job create significant skill decay. Bowers et al. (2009) points out that a practice environment must provide the appropriate set of cues – at an adequate level of fidelity – to elicit targeted behaviors. Driskell et al. (1992) suggest a correlation between the concept of over-learning and retention, while Ricci et al. (1996) proffer the efficacy of games on knowledge acquisition and retention.

Especially relevant to this study, are studies that suggest targeted practice improves performance at the Novice Stage of skills acquisition. Several studies describe how additional practice using a low fidelity game-based virtual environment improved novice training performance both at the individual level and the team level.

For instance, Hussain et al. (2009) report on the use of a low-fidelity game-based virtual environment to reinforce damage control skills that Navy recruits had been exposed to in lectures, but had not had a chance to practice in context. Taking another perspective, Bowers and colleagues (Bowers et al., 2009), asserted that a game-based "pre-experience" can lead to superior performance and fewer critical errors. Citing Salas and Cannon-Bowers (2003), Bowers et al. (2009) observe that although there are a few factors associated with effective training, the opportunity to practice may be the most overlooked element in training research and development.

Additionally, Hussain et al. (2011) link individual cognitive readiness to improved team performance, noting that not only did performance of individuals acting independently on a real-world, near transfer task show broad-based improvement, but that team performance on several real-world far transfer tasks also showed significant improvement after practicing as individuals.

The preceding cited research was targeted at novice learners. My research focuses on individual performance at levels beyond the Novice Stage (e.g., Competent Stage) of skills acquisition as defined in Dreyfus and Dreyfus (1986). My research investigates individual performance while participating in team training. Although team performance improvement is a very relevant, important topic and certainly worthy of consideration for future work, it is not specifically targeted in this investigation.

Thesis Statement

A framework for delivering contextually appropriate opportunities for warfighter practice can be applied to discover whether modeling, simulation and game-based virtual environments have the potential to improve individual performance for learners beyond the Novice Stage (e.g., Competent Stage) of skills acquisition. Furthermore, this conceptually appropriate practice (CAP) framework can be used to assess the potential of low fidelity virtual environments to provide targeted practice and to improve individual performance, not only during training in high-fidelity virtual environments (near transfer) but also in the live environment (far transfer).

Research Questions

An empirical study associated with this research seeks to determine whether a CAP framework can be applied to answer the following questions:

1. Given that delivering targeted practice via Virtual Environments (VE) contributes to improved individual performance at the Novice Stage of skills acquisition (Hussain et al., 2009), does VE-delivered targeted practice also contribute to improved individual performance at more advanced (e.g., Competent) stages of skills acquisition?
2. Does low-fidelity VE-delivered targeted practice contribute to improved individual performance during training in high fidelity simulators? During Live training?

Research Objectives

Objectives of this study include investigating theories and principles associated with:

1. Effective training
2. Training/learning/instructional design

3. Performance measurement
4. Training technologies
5. Training/learning/instructional system design
6. Training system attributes.

Additionally, this study investigates how theories and principles are applied in military environment, specifically in the areas of:

1. Training and readiness
2. Training systems
3. Current problems.

Approach/Methodology

This observational study, conducted in the context of helicopter aircrews performing two Anti-Submarine Warfare (ASW) mission tasks (1. Detect / Localize and 2. Track) during regularly scheduled training, also investigates individual aircrew member Crew Resource Management (CRM) interactions; specifically communications, leadership / teamwork, and situation awareness while performing the mission tasks. A training intervention was introduced, in which a treatment group practiced the same tasks in a low fidelity VE trainer prior to both treatment and control groups practicing the tasks in a high fidelity VE. All participants were provided two 2-hour practice sessions in the high fidelity VE, followed a few weeks later by a live assessment in helicopters on an instrumented range located at the Atlantic Undersea Test and Evaluation Center. During both VE practices and during the live event, instructors assessed aircrew members' mission task performance and CRM skills using four data collection instruments.

Contributions

This study contributes to filling a gap in the training research and development literature concerning the use of modeling, simulation and gaming technologies to deliver contextually appropriate opportunities to practice. Toward that end, this effort develops a multi-discipline Contextually Appropriate Practice – Delivery (CAP-D) framework, which leverages learning theory, principles of instructional design and instructional system design, along with concepts of training and transfer effectiveness. The framework simultaneously captures the key elements of learner development from a theoretical standpoint while offering the flexibility to modify elements of the model to accommodate changes in modeling and simulation content delivery mechanisms, context and learner skills acquisition stage.

A complementary Contextually Appropriate Practice – Assessment (CAP-A) framework provides a means to evaluate a system's suitability to deliver contextually appropriate targeted practice. Within the CAP-A construct, practitioners can use data collection instruments to capture quantitative and/or qualitative measures that describe performance in selected mission task areas along with other targeted learning objectives of interest to determine if the combination of content, delivery mechanism, context and targeted practice achieve the desired learning outcomes. More detailed evaluation of assessment scores can point to strengths, weaknesses and potential areas for remediation.

This study demonstrates the flexibility of the CAP-D and CAP-A frameworks by showing how they apply to the use of a game-based delivery system within the context of Novice Stage training and subsequently to Competent Stage training using low-fidelity and high-fidelity VE training systems.

Dissertation Organization

The following provides a description of the remaining chapters included in this dissertation.

Chapter 2: Background – presents selected theoretical research related to the science of learning, principles of instructional systems design, and training effectiveness and evaluation.

Chapter 3: Related Research – presents relevant applied research in training effectiveness and evaluation of learners at the Novice Stage of skills acquisition and includes a description of the Contextually Appropriate Practice – Delivery (CAP-D) and Contextually Appropriate Practice – Assessment (CAP-A) frameworks and the application of CAP-D and CAP-A frameworks to Novice Stage learners.

Chapter 4: Military Training Domain – provides the operational context for this study. The chapter discusses current problems and introduces examples of modeling, simulation, and serious game-based virtual environments that have been used to augment more traditional means of military training. The chapter concludes with a discussion on preparatory activities and events related to this investigation's empirical study.

Chapter 5: Method – details the experimental design, methodology, techniques and instruments used to conduct the investigation.

Chapter 6: Results – presents findings and data analysis of the empirical study and application of CAP-D and CAP-A frameworks to Competent Stage learners.

Chapter 7: Conclusions and Recommendations – discusses significance of the study and presents a suggested roadmap for future research.

CHAPTER 2

BACKGROUND

This chapter presents selected theoretical research related to the science of learning, principles of instructional systems design and training effectiveness and evaluation. Although not all the concepts provided in this background discussion are applied to the CAP investigation, many could be topics of interest for future research.

Technology, Learning Science and Military Training

The role of technology is especially evident through its impact on learning science in general, and military training, in specific. Since the advent of the electronic digital computer, military and civilian educators have used computer based modeling and simulation for training, education, learning and development. The proliferation of cheaper, faster, more portable computers has enabled educators and trainers to reach a much broader audience using modeling, simulation and serious games to augment more traditional instructional delivery methods.

Until recently, there seemed to be a reticence among military leaders to employ "games" for training. Smith (2008) reports that many people in non-entertainment industries find it difficult to use terms like "games" and "game technologies" when referring to the education and training domains. He goes on to cite Ghamari-Tabrizi's (1995) observation that the military initially kept secret the use of serious games for fear of public ridicule. However, over the past decade the military has welcomed *serious games* as a cost-effective addition to the trainer's tool set. For instance, since "America's Army" was released in 2002, nearly 9.5 million people have played the game. Although its original purpose was to serve as a recruiting tool, the Army has adapted the game for such varied uses as weapons prototyping, helping wounded soldiers adjust to their injuries and for training (Testa, 2009). Additionally, the Army has adopted the Virtual Battle Space 2 as a low cost, low overhead training system for soldiers at company level and below (King, 2009).

Theoretical Foundations of Learning Science

In the civilian realm, an example of contemporary instructional delivery methods involves a three-pronged approach; 1) Informational - instructor-led lecture, discussion, debate, and presentation, enhanced by audio-visual and/or computer-based media, and individual reading/self-study; 2) Attitudinal – role playing, simulation games, task groups or skits; and 3) Behavioral - role playing, simulation games, case study, demonstration, and skills practice lab (I-TECH, 2004). This approach is reminiscent of the work Benjamin Bloom and David Krathwohl pioneered in the late 1940's to mid-1950's that has been widely adopted by educators. Known as Bloom's Taxonomy, the theory presents a set of process classifications associated with educational goals and objectives featuring three learning domains: Cognitive, Affective and Psychomotor (Bloom & Krathwohl, 1956).

Notwithstanding the oft misunderstood Marshall McLuhan quote "the medium is the message" (McLuhan, 1964), the method of delivery does not supersede the content in importance (Federman, 2004), although presentation does have a great deal of influence on how content is perceived and received. Instructional content consists of a variety of knowledge types. Bloom and Kratwohl (1956) originally offered three knowledge dimensions; *factual*, *conceptual* and *procedural*. Anderson and Kratwohl (2001) more recently updated the list by adding the *metacognitive* dimension. Allen Newell in 1972 and later, John Anderson in 1983 discussed the distinctions between declarative and procedural knowledge, defining *declarative knowledge* as "knowing what" and *procedural knowledge* as "knowing how" (Clark, 2011).

Although the preceding discussion provides a brief overview of learning science, it is by no means exhaustive. Not all learners come to the learning experience at the same level of skills and abilities. In 1979, Steinaker and Bell published their conceptualization of a five-stage experiential learning taxonomy. A few years later, Hubert and Stuart Dreyfus (1981; 1986; 2004) devised a learner centric model of progressive skills acquisition, describing five stages of learner development. Although both efforts were independent of one another, each has been foundational for other researchers interested in staged models of skills acquisition.

Taxonomy of Experiential Learning

Steinaker and Bell's (1979) taxonomy presented experiential learning as a sequence of progressive stages: exposure, participation, identification, internalization, and dissemination. According to the authors, exposure represents a consciousness of an experience on two levels (*sensory* and *response*) and a subsequent anticipation (*readiness*) towards participation in the experience. Participation is involvement in the experience. Identification occurs when the participant starts to apply previously-learned experiences and knowledge to a situation. Internalization occurs when the participant can apply concepts to other situations. The *dissemination* stage is achieved when the participant is able and motivated to teach others.

The Irish Nursing Board (2001) (as cited in UMDNJ, 2011) modified Steinaker and Bell's taxonomy and applied their version to focus on assessment and evaluation of nursing students (Table 1). Their taxonomy goes beyond the "see one – do one – teach one" concept, which has become a less acceptable approach in the medical community, particularly when invasive procedures and high-risk care are required (Vozenilek et al, 2004). The key difference between the Nursing Board model and the original Steinaker and Bell taxonomy is the shift of participation to a later stage in the progression.

Table 1: The Irish Nursing Board Experiential Taxonomy (2001 as cited in UMDNJ, 2011)

Exposure	The student observes a competent practitioner carry out aspects of nursing care, shows a willingness and ability to relate the practice observed and its underlying theory to previous experience. Is able to analyze and discuss with the practitioner why and how certain aspects of care were carried out, and identifies sources and types of information required to enhance further application of knowledge to the practice observed.
Identification	The student now shows the ability to participate in the delivery of care under supervision on more sustained basis with less prompting and greater confidence. Shows greater ability to communicate effectively. Demonstrates a wish to acquire further information and ability to analyze and interpret information. Applies problem solving skills and knowledge base to meet different situations.
Internalization	The student is able to explain the rationale for nursing action. Requires less supervision while caring for a group of patients/clients, is able to transfer knowledge to new situations. Seeks and applies new knowledge and research findings, demonstrates ability to use problem solving skills, critical analysis and evaluation.
Participation	The student is able to participate under close supervision of a competent practitioner in carrying out aspects of care, having demonstrated knowledge by analysis. Questions practitioner on aspects of care and its rationale, decision-making, practical skills, and means of acquiring further information and opportunities for practice. Shows ability to perform manipulative skills, operationalized communication and problem solving skills with guidance.
Dissemination	Plans, implements and evaluates care for group/clients under minimal supervision. Advises others, shows ability to teach junior colleagues identifies personal management style and shows ability to manage care delivery by junior staff. Critical analysis, evaluation and decision-making skills demonstrated.

The Five-Stage Model of Adult Skills Acquisition

The Dreyfus' learning continuum provides the impetus for instructional design and training system design (also known as instructional system design). Instructional design encompasses curriculum development, while instructional system design focuses on the mechanism for curriculum delivery. According to the Dreyfus model, the five stages of skills acquisition are: Novice, Advanced Beginner, Competent, Proficient and Expert (Dreyfus & Dreyfus, 1986 as cited in UMDNJ, 2011; Dreyfus, 2004). As individuals begin to learn a skill (whether cognitive or psychomotor), they progressively master the rules and then learn how and when to apply them. As skill levels improve, learners tend to rely less and less on rules and can handle more complex interactions between previously learned skills. At the higher levels of skills development, actions become more intuitive rather than simply applying rules and accepted standards. At advanced stages, individuals perceive patterns and automatically know what actions are appropriate. Table 2 depicts key identifiable learner characteristics associated with each stage of the Dreyfus model (Eraut, 1994 as cited in UMDNJ, 2011; Dreyfus, 2004).

Table 2: Dreyfus Model of Skills Acquisition Stages (adapted from Eraut, 1994 as cited in UMDNJ 2011) and Dreyfus, 2004)

Novice	Rigid adherence to rules Lack of context Little situation awareness No discretionary judgment
Advanced Beginner	Guidelines for action based on examples, recognizable cues Some contextual understanding; begins to use maxims Situation awareness still limited Attributes and aspects are treated separately and given equal importance
Competent	Coping with "crowdedness" (i.e., task overload) Emotional involvement; sense of responsibility Sees actions at least partly in terms of longer-term goals Conscious deliberate planning Standardized and routine procedures
Proficient	Sees situations holistically rather than in terms of aspects Sees what is most important in a situation Perceives deviations from the normal pattern Decision-making less labored, but not yet automatic Still reverts to rules and maxims for guidance, to make up for lack of generalizable experience
Expert	No longer relies on rules, guidelines or maxims Intuitive situational response based on deep understanding Analytic approaches used only in novel situations or when problems occur Recognizes what needs to be done and does it

Using the Dreyfus model as a foundation, Patricia Benner (1984; 2004) defined stages of clinical competence (UMDNJ, 2011). Table 3 demonstrates Benner's contribution to the body of knowledge.

Table 3: Benner's Stages of Clinical Competence (1984; 2004 as cited in UMDNJ, 2011)

Novice	Beginners have had no experience of the situations in which they are expected to perform. Novices are taught rules to help them perform. The **rules** are context-free and independent of specific cases; hence the rules tend to be applied universally. The rule-governed behavior typical of the novice is extremely **limited and inflexible**. As such, novices have no "life experience" in the application of rules. "Just tell me what I need to do and I'll do it."
Advanced Beginner	Advanced beginners are those who can demonstrate marginally acceptable performance, those who have coped with enough real situations to note, or to have pointed out to them by a mentor, the recurring meaningful situational components. These components require prior experience in actual situations for recognition. **Principles** to **guide actions** begin to be formulated. The principles are **based on experience**.
Competent	Competence, typified by the nurse who has been on the job in the same or similar situations two or three years, develops when the nurse begins to see his or her actions in terms of long-range goals or plans of which he or she is consciously aware. For the competent nurse, a plan establishes a perspective, and the plan is based on considerable conscious, abstract, analytic contemplation of the problem. The **conscious, deliberate planning** that is characteristic of this skill level helps achieve efficiency and organization. The competent nurse lacks the speed and flexibility of the proficient nurse but does have a feeling of mastery and the ability to cope with and manage the many contingencies of clinical nursing. The competent person does not yet have enough experience to recognize a situation in terms of an overall picture or in terms of which aspects are most salient, most important.
Proficient	The proficient performer perceives situations as wholes rather than in terms of chopped up parts or aspects, and performance is guided by **maxims**. Proficient nurses understand a situation as a whole because they perceive its meaning in terms of long-term goals. The proficient nurse learns from experience what typical events to expect in a given situation and how plans need to be modified in response to these events. The proficient nurse can now recognize when the expected normal picture does not materialize. This holistic understanding improves the proficient nurse's decision making; it becomes less labored because the nurse now has a perspective on which of the many existing attributes and aspects in the present situation are the important ones. The proficient nurse uses maxims as guides which reflect what would appear to the competent or novice performer as unintelligible nuances of the situation; they can mean one thing at one time and quite another thing later. Once one has a deep understanding of the situation overall, however, the maxim provides direction as to what must be taken into account. Maxims reflect nuances of the situation.
Expert	The expert performer no longer relies on an analytic principle (rule, guideline, or maxim) to connect her or his understanding of the situation to an appropriate action. The expert nurse, with an enormous background of experience, now has an intuitive grasp of each situation and zeroes in on the accurate region of the problem without wasteful consideration of a large range of unfruitful, alternative diagnoses and solutions. The expert operates from a deep understanding of the total situation. The chess master, for instance, when asked why he or she made a particularly masterful move, will just say: "Because it felt right; it looked good."

	The performer is no longer aware of features and rules;' his/her performance becomes fluid and flexible and highly proficient. This is not to say that the expert never uses analytic tools. Highly skilled analytic ability is necessary for those situations with which the nurse has had no previous experience. Analytic tools are also necessary for those times when the expert gets a wrong grasp of the situation and then finds that events and behaviors are not occurring as expected When alternative perspectives are not available to the clinician, the only way out of a wrong grasp of the problem is by using analytic problem solving.

RCOG Levels of Competence

To assist residents in achieving its specified competencies, the Royal College of Obstetricians and Gynecologists (RCOG) developed a five-level model of competence RCOG, 2006 as cited in UMDNJ, 2011). The following table (Table 4) illustrates the RCOG approach to progression, based on the degree of supervision needed:

Table 4: RCOG Levels of Competence Model (2006) (as cited in UMDNJ, 2011)

Level 1	Observes	Observes the clinical activity performed by a colleague
Level 2	Assists	Assists a colleague perform the clinical activity
Level 3	Direct Supervision	Performs the entire activity under direct supervision of a senior colleague
Level 4	Indirect Supervision	Performs the entire activity with indirect supervision of a senior colleague
Level 5	Independent	Performs the entire activity without need for supervision

Summary

Although underlying premises and definitions differ, all 'staged' models offer a stepped progression of skills acquisition. Coincidentally, all the models appear to agree that progression involves five stages, as shown in the comparative table below (Table 5):

Table 5: Staged Models of Skills Acquisition (as cited in UMDNJ, 2011)

Dreyfus (1981; 2004) / Benner (1984; 2004)	Steinaker & Bell (1979)	RCOG (2006)
Novice	Exposure	Observes
Advanced Beginner	Participation	Assists
Competent	Identification	Direct Supervision
Proficient	Internalization	Indirect Supervision
Expert	Dissemination	Independent

Training Effectiveness and Evaluation

Training effectiveness and training evaluation are terms often used interchangeably, and although various authors approach the topic from different perspectives, generally there is agreement in most key areas. According to Alvarez et al. (2004), training evaluation is a measurement technique that examines the impact of training programs on learning outcomes and training effectiveness is the study of the variables that likely influence the learning system as a whole. Describing their thoughts on the practical application of training effectiveness and evaluation, Cohn et al. (2008) offer a similar definition and in addition, present a lifecycle approach to training system design and evaluation. Borrowing from the preceding, as well as other studies (refer to Figure 1 for references) Figure 1was created to illustrate theoretical foundations of training effectiveness and evaluation.

As Figure 1 illustrates, four foundational activities (*training needs analysis*, *task analysis*, *training system design* and *instructional design*) form the basis for learning outcomes, which can be measured in terms of *training performance*; how well the training audience performs during training, and *transfer performance*; how well the training audience performs on the job after training (Alvarez et al., 2004). In the layer between the foundational activities and the learning and development outcomes, this model addresses internal learner outcomes (*reactions*, *post-training attitudes* and *cognitive learning*) that the training systems engender (Alvarez et al., 2004; Cohn et al., 2008; Salas & Cannon-Bowers, 2001).

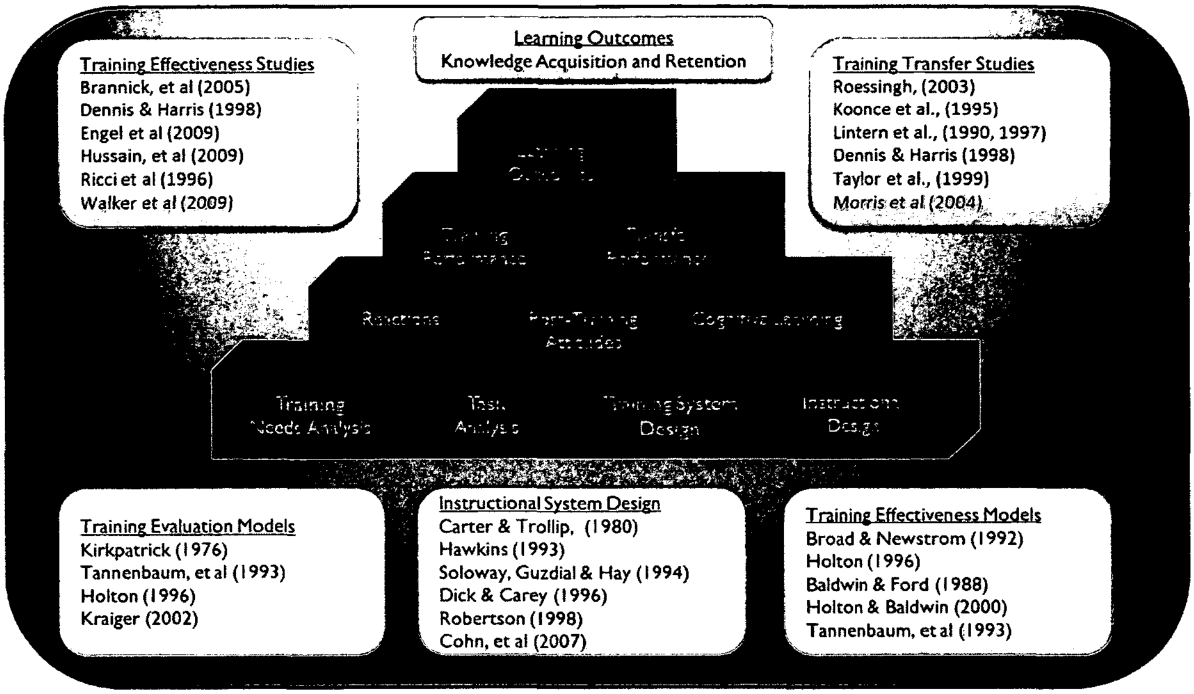

Figure 1: Theoretical Foundations of Training Effectiveness and Evaluation

Training Needs Analysis

Cohn et al. (2008) presents Brown's (2002) description of training needs analysis (TNA) as "the process of collecting data to determine what training needs exist to allow development of training that facilitates accomplishment of an organization's goals" (p.159). The authors offer some practical TNA steps to take, including doctrine review, instructor interviews and direct observation. They go on to point out that input garnered from the aforementioned sources and especially from subject matter experts (SMEs), is critical to understanding training objectives, which in turn serve as the basis for the next step, a more detailed Task Analysis.

Task Analysis

Task analysis (TA) enables "drilling down" to detailed task training objectives (task/subtask) and is targeted at trainee (learner) expertise level, current and desired knowledge, skills and attitudes (KSAs).

Several approaches to TA have emerged. For instance, sensory task analysis (STA) includes identification of required cues (visual, audio, or haptic), which can be quite useful in curriculum development as well as instructional system design, especially as they apply to enhancement of situation awareness in a training virtual environment (Hale, 2006).

Cognitive task analysis (CTA) is another TA approach. In their review of the science of training literature produced during the previous decade, Salas and Cannon-Bowers (2001) define CTA as a set of procedures for understanding the mental processing and mental requirements for job performance. The authors report on research (Rogers et al., 1997) which suggests that continued practice enables learners to automatize complex behaviors. This automaticity in turn, frees up cognitive resources which can then be applied to monitoring and evaluating behavior. Citing Klein (1993), Brenner et al., list five steps for conducting a CTA:

1. Map the task using traditional TA
2. Identify critical decision points
3. Cluster and link the decision points
4. Prioritize the decision points
5. Diagnose and characterize decisions in view of strategies used, cues that signal decision points, and inferences made regarding cues and decision points.

Again, citing Klein (1993), the authors differentiate between traditional TA and CTA. Traditional TA focuses on observable task behavior and does not take into consideration the overall organization of knowledge, while CTA concentrates on the critical decisions and cognitive processes that separate the expert from the novice.

Instructional System Design

Instructional system design translates TNA and TA results into system requirements. One instructional system design method is user-centered design (UCD). UCD identifies interactive, egocentric, and affective cues that support training, and system attributes (functionalities), sufficiently similar to real-world systems to enable learners to practice key competencies, thus supporting transfer of training. (Soloway, Guzdial, & Hay, 1994; cited in Cohn et al., 2008).

Usability analyses are useful in determining needed training system functionalities. Examples are: 1) Heuristic evaluations identify existing usability issues, such as perception or interaction errors. 2) Cognitive walkthroughs employ a “thinking aloud” protocol to identify a rationale for interaction with the training system’s interface. 3) User testing can validate and extend findings from the heuristic and cognitive evaluations. Problem / solution tables describe interaction issues and recommendations to enhance training system design.

Selecting Training System Attributes for Training Outcomes

Training system attributes directly impact learning and development as discussed in a paper on the effectiveness of using games for training by Engel et al. (2009) that addresses how gaming principles can be applied to training applications to retain trainee interest. Founded on the psychology and education research of Kraiger, Ford and Salas (1993) and Bloom (1994), the Engel team identified 14 applicable training outcomes and 11 gaming attributes (Table 6).

Table 6 Comparison: Training Outcomes versus Gaming Attributes (adapted from Engel et al., 2009)

Training Outcomes	Gaming Attributes
Application	Fantasy
Cognitive strategies	Environment
Declarative knowledge	Conflict
Knowledge organization	Assessment
Adaptation	Action language
Automaticity set	Rules/goals
Compilation	Adaptation
Origination	Challenge/surprise
Psychomotor	Immersion
Attitudinal valuing	Human interaction
Internalizing values	Feature control
Motivation	
Organization	
Receiving/responding phenomena	

The objective of their research was to learn more about how gaming attributes, such as feature control, challenge and / or immersion, influence specific training outcomes. In particular, their paper describes the methods, measures and results of experiments investigating the impact of feature control on three training outcomes: *declarative knowledge, motivation and application.* Although an unintended by-product of their research, the authors found that previous experience / training along with greater feature control resulted in a significant increase in training benefit, leading the researchers to conclude that experience level should be one of the factors considered in design of serious games for training. Certainly, not all virtual environments share all the attributes of a gaming environment, but attributes such as immersion, human interaction, conflict and adaptation are often implicit components of virtual environments used for training.

Instructional Design

Instructional design is predicated on a number of factors; training needs analysis, training transfer and knowledge acquisition and retention. Training needs analysis, deciding who and what should be trained, leads to establishment of training objectives and formulation of criteria for meeting the objectives. Inherent in training needs analysis is determination of organizational goals, available resources, constraints and organizational support. Once organizational factors are considered, instructional designers usually conduct a job / task analysis, identifying the desired knowledge, skills and attitudes (KSA) required. Often, instructional designers conduct a cognitive task analysis (CTA) to understand the mental processing / mental requirements for performing the tasks. Using knowledge elicitation techniques, cognitive scientists construct mental models, cues for decision-making, potential cues for training scenario development, as well as performance and feedback protocols (Salas & Cannon-Bowers, 2001).

Instructional Design Considerations for Crew Resource Management

Effective crew level performance is contingent on individual crew member proficiency. Many previous studies focus on individual proficiency alone, while others have investigated team proficiency and the complex dynamics associated with collective performance. For example, Seamster and Kaempf (2001) address the topic of crew resource management (CRM) training for airline crews, citing Katzenback and Smith's (1993) observation that "a key distinguishing element of the team is that the team has both individual-participant and team accountability" (p.13). Further, the authors present Guzzo and Dickson's (1996) definition of team as the preferred organizational psychology term for "small groups that work together on a common set of tasks" (p.13). The terms team and crew are used interchangeably throughout the literature, so for the purposes of this study any reference made to teams also applies to crews. Throughout the literature surveyed in (Salas et al., 2001), some common CRM skills emerge. These skills include situational awareness (Seamster and Kaempf, 2001), communication (Kanki & Smith, 2001), teamwork (Smith-Jentsch et al., 2001), coordination and planning (Seamster & Kaempf, 2001), and decision making (Flin & O'Connor, 2001).

Instructional Design Considerations for Knowledge Acquisition and Retention

Although the objective of training is skills acquisition, equally, if not more important is the retention of those skills. Significantly relevant to this study, a report by

Wetzel, Konoske and Montague's (1983), cited in Ricci et al. (1996), found that anti-submarine warfare sensor operators' performance degraded between successive training events; in this case, the 25-day interval between basic acoustic analysis course and the follow-on operator course. Based on that and other reports on knowledge decay, the authors postulate the need for "training methodologies or enhancements that aid in knowledge retention" (p.296). Their research presents six computer-based gaming attributes they think contribute to training effectiveness: active participation, immediate feedback, dynamic interaction, competition, novelty and goal direction. Based on their investigation, the researchers conclude that computer-based gaming has the potential to enhance learning and retention. The implication for this research is the potential not only for improving the effectiveness of high fidelity simulators and live training while at home station, but also for reducing skills atrophy / degradation while ASW crews are deployed.

During and After Training

The building block approach to training effectiveness and evaluation, as depicted in Figure 1, illustrates the capstone type relationship of learning outcomes. Once the foundations of training needs analysis, task analysis, training system and instructional design have been laid through careful thought and according to theoretical principles, training is delivered. Both during and after training delivery, certain outcomes are evident. These include learner reactions, attitudes, progressive learning and the two performance types, training and transfer. The following sections address these topics and their impact on training and transfer performance.

Reactions

In presenting their integrated model for training effectiveness and evaluation (IMTEE), Alvarez et al. (2004) observe that the first outcome from training is the learner's reaction. Reactions can include the learner's attitude on the utility, relevance, value and whether or not the training might be transferable to their specific domain. Presenting *affective reactions* as an example, the authors contrast Kirkpatrick's (1976) view that positive *affective reactions* are directly related to learning versus Alliger et al.'s (1997) meta-analysis that found no correlation between *affective reactions* and learning or any other training outcome, but others share a different perspective as discussed in the following section.

Reactions: Involvement, Immersion, Presence and Learning

Witmer and Singer (1998) suggest that reactions to training system attributes and their ability to induce *involvement*, *immersion* and *presence*, play a prominent role in learning. Witmer and Singer (1998) define *presence* as the sense (subjective experience) of being in one place while physically located in another, contending that *presence* is dependent on two factors: *involvement* and *immersion*, both of which are required to generate *presence*. Similarly, Bernatovich (1999), in his master's thesis at the Naval Postgraduate School, defines *presence* as the sense of being there; his example: the combination of spatial knowledge (such as landmark recognition), procedural knowledge and survey knowledge. Witmer and Singer (1998) define *involvement* as the mental (psychological) "state experienced as a consequence of focusing one's energy and attention on a coherent set of stimuli" (Witmer & Singer, 1998). Further, the authors define *immersion* as a mental (psychological) "state characterized by perceiving one's self to be enveloped by, included in, and interacting with an environment" (p. 227).

Equally important to the training system's ability to induce presence-related factors, is the learner's willingness (attitude) and ability to experience *involvement*, *immersion* and *presence* within a VE. In their contribution to Massachusetts Institute of Technology's (MIT) Presence periodical, Witmer and Singer (1998) report on the relationship between a user's sense of *presence* and the training effectiveness of virtual environments (VE). The authors postulate that *presence* requires directed attention and is based on interactions between sensory stimulation, environmental factors and users' tendencies to become involved. Four factors the authors believe contribute to a sense of *presence* are: *control*, *sensory*, *distraction* and *realism*.

Control Factors – Witmer & Singer (1998) contend that the degree, immediacy and mode of control a learner can exercise over the virtual environment (relative to that exercised in the real-world) is directly related to the degree of *presence* the learner experiences.

Sensory Factors – Witmer and Singer's (1998) research suggests that the sensory factors tied to presence are primarily visual (e.g., environmental richness and movement perception), however they agree that the addition of other senses (e.g., aural and kinesthetic) also contribute to experiencing *presence*.

Distraction Factors – Witmer & Singer (1998) point to Held and Durlach's (1992) assertion that unnatural, clumsy, artifact-laden interface devices interfere with the direct and effortless interpretation of (and interaction with) a VE and hence diminish *presence*. They also mention the potential contribution of relative isolation engendered by the use of head-phones or head mounted displays to the sense of *presence*. Another very important factor is the learner's willingness/ability to ignore distractions.

Realism Factors – From a sensory perspective, Witmer and Singer (1998) list scene realism attributes (content, texture, lighting, resolution and field of view, to name a few) along with real-world consistency and post-VE related disorientation as significant influences on a learner's experience of *presence*.

Citing previous investigations by Witmer and Singer (1994) and Bailey and Witmer (1994), the authors acknowledge mixed results from efforts to establish direct correlation between *presence*, learning and performance within a VE vis-à-vis learning and performance in the real world. However, they cite apparent coincidence of factors affecting presence that also affect learning and performance and conclude with their assertion that "manipulating factors that increase *presence* will increase learning and performance" (p. 239). In their article published in the June 1998 MIT Presence periodical, Witmer and Singer (1998) present tandem data collection instruments they used to support their assertions concerning the relationship between *presence* and learning. Their Immersive Tendencies Questionnaire poses a series of questions designed to elicit a learner's propensity to become immersed. The Presence Questionnaire presents a series of questions eliciting the learner's attitudes toward the VE-delivered training and their related subjective self-reported (affective reaction) sense of *presence*.

Other interesting perspectives on presence include Morris et al.'s (2004) concept of stress inducement as a potentially beneficial element of presence, stating "Contextually correct simulations generate higher levels of stress during training, thereby reducing the user's level of stress in actual operational performance" (p. 138). Two other approaches to presence provide additional perspective on the topic. Alexander et al. (2005) linked *presence* with a simulation's fidelity in terms of two separate perspectives – physical (looks, sounds, feels like) and functional (acts like) while Bell and McNamara (2005) touted the importance of cognitive fidelity because it elicits decision-making and team behaviors that match real-life mental processes.

Although these topics are of great interest and worthy of further in-depth research, this study does not explicitly investigate involvement, immersion, stress inducement and presence. Witmer and Singer's (1998) presence-related data collection instruments would undoubtedly be valuable in assessing whether a training system induces sufficient *involvement*, *immersion* and *presence* to deliver contextually appropriate practice.

Post-Training Attitudes

While Kirkpatrick (1976) classified changes in post-training attitude as a component of learning, Tannenbaum et al. (1993) and Kraiger (2002) treat attitude change as a separate category, indicating possible outcomes such as increase in self-efficacy, motivation and organizational commitment (Alvarez et al., 2004). Citing eleven sources, Alvarez et al. (2004) link self-efficacy to cognitive learning, training performance and transfer performance. Although there seems to be a strong relationship between the factors, the reports do not indicate the relationships' direction, so the authors conclude that the relationship may be reciprocal. Alvarez et al. (2004) found only a couple investigations on post-training changes in motivation linking training effectiveness models with motivation to learn and motivation to transfer (Baldwin & Ford, 1988; Holton, 1996; Holton & Baldwin, 2000; and Tannenbaum et al., 1993).

During the U.S. Navy Recruit Training Command study presented in Chapter 3, Hussain et al. (2009) address the implicit objective of increased learner motivation and organizational commitment, but the researchers did not explicitly measure either of those factors, nor did they administer an instrument designed to discover the recruits' self-efficacy. In this effort's study of helicopter aircrew members, motivation and organizational commitment were not factors of interest. However, the adapted Mission Awareness Rating Scale (MARS) data collection instrument used in this study contains elements of self-efficacy as they relate to situation awareness.

Cognitive Learning

Cognitive learning is characterized by increase in knowledge, change in mental structure of knowledge or both (Alvarez et al., 2004; Tannenbaum et al., 1993). Cognitive learning encompasses structural knowledge, problem-solving, self-knowledge and executive control (Kraiger, 2002). Alvarez et al. (2004) reassert that post-training self-efficacy and cognitive learning have a reciprocal relationship and that cognitive learning impacts *training performance* and *transfer performance* citing Alliger et al. (1997), Kraiger et al. (1993) and Tannenbaum et al. (1993). As depicted in Figure 1, *cognitive learning* leads to *training performance* and *transfer performance*. The following two sections should clarify any misunderstandings surrounding the concepts of *training performance* and *transfer performance*.

Training Performance

Training performance has to do with how well the training audience performed during the training. Training performance evaluates whether and to what extent learning occurred while using the training system. In addition to instructor assessments of performance, metrics associated with training performance should also indicate what knowledge was acquired, what skills were developed or enhanced and what attitudes were changed (Kraiger et al., 1993). Cohn et al. (2008) proffer the potential of virtual environments to develop complex cognitive task behaviors; for example: situational awareness and cognitive expertise. In addition to these two factors, this dissertation investigates and reports on training performance related to several other complex task behaviors, including two mission-related tasks and leadership / teamwork and communications behaviors.

Transfer Performance

Transfer performance indicates how well the training audience performed on the job, post training. Training Transfer Evaluation measures the learner's ability to perform learned skills in the operational environment (Kirkpatrick's Level 3); requires establishing relationship between training objectives and operational requirements through defined process and outcome measures. Useful tools include a taxonomy derived from documentation review, task analysis and interviews. The objective is a set of metrics trainers can use to predict how effectively a training solution transfers to operational performance. Training transfer performance is the gold standard of training (Alvarez et al., 2004). Numerous studies have investigated training efficacy and transfer of training, but Champney et al. (2006) developed methods to adapt transfer of training methods to enable evaluation within operational constraints.

The concept of transfer of learning is foundational for determining training effectiveness. According to Perkins and Salomon (1992); "Transfer of learning occurs when learning in one context or with one set of materials impacts on performance in another context or with other related materials" (p.1). Ideally, transfer of learning is the desired effect of instruction. However, the authors contrast transfer with ordinary learning, citing a case of a student demonstrating understanding of grammar on an English test, but not applying the desired skill in daily life. Four related concepts are particularly applicable to this study. They are positive, negative, far and near transfer.

Positive Transfer – Learning in one context improves performance in another

Negative Transfer – Learning in one context does not accrue to or actually has a negative impact in another context

The authors downplay the impact of negative transfer in education in general. However, there is a very real (and potentially catastrophic) impact in the realm of military training.

Near Transfer – Skills learned in training context transfer directly to a similar performance context; e.g., a person who has learned to drive a car can readily adapt those skills to driving a light truck.

Far Transfer – Skills acquired in a training context may be applied to seemingly dissimilar contexts; rather than relying on rote learning for every possible use case, far transfer is the ability to generalize and apply principles and strategies to new and unique situations (Wallace, 1997). To continue the driving analogy, a person who previously learned to drive a car can adapt those basic skills to operate the vehicle under a variety of conditions (e.g., driving on highways, in urban areas, during periods of inclement weather, etc.).

Learning Outcomes

Training is not an end in itself. It is a means to achieving some kind of learning outcome. Learning outcomes are the result of the preceding activity levels depicted in Figure 1. They include acquisition and retention of knowledge, skills and abilities. According to Peter Ewell (2001), "...learning outcome...is...defined in terms of the particular levels of knowledge, skills and abilities... attained at the end (or as a result) of his or her engagement in a particular set of... experiences" (p.4). In the military context, learning outcomes equate to unit proficiency and readiness.

How do we determine whether and to what extent desired outcomes in terms of knowledge, skills and abilities have been achieved? Assessment of learning outcomes naturally follows delivery of training. One method of assessment, Bloom's taxonomy provides a framework for assessing learning outcomes: 1) cognitive outcomes deal with a learner's acquisition of declarative, procedural and strategic knowledge; 2) psychomotor outcomes include a learner's perceptual, response selection, motor and problem solving skills; 3) affective (attitudinal) outcomes involve factors such as a learner's task persistence in the face of adversity, physiological responses such as stress and targeted emotional responses (remaining calm under pressure) – Kraiger et al. (1993). The following section discusses theoretical foundations for performance measurement.

Theoretical Foundations of Performance Measurement

Generally, the purpose of performance measurement is to describe, evaluate / assess and diagnose performance with the goal of remediating poor or inadequate performance. Literature surrounding the topic of performance measurement abounds. This section summarizes the review of theory and best practices Salas et al. (2009)

contributed in their discussion on performance measurement in the context of simulation based training (SBT). The authors value SBT for its ability to provide practice opportunities in virtual environments that represent important features similar to those experienced in the "real world" environment. Their study presents theory to help answer the question of "What" to measure and best practices to address "How".

Individual Performance Theories

This section introduces individual performance theories as presented in Salas et al. (2009). Theories presented include frameworks of learning outcomes, a general theory of performance, human information processing (HIP) and expertise.

Frameworks of learning outcomes

Learning outcomes are "persistent states that make possible a variety of human performance" (Gagne, 1984, p. 377). Citing research by Bloom (1956), Gagne (1984), Kraiger, and Ford and Salas (1993), the authors seem to opine that classes of learning outcomes are diverse and that theories can be generalized within, but not across categories. Their example is that theories of motor performance are not helpful in understanding coordination or decision-making tasks. The researchers offer Kraiger et al.'s (1993) framework that features three general heterogeneous categories of outcomes:

- Cognitive outcomes – declarative knowledge, knowledge structure, cognitive strategies
- Skill-based outcomes (psychomotor) – procedural; skill compilation; automaticity
- Affective outcomes – attitudes, motivation

In their discussion on application of learning outcome theory, Salas et al. (2009) imply a certain hierarchy amongst the outcomes. Presenting a novice's acquisition of driving skills, the initial focus for targeted training might involve coordination of motor skills required, such as the manipulation of steering, shifting, braking and acceleration controls. As the learner masters the basic psychomotor skills, she begins to develop cognitive strategies to deal with operating a vehicle under more challenging conditions (such as inclement weather, high density traffic, etc.), which in turn could be the next targeted competency.

General theory of performance

Campbell, McCloy, Oppler, and Sager (1993) present a high level theory describing performance as a multidimensional construct comprising distinct components that can be accounted for by three primary factors (Campbell et al., 1993):

Declarative Knowledge – facts, principles, goals, self-knowledge associated with a given task.

Procedural Knowledge and Skill – cognitive, psychomotor, physical, self-management and interpersonal task-related skills.

Motivation – choice behaviors; e.g., choosing to expend effort, the level of effort expended and persistence of the effort.

Human information processing (HIP)

HIP theories focus on cognitive learning and address how people internally manipulate information from the environment and how the transformed information determines performance. Although there are numerous HIP models, Naturalistic Decision

Making (NDM) models have come to the forefront as researchers have expressed interest in contextually appropriate approaches. The Recognition Primed Decision (RPD) model (G. Klein, 1993) in particular, has gained in popularity.

The RPD model views the decision-making process in terms of two sub-processes: 1) cue and pattern recognition, and 2) mental simulation of courses of action. As a decision maker senses cues in the environment, he seeks to match the pattern with similar past experiences, mentally developing expectancies, goals and potential courses of action. If there is a match with previous experiences, the decision maker mentally simulates viable courses of action and either adopts, modifies or rejects them as solutions. Situation assessment and situation awareness (SA) play a crucial role in contextually appropriate decision making, thus Endsley's (1997) three levels of SA are essential inputs to the decision making process. The three levels of SA are:

- Perception of environmental elements
- Comprehension of the current situation
- Projection of future status.

Unless articulated verbally, internal thought processes (e.g., pattern recognition, internal strategies, and expectancies) are not directly observable. Therefore, cognitive performance measurement is normally inferred through mechanisms such as knowledge tests. Salas et al. (2009) suggest that SBT scenarios can be an excellent way to elicit information on behaviors that rely on "components of covert performance" such as RPD pattern recognition or mental simulation sub processes. This technique, known as event-based measurement, is discussed in the section on best practices.

Expertise

Superior performance is a function of task specific, specialized psychological mechanisms developed during extended practice activities. Although expertise research has historically been accomplished in task specific domains such as chess, sports and music, Salas observes that research into expertise developed in narrowly defined tasks has extended understanding of expertise in less well-defined tasks and researchers have begun to study more cognitively complex tasks, e.g., the expertise study involving fire fighter performance from which the RPD model emerged.

Salas and colleagues present a framework of expertise components encompassing a set of characteristics gleaned from over a dozen research efforts. Key concepts include:

- Expert performance is a function of extended experience within a domain (Bordage & Zacks, 1984)
- Simulation provides the means to study expertise and to develop expertise through increased <u>opportunities to practice</u> (Ericsson & Ward, 2007)
- Intentional practice yields:
    - A larger, more organized knowledge base (Chi, Fletovich & Glaser, 1981)
    - Memory skill (Ericsson & Kintsch, 1995)
    - Improved pattern recognition and problem representation (Glaser & Chi, 1988; Simon & Chase, 1973; Zeitz, 1997)
    - Improved self-monitoring and automaticity (Lesgold et al., 1988; and Moors & De Houwer, 2006).

Team Performance Theories

Introducing their discussion on team performance, Salas et al. (2009) present their rationale for the breadth of their review. They observe that SBT enables targeted practice of dynamic interactions – key to developing teamwork skills. Teamwork is complex, involving multiple personnel performing individual tasks with the added element of coordinating their individual actions, thus justifying the need for good theory as a basis for developing performance measures. As with targeted practice, targeted performance measurement, processes and outcomes should be contextually relevant and customized to the needs of the particular team (Kendall & Salas, 2004; MacBryde & Mendibil, 2003).

Several research groups have developed frameworks for team performance measures. One specific example is Smith-Jentsch et al.'s (1998) proposed four dimensions of team performance to consider when developing a measurement system: team processes, individual processes, team outcomes and individual outcomes. Four team processes that have been empirically proven to be related to performance outcomes are offered for consideration when observing team performance: information exchange, communication, supporting behavior, and team leadership.

Building on the premise that performance measurement needs to have a solid theoretical basis, Salas and his colleagues present what they view to be the major contributions to the field, including the concepts of Input, Process, Output (IPO) models, shared mental models, team adaptability, the "big five" of teamwork, and macro cognition / shared team cognition.

IPO Models

Input - Process - Outcome (IPO) models and theories characterize relationships between input variables; e.g., individual and team characteristics; process variables – shared cognition, leadership, communication, coordination, decision making and back-up behavior; and outcome variables – performance outcomes, productivity and satisfaction (Gersick, 1988; Hackman, 1983; Marks, Mathieu & Zaccaro, 2001; Salas, Dickinson, Converse & Tannenbaum, 1992; and Salas, Priest & Burke, 2005) as cited in Salas et al. (2009).

Tannenbaum, Beard and Salas (1992) developed and evaluated an integrative IPO framework that segments input variable into four classes (individual characteristics, team characteristics, work structure, and task characteristics); addresses team process variables and team interventions as throughput components; and identifies output variables in terms of individual changes, team changes, and team performance.

Shared Mental Models

Salas et al. (2009) describes shared mental models as "shared knowledge held between team members" (p. 333). According to Johnson-Laird (1983), shared mental models can be construed as knowledge structures involved in the integration of information and the comprehension of a given phenomenon. Another definition, a representation of information held by more than one person – team members that facilitates interactions between team members is presented by Klimoski and Mohammed (1994). From a practical standpoint, shared mental models allow team members to predict the needs (e.g., information, material resources, present workload) of their team mates (Mathieu et al., 2000).

Adaptability

The team adaptation model was developed by Burke et al. (2006) and takes an IPO approach to address the gap in research related to team performance over time. The centerpiece of the theory is the four-phased Adaptive Cycle.

Phase 1 – Situation Assessment: cue recognition / ascription of meaning to cue patterns

Phase 2 – Planning: develop and decide courses of action

Phase 3 – Execution: team coordination – mutual monitoring, communication, back-up behavior, leadership

Phase 4 – Team Learning: performance evaluation.

Shared Mental Model and Team Situational Awareness are integral to each phase and are the basis for each subsequent phase.

Big Five Model of Teamwork

In an effort to combine disparate thoughts on teamwork and team effectiveness, Salas, Sims, and Burke (2005) presented a concept they labeled the “big five” of teamwork. The “big five” are Leadership, Adaptability, Mutual Performance Monitoring, Back-up Behavior, and Team Orientation. The five factors (discussed in more detail below) require three coordination elements – shared mental models, closed loop communication and trust – which are needed for effective teamwork.

Leadership – the concept of shared leadership focuses on team members shifting leadership responsibilities as task demands change. Rather than supplanting formal leadership structure, this technique enables team members to exercise their own unique expertise, knowledge and skills to maximize team performance.

Adaptability – as discussed in the previous section, this element enables the team to adjust to the changes in a dynamic environment.

Mutual Performance Monitoring – maintaining awareness of other team members' activities, while performing one's own responsibilities (McIntyre & Salas, 1995). Mutual performance monitoring requires team members to have a shared mental model and a common understanding of what "right" looks like. When applied in a climate of mutual trust, this principle is a valid way of elevating levels of team performance.

Back-up Behavior – according to Porter et al. (2003), back-up behavior is "the discretionary provision of resources and task-related effort to another… (when) there is a recognition by potential back-up providers that there is a workload distribution problem in their team" (pp. 391-392) . In other words, when engaged in mutual performance monitoring, one team member "picks up the slack" of another if an overload is detected (Marks, Zaccaro & Mathieu, 2000).

Team Orientation – the tendency to coordinate, evaluate and use task inputs of other team mates and to engage in patterns of behavior that improve team and individual performance (Driskell & Salas, 1992).

Coordinating Mechanisms that support the "big five" are shared mental models, closed loop communication and mutual trust. Shared mental models are discussed earlier in this section. Closed loop, a type of explicit communication, consists of a message initiated by a sender; receipt, interpretation and acknowledgement by the receiver and follow-up by the sender to ensure appropriate interpretation (McIntyre & Salas, 1995). Mutual trust is the "shared perception…that individuals in the team will perform particular actions important to its members and... will recognize and protect the rights and interests of all the team members engaged in their joint endeavor" (Webber, 2002, p. 205).

Macro-cognition/Shared Cognition

Investigators, such as Hinsz, Tindale and Vollrath (1997) and others have viewed groups / teams as information processing systems. As such, teams engage in team-level macro-cognitive processes and cognitive activities such as assessing situations, making decisions, designing and monitoring plans and gathering knowledge, thus justifying the need to understand and measure cognition at the team level.

There are three categories of measures to assess team cognition (Cook, Salas, Cannon-Bowers & Stout, 2000). All three share the underlying assumption that team knowledge is a collection of individual-level knowledge. The three categories are cumulative, each building on the foundation of the previous. The categories and a general description follow:

Elicitation Methods – observation, interview, mapping, protocol analysis

Team Metrics – quantifies information gathered through knowledge elicitation; difficult to analyze in terms of accuracy, team member overlap, and aggregation of individual level data to team level

Aggregation – methods include averaging individual scores, or taking, median, maximum or minimum value.

Performance Measurement Methods

Offering Campbell's (1990) thoughts; "Performance is not the consequence or result of action; it is the action itself" (p. 174), Salas (2009) contends that although outcomes are important, the process is equally if not more important and that improved process improves outcomes.

Qualitative Methods – what to measure – are best for understanding processes. Salas (2009) offers three examples: protocol analysis, critical incident technique and concept mapping.

Protocol Analysis uses verbal reports as indicators of cognition (Ericsson & Simon, 1993). Shadbolt (2005) offers two categories:

- on-line – participant verbalizes thoughts during (concurrent with) task performance
- off-line – participant verbalizes thoughts after performance, perhaps while viewing video recording of the performance session.

While the on-line technique is more naturalistic, off-line interferes less with task performance. The protocol analysis process evolves through three stages: data collection – verbalizations recorded; analyst identifies meaningful units representing separate mental processes; examination of explicit and implicit content. Cognitive walkthrough is a related technique in which participants talk through how they would perform the task, without doing it (Bainbridge & Sanderson, 2005).

Critical Incident Technique is a set of procedures developed by the U.S. Air Force during World War II to observe behavior in an effort to solve practical problems (Flanagan, 1954). A critical incident is defined as follows: intent of performer is clear to the observer, who can identify consequences of behavior. The technique features a flexible set of principles: observers are only required to make simple judgments, observers must be qualified, and there is an agreed-upon purpose to guide evaluation.

Concept Mapping is a knowledge elicitation method that measures structure, content and robustness of the individual and team knowledge, resulting in a graphical representation of their inter-relationships. Maps can be generated through questionnaires and interviews, analysis of pre-existing data or through participants generating structure and content.

Quantitative Methods – How to measure.

- Behaviorally anchored rating scales (BARS) consist of a series of brief descriptions characterizing levels of performance. Each description corresponds to a scale value representing excellent, acceptable and poor performance. The purpose of BARS is to facilitate observer ratings of individual and team performance. Reports by Flanagan (1954), Fowlkes et al. (1994) and Morgan et al. (1989), indicate that BARS are often developed with the assistance of Subject Matter Experts (SME) using the Critical Incident Technique.
- Behavioral observation scales (BOS) are similar to BARS, but assess typical rather than isolated incidents of performance. BOS assists observers to capture the frequency of team behaviors of interest by employing range of values similar to a Likert scale. One example of how the scale is used: during a training event, an observer assigns a numerical score (e.g., 0 = never, 9 = always) describing how often the team engages in a behavior.

- Communication analysis, according to the authors, is the best method to measure team processes. Although observers measure the simple frequency of communications, Salas et al. (2009) cites several research efforts that present mixed results in their attempts to correlate frequency of communication with outcomes. Content analysis aims to measure elements of verbal exchange related to team processes. The technique looks at types of communication, such as commands, questions, acknowledgments, etc. (Kanki, Lozito & Foushee, 1989), analyzing the exchanges using matrices of initiating and responding speech (Kanki et al., 1989).
- Event-based measurement is an observation technique that uses contextually relevant scenarios containing cues designed to elicit important team behaviors that occur relatively infrequently during the normal course of operations/training. Two techniques mentioned are Targeted Acceptable Responses to Generated Events (TARGETs; Fowlkes et al., 1994) and an event-based approach to training (EBAT) tool known as Simulation Module for Assessment of Resident's Targeted Event Responses (SMARTER). SMARTER is an example of EBAT methods being employed in the health care field. Within a simulation scenario, EBAT compares learning objectives and desired competencies with performance measures. Salas et al. (2009) observe that EBAT is well suited for measuring team performance, since the technique can be used during practice.
- Structural knowledge assessment addresses both individual (expertise) and team (shared mental model) performance. One approach, known as Pathfinder, is similar to concept mapping. The technique represents structural properties of knowledge as a

network. The knowledge concepts are portrayed as nodes and degree of relatedness between concepts as links or edges.

- Self-report measures are generally elicited through the use of questionnaires to determine individual and team level cognitive and affective states and processes. The authors cite several studies that have attempted to translate results of individual self-reports into team-level characteristics, such as team orientation (Driskell & Salas, 1992), collective efficacy (Tagger & Seijts, 2003) and team cohesion (Gully, Devine & Whitney, 1995).
- Automated performance recording and measurement assays to capture individual and team behaviors in a less intrusive manner that features fewer biases. Although the technique is useful for observing quantifiable behavioral responses, it does not readily lend itself to measuring cognitive processes. One potential exception to this deficiency has been presented in communication analysis literature, specifically Keikel et al. (2001). The authors mention a significant downside to this methodology; implementation cost.

Best Practices

To organize and present best practices, the authors use a framework for describing requirements for performance measurement they presented in Cannon-Bowers and Salas (1997). Table 7 summarizes their conclusions.

Table 7: Best Practices for Performance Measurement in Training Simulation (adapted from Salas et al., 2009)

Measurement System Requirement	Best Practice Performance Measurement Characteristics
Multiple levels of measurement	1. Captures multiple dimensions of performance at appropriate levels of analysis
	2. Event-based techniques are used to capture data at multiple levels of analysis
	3. Captures multiple measures from various sources
	4. A systematic plan integrates data from multiple measures
Address process as well as outcome	1. Captures the processes of performance
	2. Expert models of the task are used to compare and evaluate performance processes and outcomes
	3. The collection and transmission of objective measures is automated
Describe, evaluate and diagnose performance	1. Measures are descriptive of performance
	2. Performance can be compared to standards for desired levels of performance
	3. Diagnostic – provides insight into the causes of performance
	4. Allows for performance diagnosis to be partially or fully automated
	5. Allows for performance evaluation to be partially or fully automated
	6. Flexibility is designed into embedded measures (different measures can be substituted)
	7. Captures a broad spectrum of measures and the context of performance
	8. Observers are trained to high levels of reliability
	9. Observers use protocols
Basis for remediation	1. Performance measurement supports learning
	2. Allows for the automated and manual creation of AAR aids for training remediation and feedback
	3. Enables automated scaffolding and performance-based coaching
	4. Drives real-time corrective feedback
	5. Integrated with trainer controls and feedback generation

CHAPTER 3

RELATED RESEARCH

Several researchers have described how the addition of a low fidelity game-based simulation improved Novice Stage training performance both at the individual level and the team level. Among them, Hussain et al. (2009) report on the Recruit Training Command (RTC) / Office of Naval Research (ONR)-sponsored research and development of a low-cost training system for practicing cognitive and procedural skills previously presented in classroom instruction, but which were shown to be deficient during an end of training cycle evaluation in a full scale high fidelity simulator. Specifically, their paper focuses on the use of a low-fidelity game-based virtual environment to reinforce damage control skills Navy recruits had been exposed to in lectures, but had not had a chance to practice in context.

In their 2009 report, Hussein et al. observe: "The FCT (*Flooding Control Trainer*) targets students that are novices who have declarative knowledge about ships and basic damage control procedures. The FCT design assumes that students have been through formal training to acquire the declarative knowledge, but that they have not been required to apply this information or to draw on this information to solve problems. In other words, the recruits have classroom instruction, but very little hands-on experience with flooding control and related skills. The expectation is that learners may not fully understand the information and that they require practice in realistic contexts to build sound mental models" (p.9).

Technical curricula presented in the low-fidelity game-based VE offered a relevant backdrop for providing contextually appropriate opportunities to practice key cognitive skills which RTC instructors had identified as being deficient during previous exercises in the high-fidelity Battle Stations 21 simulator: i.e., situation awareness, communication protocols and decision-making. Additionally the serious game environment provided appropriate context to promote core values, and reinforce the culture of the service (Bloom's Affective Domain).

In an effort to improve the quality of training for every recruit, the U.S. Navy, invested approximately $80 million to develop a full-scale replica of an Arleigh-Burke class ship. The purpose was to give the recruits an opportunity to implement the knowledge and skills they had been presented (primarily through lecture) during the previous eight or so weeks of Navy Boot Camp. In spite of the Navy's investment in DoD's largest, most expensive simulator, recruit performance was deficient, especially in the areas of situation awareness, communications and decision-making. This gap in training performance led to the decision to develop a game-based trainer that would enable recruits to engage in targeted practice of the skills identified as deficient.

A multi-discipline (e.g., learning science, instructional curriculum, instructional systems and interactive multi-media design) research and development team comprised of industry and academia, designed, developed and validated a 3D immersive training game in just over 14 months. The ground rules established at the beginning of the study were: 1) researchers could have no more than one or two hours of the recruits' time and 2) the use of the game could cause no additional workload on the instructors.

The resultant immersive game-based virtual environment served as the basis for testing the researchers' hypothesis that game-based training should be able to create a strong positive learning effect with minimal instructor involvement. A usability study of

70 respondents produced very favorable results, with the majority (93%) of those questioned rating the game 7 or higher on a scale of 0-9. The validation study revealed a 50% reduction in the treatment group's decision-making errors; an 80% reduction in communication errors and 50% improvement in navigation and situation awareness.

From another perspective, Bowers et al. (2009) assert that a game-based "pre-experience" can lead to superior performance and fewer critical errors. Although the manuscript as of yet remains unpublished, the authors offer several insights relevant to this study, not covered in the companion RTC articles. Citing Salas and Cannon-Bowers (2003), the authors observe that although there are a few factors associated with effective training, the opportunity to practice may be the most overlooked element in training research and development. The authors present a compelling case for the use of gaming technologies as delivery mechanisms for "pre-training" intervention, exploring the training value of lower fidelity devices. Referring to Jentsch and Bowers (1998) literature review, they observe that game-based simulations can create reasonable manipulations of independent variables and that those manipulations produce behavioral changes similar to experience in the operational environment. Furthermore, they observe that manipulations using game-based systems generally elicit only targeted behaviors, an effect they refer to as *appropriate discriminate validity*.

Contrasting the preponderance of literature focused on the use of game-based tools "during training", Bowers et al. (2009) argue for increased investigation across the entire training spectrum, from pre-training to post-training as presented in Cannon-Bowers, Rhodenizer, Salas, and Bowers (1998). In particular, they contend that the use of high fidelity simulation / simulators presumes learners have the requisite knowledge and skills required for a training event, when in fact, that often is not the case. Their study involved 30 recruits at the Navy's Recruit Training Command, the purpose of which was to assess the feasibility of game-based training technology as pre-training intervention for Navy recruits (Novice stage of skills acquisition using the Dreyfus model). The authors concluded that the results of their study appear to support the hypothesis that recruits who experienced the game would perform better on targeted skills in the simulator than would a no-game control group.

Taking another viewpoint, Hussain et al. (2011) link individual cognitive readiness to improved team performance, noting that not only did performance of individuals acting independently on a real-world, near transfer task show broad-based improvement, but that team performance on several real-world transfer tasks also showed significant improvement after practicing as individuals.

Introducing the interdependencies of individual cognitive readiness skills and general teaming skills, the authors define cognitive readiness (in part) as the mental preparation an individual needs to establish and sustain competent performance. Citing the work of Bowers and Cannon-Bowers (2011), Hussain et al. (2011) point out that 1) the situation awareness cognitive skill is directly related to team monitoring; 2) the decision-making cognitive skill is an important component of team leadership; and 3) communication skills are fundamental to team coordination and shared mental models.

A couple of ancillary observations are worthy of note. First, the authors, referring to Glaser (1989), point out that novice learners often develop only a superficial understanding of targeted skills as a result of traditional lecture-based delivery. Second, they share the assertions of Cannon-Bowers & Bowers (2009) and Kolb (1984) that experiential learning may be more effective in achieving positive training outcomes, that experiential learning requires meaningful context (contextually appropriate) in order to develop new knowledge and skills, which in turn is more effective in developing accurate mental models of a complex environment.

The authors hypothesized that providing individual training (opportunities to practice) would improve the sailors' cognitive readiness and lead to improved team performance in the capstone Battle Stations 21 event. To validate their hypothesis, the researchers conducted two assessments. In the first, they evaluated the impact of playing the immersive virtual environment game on near transfer of damage control task skills. Their findings indicate significant improvement in cognitive readiness and task performance. The second assessment evaluated impact of individuals playing the game on team performance of both damage control and non-damage control tasks. Again, they observed significant team performance improvements as well as several indicators of enhanced team cognitive readiness.

Other Training Effectiveness Studies

Numerous other training effectiveness studies have been performed over the last two decades. Some involved PC-based simulations and some involved games; others virtual environments of various kinds, but nearly all of them focused on individual knowledge, skills and attitudes (KSA). Appendix D provides a representative sample of studies investigating training effectiveness and training transfer. Only a few of these training effectiveness / transfer studies specifically targeted crew training (e.g., Brannick et al., 2005; Proctor et al., 2004; Prince & Jentsch, 2001), and a couple addressed low-fidelity training devices (Prince & Jentsch, 2001; Giovanni et al., 2009).

According to Salas et al. (2009), "The value of simulation based training (SBT) primarily stems from its ability to provide practice opportunities in environments that replicate important features of the "real world" environment." (p. 329). The authors assert that since learning and performance are task dependent, the ability to transfer learned knowledge and skill depends on the similarity of the practice and transfer tasks. SBT must provide guided practice to ensure correct competencies are acquired. Therefore, training systems that enable dynamic interaction are a key contributor to developing team skills.

Several researchers including Salas et al. (2009), Hussain et al. (2009), and Bowers et al. (2009), have recognized a gap in the literature surrounding the introduction of modeling, simulation and serious games-based virtual environments into the learning environment to provide additional contextually appropriate opportunities to practice. Although a significant amount of research has been conducted concerning the use of simulation and games for education and training, little has specifically focused on transforming conceptual, declarative, procedural and metacognitive knowledge into cognitive and physical skills through simulation-supported contextually appropriate targeted practice. The literature research revealed no existing framework for delivering and assessing contextually appropriate opportunities for warfighter practice. To address that gap (at least partially), my research developed a Contextually Appropriate Practice – Delivery (CAP-D) framework and a Contextually Appropriate Practice – Assessment

(CAP-A) framework to assist in understanding and bridging the gap. A more detailed discussion on CAP-D and CAP-A frameworks follows.

Contextually Appropriate Practice (CAP) Framework

In an attempt to close the gap in the training research and development body of knowledge concerning the use of modeling, simulation and gaming technologies to deliver contextually appropriate opportunities to practice, this research generated a framework for contextually appropriate practice (CAP). The CAP framework consists of two parts: the Contextually Appropriate Practice – Delivery (CAP-D) framework and the Contextually Appropriate Practice – Assessment (CAP-A) framework.

The CAP-D, shown in Figure 2, is a multi-disciplinary framework which leverages learning theory, principles of instructional design and instructional system design, simultaneously capturing key elements of learner development from a theoretical standpoint while offering the flexibility to modify elements of the model to accommodate changes in modeling and simulation content delivery mechanisms, context and learner skills acquisition stage. Using learning theory as a foundation, CAP-D consists of three primary elements; content, delivery and outcome. The framework also includes three process activities; decide, deliver and assess. Embedded in the processes are the activities and detailed descriptions found in Figure 1 and elsewhere in the Training Effectiveness and Evaluation section of Chapter 2.

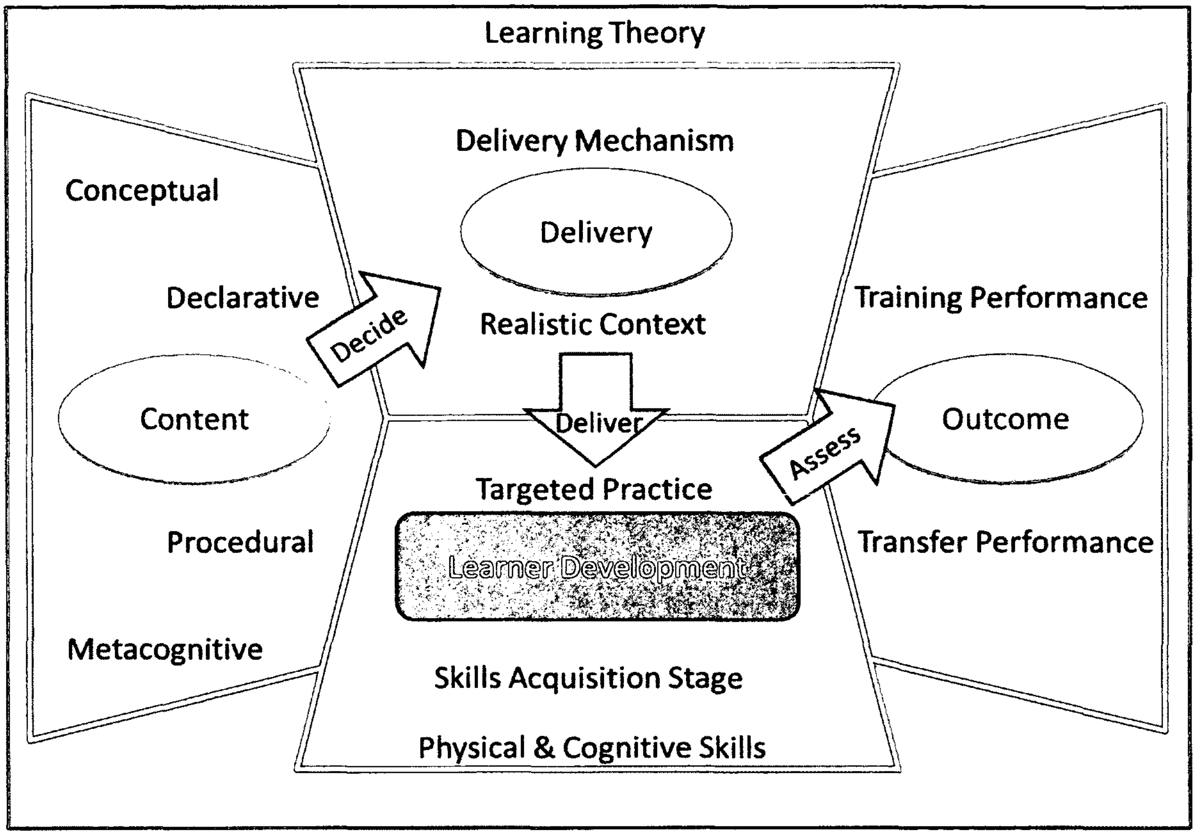

Figure 2: Contextually Appropriate Practice – Delivery (CAP-D) Framework for Delivering Warfighter Practice

When presented with a training opportunity, practitioners conduct a training needs analysis (Brown, 2002, as cited in Cohn et al., 2008) and a task analysis (e.g., Hale, 2006; Salas & Cannon-Bowers, 2001) to evaluate the type of knowledge content they wish to deliver; conceptual, declarative, procedural, metacognitive, or a combination of the four. Along with the training audience and their skills acquisition stage (Dreyfus & Dreyfus, 2004), other considerations include the physical and / or cognitive skills to be trained, the most appropriate context for addressing the skills and which ones are the best candidates for targeted practice. With these factors and the desired outcome in mind (Ewell, 2001), the next step is to decide the best delivery mechanism for the selected knowledge content.

Engel et al. (2009) present a method of selecting training system attributes (described in Chapter 2). Other factors worthy of consideration while deciding on the delivery mechanism, is the extent to which it aids in knowledge retention as well as knowledge acquisition. For instance, Ricci et al. (1996) suggest six computer-based gaming attributes that may contribute to training effectiveness: active participation, immediate feedback, dynamic interaction, competition, novelty and goal direction. Also worth considering, Witmer and Singer (1998) address involvement, immersion and presence factors in their contribution to VE learning.

The third process activity of CAP-D (Assess) necessitated development of a Contextually Appropriate Practice – Assessment (CAP-A) framework. The CAP-A framework (Figure 3) provides a means to evaluate the outcome of the targeted practice. Within the CAP-A construct, practitioners can use data collection instruments that capture quantitative and/or qualitative measures that describe performance in selected *warfighter mission task* areas along with other targeted learning objectives of interest.

Once the targeted learning objectives and tasks are identified, the CAP-A user can fill in the areas to be assessed along with the targeted tasks associated with those areas. Generally, a user will identify measures by which tasks are to be assessed and will select an appropriate data collection instrument. If an instrument that enables targeted data collection does not exist, the practitioner may modify an existing instrument or if necessary, create one. Chapter 2 provides the theoretical foundations of performance measurement, summarizing the Salas et al. (2009) review of theory and best practices, from which a practitioner may select theory and practice best suited for the intended use.

Contextually Appropriate Practice - Assessment (CAP-A) Framework					
Assessed Area	Task	Measure	Data	[illegible]	[illegible]
Warfighter Mission Task Performance	Mission Task 1	Measure used to assess Mission Task 1 performance	Type of data collected: e.g., Q-O, q-s	Data Collection Instrument	P/F/I
	Mission Task n	Measure used to assess Mission Task n performance	Type of data collected: e.g., Q-O, q-s	Data Collection Instrument	P/F/I
Targeted Area 1 e.g., Situation awareness, communications, decision-making	Targeted Area 1: Task 1	Measure used to assess Targeted Area 1: Task 1 performance	Type of data collected: e.g., Q-O, q-s	Data Collection Instrument	P/F/I
		Measure used to assess Targeted Area 1: Task 1 performance		Data Collection Instrument	P/F/I
	Targeted Area 1: Task n	Measure used to assess Targeted Area 1: Task n performance	Type of data collected: e.g., Q-O, q-s	Data Collection Instrument	P/F/I
		Measure used to assess Targeted Area 1: Task n performance		Data Collection Instrument	P/F/I
Targeted Area n User Defined	Targeted Area n: Task 1	Measure used to assess Targeted Area n: Task 1 performance	Type of data collected: e.g., Q-O, q-s	Data Collection Instrument	P/F/I
		Measure used to assess Targeted Area n: Task 1 performance		Data Collection Instrument	P/F/I
	Targeted Area n: Task n	Measure used to assess Targeted Area n: Task n performance	Type of data collected: e.g., Q-O, q-s	Data Collection Instrument	P/F/I
		Measure used to assess Targeted Area n: Task n performance		Data Collection Instrument	P/F/I

Q-O: Quantitative data collected objectively
q-s: Qualitative data collected subjectively

Figure 3: Contextually Appropriate Practice – Assessment (CAP-A) Framework

The CAP-A can serve as an artifact, providing a record of areas assessed, specific tasks assessed, measures by which those tasks are evaluated, data types, and data collection instruments used. The CAP-A framework also provides a means to assess the delivery mechanism's suitability to deliver contextually appropriate training. Evaluators can assign a Pass, Fail or Inconclusive score based on learner performance derived through the use of the data collection instruments. Once filled in, the CAP-A can serve as a formal deliverable.

Of course, the obvious question is: what to do with the assessment results? The CAP-D and CAP-A frameworks are iterative in nature. Depending on the results, an investigator can cycle back through the framework, and review and modify content, delivery mechanism, targeted tasks and/or data collection instruments. For example:

suppose while assessing a low fidelity VE, an investigator observes numerous failures in Situation Awareness tasks. This could be an indication that the delivery system does not sufficiently replicate important features of the "real world" environment (Salas et al., 1991) to deliver contextually appropriate practice. Such a finding may generate a closer inspection of the system to determine what, if any remedial actions can/should be taken to enhance environmental cues. Otherwise, one may need to seek a higher fidelity solution to deliver that specific training.

This study's thesis is that a framework for delivering contextually appropriate opportunities for warfighter practice can be applied to discover whether modeling, simulation and game-based virtual environments have the potential to improve individual performance for learners beyond the Novice Stage (e.g., Competent Stage) of skills acquisition, and that the conceptually appropriate practice (CAP) framework can be used to assess the potential of low fidelity virtual environments to provide targeted practice and to improve individual performance, not only during training in high-fidelity virtual environments (near transfer) but also in the live environment (far transfer).

This study investigates the relationship of technology and learning science, and seeks to determine whether a CAP framework can be applied to answer the following research questions:

1. Given that delivering targeted practice via Virtual Environments (VE) contributes to improved individual performance at the Novice Stage of skills acquisition (Hussain, et al, 2009), does VE-delivered targeted practice also contribute to improved individual performance at more advanced (e.g., Competent) stages of skills acquisition?

2. Does low-fidelity VE-delivered targeted practice contribute to improved individual performance during training in high fidelity simulators? During Live training?

Applying the CAP-D and CAP-A Frameworks to Novice Stage Learners

How do the Recruit Training Command (RTC) studies (Hussain et al., 2009; Bowers et al., 2009; Hussain et al., 2011) fit within the CAP-D framework? Using the RTC investigations as a pilot study, Figure 4 depicts the modifications to the baseline CAP-D framework to demonstrate the fit with the RTC studies. The modifications are highlighted in bold text. The researchers elected to address metacognitive and procedural content since learners had already been exposed to conceptual and declarative knowledge in a classroom environment. They selected a game-based virtual environment as the delivery mechanism based on its potential to create a strong positive learning effect with minimal instructor involvement (Hussain et al., 2009). According to the authors' report, researchers designed and built the game-based VE on fundamentals of learning theory, which in turn reinforced metacognitive organization and delivered the procedural content.

Observed shortfalls in recruit performance during previous training in the high-fidelity Battle Stations XXI simulator led the researchers to select three specific targeted skills; communications, situation awareness and decision-making. The multi-discipline team designed the game to weave instruction throughout the gaming experience and to vary the instruction based on student performance. The game-based VE delivered opportunities for Novice Stage learners to practice the targeted skills while performing tasks within the context of an overarching ship damage control mission scenario.

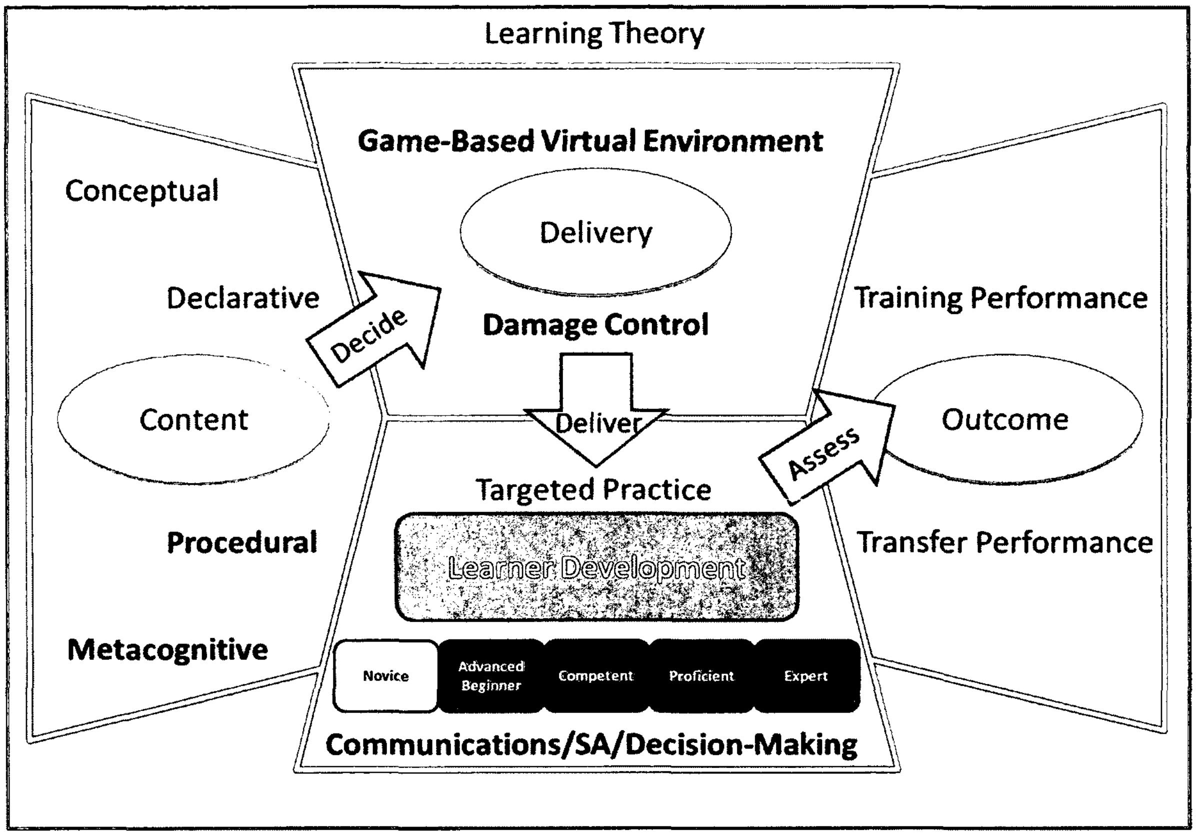

Figure 4: CAP-D Framework Applied to Learners at the Novice Stage of Skills Acquisition

How does the CAP-A framework apply to performance assessments conducted during the RTC studies? Once again, the RTC investigation served as a relevant pilot study of the CAP-A framework. Instructor assessments and empirical studies evaluated learner development and resultant training outcomes in terms of training performance and transfer performance. As shown in Figure 5, researchers used four data collection instruments to assess learner performance in three mission-related tasks while observing and recording results of 12 tasks associated with Communications, Situation Awareness, and Decision-Making.

[illegible]	[illegible]	Measure	Data	Instrument	[illegible]
Mission Task Performance	Flooding Magazine Compartment	Behaviorally Anchored Rating Scales (BARS) targeting Situation Awareness, Communications and Decision-Making	Instructor assessed score (Q-O)	Observer Form for Individual Study	P
				Flooding Magazine Compartment Scenario Observer Form	P
	Bridge Watch	Behaviorally Anchored Rating Scales (BARS) targeting Communications	Instructor assessed score (Q-O)	Bridge Watch Scenario: Bridge Talker Observer Form	N/A
				Bridge Watch Scenario: Conning Officer Observer Form	N/A
	Roving Security Watch	Behaviorally Anchored Rating Scales (BARS) targeting Situation Awareness, Communications and Decision-Making	Instructor assessed score (Q-O)	Roving Security Watch Scenario Observer Form	N/A
Situation Awareness	Recognize Abnormal Condition	Behaviorally Anchored Rating Scales (BARS) targeting Situation Awareness	Instructor assessed score (Q-O)	Flooding Magazine Compartment Scenario Observer Form	P
	Assess Flooding Situation	Behaviorally Anchored Rating Scales (BARS) targeting Situation Awareness	Instructor assessed score (Q-O)	Flooding Magazine Compartment Scenario Observer Form	P
	Recognize Shipboard Navigation Cues	Behaviorally Anchored Rating Scales (BARS) targeting Situation Awareness	Instructor assessed score (Q-O)	Observer Form for Individual Study	P
				Roving Security Watch Scenario Observer Form	N/A
	Anticipate Consequences of Actions	Behaviorally Anchored Rating Scales (BARS) targeting Situation Awareness	Instructor assessed score (Q-O)	Flooding Magazine Compartment Scenario Observer Form	P
Communications	Recognize when Situation Warrants Communication	Behaviorally Anchored Rating Scales (BARS) targeting Communications	Instructor assessed score (Q-O)	Observer Form for Individual Study	P
				Flooding Magazine Compartment Scenario Observer Form	P
				Bridge Watch Scenario: Bridge Talker Observer Form	N/A
	Report Appropriate Information	Behaviorally Anchored Rating Scales (BARS) targeting Communications	Instructor assessed score (Q-O)	Observer Form for Individual Study	P
				Flooding Magazine Compartment Scenario Observer Form	P
	Report Information Accurately	Behaviorally Anchored Rating Scales (BARS) targeting Communications	Instructor assessed score (Q-O)	Observer Form for Individual Study	P
				Flooding Magazine Compartment Scenario Observer Form	P
	Repeat Back Accurately	Behaviorally Anchored Rating Scales (BARS) targeting Communications	Instructor assessed score (Q-O)	Bridge Watch Scenario: Bridge Talker Observer Form	N/A
				Bridge Watch Scenario: Conning Officer Observer Form	N/A
Decision-Making	Maintain Watertight Integrity	Behaviorally Anchored Rating Scales (BARS) targeting Decision-Making	Instructor assessed score (Q-O)	Observer Form for Individual Study	P
				Flooding Magazine Compartment Scenario Observer Form	P
	Follow Safety Protocols	Behaviorally Anchored Rating Scales (BARS) targeting Decision-Making	Instructor assessed score (Q-O)	Observer Form for Individual Study	P
	Take Proper Actions to Combat Flooding	Behaviorally Anchored Rating Scales (BARS) targeting Decision-Making	Instructor assessed score (Q-O)	Observer Form for Individual Study	P
				Flooding Magazine Compartment Scenario Observer Form	P
	Use Compartment Identifiers to Navigate Ship	Behaviorally Anchored Rating Scales (BARS) targeting Decision-Making	Instructor assessed score (Q-O)	Observer Form for Individual Study	P
				Roving Security Watch Scenario Observer Form	N/A

Q-O: Quantitative data collected objectively
q-s: Qualitative data collected subjectively

Figure 5: CAP-A Framework Applied to Learners at the Novice Stage of Skills Acquisition

The four data collection instruments incorporated behaviorally anchored rating scales (BARS), enabling evaluation of individual training performance and outcomes. As reported in Hussain et al. (2009), the validation study revealed a 50% reduction in the treatment group's decision-making errors; an 80% reduction in communication errors and 50% improvement in navigation and situation awareness. Extending RTC-related study observations, Hussain et al. (2011) link individual cognitive readiness to improved team performance, demonstrating improvement in near transfer individual task performance as well as far transfer team performance after practicing as individuals.

Applying the CAP-A framework to assess the suitability of the game-based delivery mechanism yields some interesting observations. As reported in Hussain et al. (2011), RTC investigators assessed training results in three separate contexts: Flooding Magazine Compartment Scenario (directly related to the game-based scenario), Bridge Watch Scenario, and Roving Security Watch Scenario. Although the game-delivered training addressed only one of the scenarios, individual performance in the targeted skills improved across all three scenarios, supporting the authors' observation of both near and far transfer performance.

In the absence of primary data, the CAP-A framework captured a subjective assessment of the game-based delivery mechanism based on RTC empirical results in all 12 of the assessed learner tasks. One can infer from the results that the improvement in learner performance indicates the suitability of the game-based delivery mechanism to deliver contextually appropriate training to learners at the Novice Stage of skills acquisition, thus the game-based delivery mechanism earned a Pass grade.

CHAPTER 4

MILITARY TRAINING DOMAIN

Commanders are constantly seeking ways to ensure that military members get the most "bang for their buck" when participating in the few virtual and even fewer live training opportunities available to them. US Navy helicopter aircrews face a number of challenges when it comes to maintaining critical mission-related competencies. According to then Commander, Naval Air Force U.S. Atlantic Fleet, current operational tempo results in squadrons spending on average, only 40% of their time in port, thus severely limiting available training time (O'Hanlon, 2010). Although aircrews are required to be proficient in more than 20 Surface Warfare (SUW) and Anti-Submarine Warfare (ASW) mission areas, opportunities to practice are infrequent.

When aircrews are not deployed, there are limited opportunities for "live" training - Live crew training is expensive and challenging to plan. Live training aids are difficult to obtain under other than prescribed, pre-designated conditions and locations specifically dedicated to live training. Live Navy helicopter aircrew training is conducted at one of two locations; the Atlantic Undersea Test and Evaluation Center (AUTEC), located in the vicinity of Andros Island, Bahamas and the Southern California Off-Shore Range (SCORE), near San Clemente Island, approximately 75 miles northwest of San Diego.

The Navy has recognized that training using high fidelity simulators is a suitable means for preparing helicopter aircrews for live training. Helicopter aircrew members do most of their initial and refresher crew training in fixed-facility high-fidelity simulators and as might be expected, high fidelity simulators are very expensive. Only a few exist and helicopter aircrews not located in the vicinity of the facilities must travel to the simulators. Inevitably, travel expenses impact unit budgets; not particularly popular during a time of fiscal constraints and increased interest in reducing costs.

US Navy Multi-Mission Helicopters

The US Navy employs a variety of surface ship and aviation platforms to hunt and attack adversary submarines as well as to detect, track and engage surface targets. Among these platforms are three variants of the Sikorsky Sea Hawk helicopter; the SH-60B, the SH-60F and the MH-60R (Sikorsky, 2010).

The SH-60B, also known as Light Airborne Multi-Purpose System (LAMPS III) can be armed with the MK-46 or MK-50 torpedo, the AGM-114 Hellfire missile and either a M-60D or GAU-16 machine gun. The LAMPS III platform is equipped with a sophisticated suite of sensors, including a towed Magnetic Anomaly Detector (MAD), air launched sonobuoys, the APS-124 search radar, ALQ-142 electronic support measures (ESM) system and an optional nose mounted forward looking infrared (FLIR) turret. Generally employed aboard smaller ships such as frigates, destroyers and cruisers, the SH-60B is crewed by a pilot, a co-pilot / Airborne Tactical Officer (ATO) and an enlisted aviation systems warfare operator, also known as the Sensor Operator (SENSO or SO).

The carrier-based SH-60F mission package and crew composition differs slightly from the SH-60B. The SH-60F uses the AQS-13F dipping sonar rather than a MAD device and carries fewer sonobuoys (12 instead of 25). Potential armament is essentially identical. Other differences include the addition of an enlisted tactical sensor operator.

The MH-60R is an upgraded version of the SH-60B and SH-60F. The Navy is in the midst of a seven year fielding effort, during which as MH-60R are delivered to a squadron, SH-60B and SH-60F aircraft are being retired. Featuring a glass cockpit (digital electronic displays, rather than mechanical gauges) the platform has an improved Airborne Low Frequency Sonar (ALFS), increased sonobuoy and acoustic signal processing, and a Multi-Mode Radar (MMR), including the Inverse Synthetic Aperture Radar (ISAR) that provides imaging and periscope detection modes. Other sensor improvements enable passive detection and targeting of radar sources and a Forward Looking Infrared (FLIR) sensor (Global Security, 2010).

Helicopter Flight Crew Training Lifecycle

The SH-60B helicopter is crewed by two officers and one enlisted person. The officers are rated aviators who are trained in the employment of acoustics sensors and weapons systems, while the enlisted sailor is a highly trained technician who operates the aircraft sensor systems used to detect and classify underwater acoustic signatures. The typical flight crew training lifecycle is shown in Figure 6.

Although the officers may come from a variety of commissioning sources, after commissioning, they attend primary flight training at either Naval Air Station (NAS) Whiting Field, near Milton, FL or NAS Corpus Christi, TX. Once they successfully complete primary flight training, they learn to fly helicopters at NAS Whiting Field.

After arriving at the Fleet Readiness Squadron (FRS) based at Naval Station, Mayport, FL, pilots receive flight instruction on the SH-60 as well as academics that include employment of acoustics sensors and weapons systems. Duration of SH-60B training is approximately six months, while MH-60R training lasts nine to twelve months.

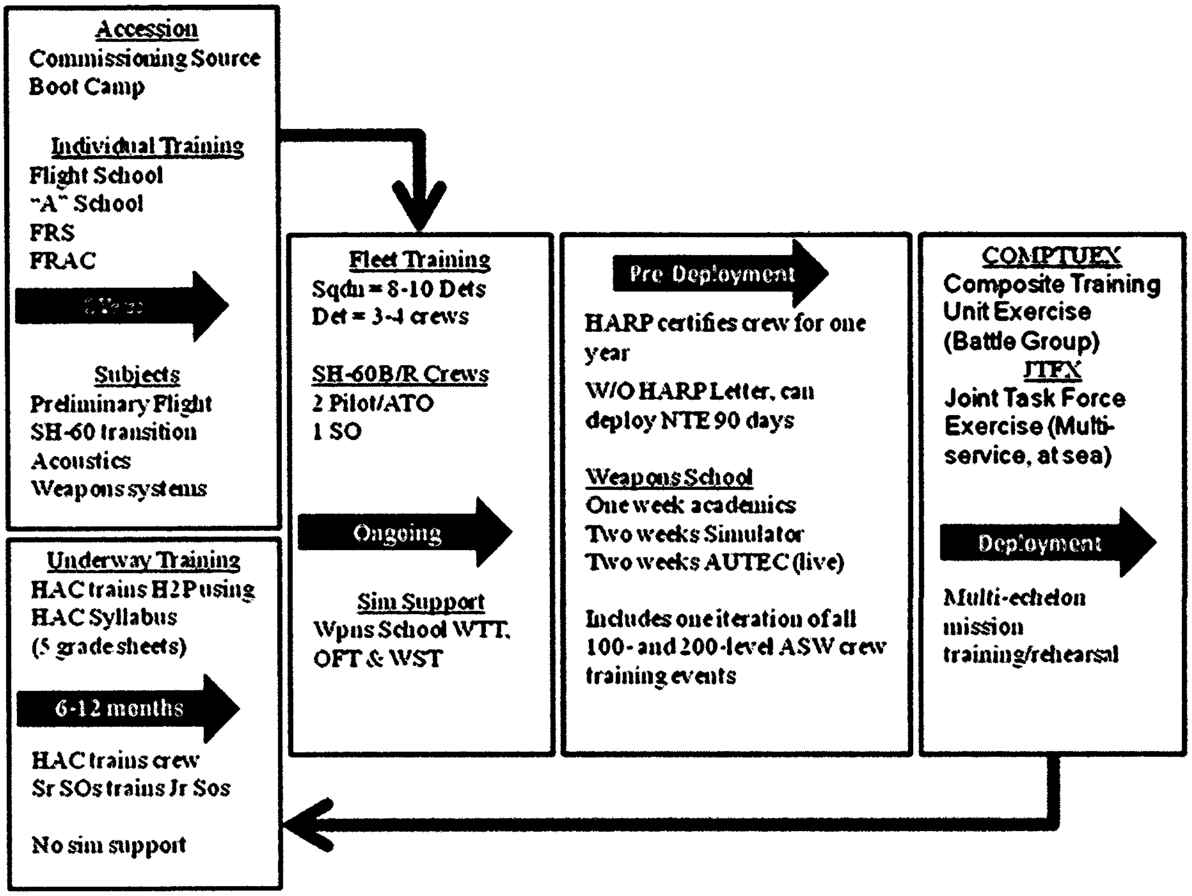

Figure 6: Helicopter Flight Crew Training Lifecycle

Enlisted sensor operator (Aviation Warfare Specialist) training commences once sailors have completed boot camp. Sailors attend "A" School at Pensacola, FL for about four months, where they learn basics of underwater acoustics and the use of sensor systems. After completing "A" School, sensor operators attend Fleet Replacement Aircrew (FRAC) training at the Helicopter Maritime Strike Weapons School, where the sailors receive advanced acoustics training and learn to operate mission-specific sensor equipment. Enlisted training on the SH-60B lasts approximately six months; MH-60R training duration is nine to twelve months. Both FRS and FRAC are administered by the Helicopter Maritime Strike (HSM) Weapons School at Naval Station, Mayport.

After officer and enlisted aircrew members have completed their initial training, they are assigned to squadrons. Currently, there are seven squadrons in the vicinity of Jacksonville(JAX) / Mayport; two MH-60R squadrons (HSL 70 and 74) two SH-60B squadrons in transition to MH-60R (HSL 42 and 46) the Fleet Readiness Squadron (HSL 40) and a Navy Reserve Squadron (HSL 60); each squadron consists of 12 crews; squadrons are made up of detachments that form a small unit for training and deployments; generally there are two to three crews per detachment; crewmembers are interchangeable within the detachment; usually detachments have one of two configurations, as depicted in Table 8.

Table 8: US Navy Helicopter Aircrew Detachments

Smaller Detachment	
Quantity	Position
1	Officer in Charge (OIC), also a Helicopter Aircraft Commander (HAC)
1	HAC
2-3	Helicopter second pilot (H2P)
2	Air Warfare Specialist (AW)
Larger Detachment	
Quantity	Position
1	OIC/HAC
2	HAC
3	H2P
3	AW

Once assigned to a detachment, air crew member training is somewhat decentralized. Primary interest is ensuring pilots maintain proficiency in their aeronautical skills, which is accomplished in "live" aircraft. Air crews periodically get opportunities to train in high fidelity simulators; usually one three to four-hour block every five to ten weeks.

In anticipation of an upcoming at-sea deployment, aircrew members participate in a five-week pre-deployment event, known as the Helicopter Aircrew Readiness Program (HARP), which is designed to ensure deploying helicopter aircrew members are at Readiness Level C-1. Operational readiness, according to Knapp (2001), is measured in four areas: personnel, equipment on hand, equipment serviceability and training. Each area is evaluated by a "C" or capability rating on a four level scale:

- C-1: unit has required resources and can perform its wartime mission
- C-2: unit can do most of its assigned wartime missions
- C-3: unit can do many, but not all assigned wartime missions
- C-4: unit is unable to perform its assigned mission due to lack of resources

Successful participation in the HARP certifies air-crew members for one year. Without a HARP certification letter, aircrew members are ineligible to deploy for more than a 90 day period.

The HARP consists of three phases: Phase 1 is a one-week academics refresher; Phase 2 consists of two weeks of mission area practice in high-fidelity simulators; and Phase 3 is a two-week long stint at the Atlantic Undersea Test and Evaluation Center (AUTEC), during which, air crews participate in "live" training in real-world helicopters. Upon completing the HARP, Detachments are attached to a deploying battle group and participate as part of the group in an ensuing Composite Training Unit Exercise (COMPTUEX).

Once underway, aircrew members' training generally consists of HAC Syllabus training (compare to Cognitive Walkthrough technique discussed in theoretical foundations sections). Unless they are called upon to participate in an at-sea multi-service Joint Task Force Exercise (JTFX), there usually are no further opportunities to practice either SUW or ASW crew level skills.

HARP Performance Measurement and Evaluation

During HARP Phases 2 and 3, instructors evaluate aircrew performance using a Master Flight Brief / Grade Sheet, a form that employs a combination of a Behaviorally Anchored Rating Scale (BARS)-variant and event-based approach to performance measurement. The form offers a six-point Likert-type scale (0-5, 2-10 or 3-15) enabling the instructor to quantify performance based on a set of guidelines provided in the Master Flight Brief / Grade Sheet Grading Guidance sheets.

Master Flight Brief / Grade Sheet Grading Guidance sheets provide specific guidance for 30 separate tasks / subtasks, including a textual description of behavior associated with each specific level of performance, e.g., to achieve a grade of 5, all performance standards are met; to achieve a grade of 3 the majority of performance standards are met, with only transitory errors that are corrected to an acceptable standard; and a grade of 0 is assessed when the evaluator needs to take positive control to avoid a safety of flight issue, the majority of performance standards are not met, and / or the aircrew does not comply with standard techniques / procedures.

In addition to assisting instructors to assess mission technical performance, the Master Flight Brief / Grade Sheet provides a means to assess Crew Resource

Management (CRM) performance in the areas of mission analysis / planning, decision making, communication and situational awareness. Descriptions of behaviors are consistent with theoretical constructs of communication content and protocol analysis, shared cognition, and human information processing.

Examples of Master Flight Brief / Grade Sheet and Master Flight Brief / Grade Sheet Guidance can be found in Appendix F and Appendix G, respectively.

Air Crew Competency Levels

Navy aircrew members are categorized in one of five general competency levels, based on their level of experience. Table 9 depicts the five experience levels categorizing Navy helicopter air crew members' competency.

Table 9: US Navy Helicopter Air Crew Competency Levels

Level	Description
Level I	Post Fleet Replacement Squadron (FRS)
Level II	Helicopter second pilot (H2P) prior to/during first cruise
Level III	Post cruise/Helicopter Aircraft Commander (HAC)
Level IV	Dept Head/OIC/FRS Tactics instructors
Level V	Seahawk Weapons and Tactics Instructor (SWTI)

Helicopter Flight Crew Training Systems

Instructors at the Weapons School conduct operational flight training for pilots and aircrew assigned to the U.S. Atlantic Fleet employing 12 live aircraft, three SH-60B Weapons Systems Trainers (WST) and 2 MH-60R WSTs. The Weapons School simulators are used by seven squadrons and are scheduled in four 4-hour blocks from 0600 to 2200 every day. Each high fidelity simulator is "available for training" 3,700 hours/year; 220 training days/year; 16 hours/day. Simulator training periods are scheduled in 2 to 4-hour blocks, each simulator can accommodate 3-4 crews/day; thus each crew can expect to train in a high fidelity simulator about once every 21-28 duty days, best case. Since the crews usually can practice only one or two mission areas during their simulator block, skills retention is a problem. Based on crew training readiness and currency requirements, there is currently a shortfall of high fidelity simulator throughput capacity of ~19,000 hours annually (Parsons, 2011). Routine priority of use is as follows (Davis, 2011):

- Naval Air Training and Operating Procedures Standardization (NATOPS) / Instrument training (requires use of High Fidelity Simulators)
- Tactical Evaluations (TAC EVAL) / BEAR TRAP
- Helicopter Aircrew Readiness Program (HARP)
- FRS / FRAC
- Fleet training
- Seahawk Weapons and Tactics Instructors (SWTI)
- Testing

System Maintenance takes precedence over everything; since an inoperable system is unusable for training. Currently, Fleet Synthetic Training (FST) events take precedence over all but maintenance. A FST usually requires two weeks for testing / integration and actual FST participation. During that timeframe, simulators are not available for training, exacerbating the throughput capacity problem.

High Fidelity Simulators

The Weapons School at NAS Mayport employs three weapons/tactics trainers (WTT) and two full-motion SH-60B operational flight trainers (OFT). An MH-60R OFT (Figure 7) is available at NAS Jacksonville, and another MH-60R OFT is currently under construction at NAS Mayport, with availability forecast for mid to late 2012.

Helicopter Operational Flight Trainers are six degrees of freedom (6 DOF) full motion simulators that feature exact replicas of aircraft cockpits, including pilot and co-pilot controls, instruments and systems monitoring panels as well as a high definition 3-D representation of the synthetic environment to provide visual references and cues. OFTs do not replicate the Air Warfare Specialist (AW) workstation, which is represented by a Weapons Tactics Trainer (WTT).

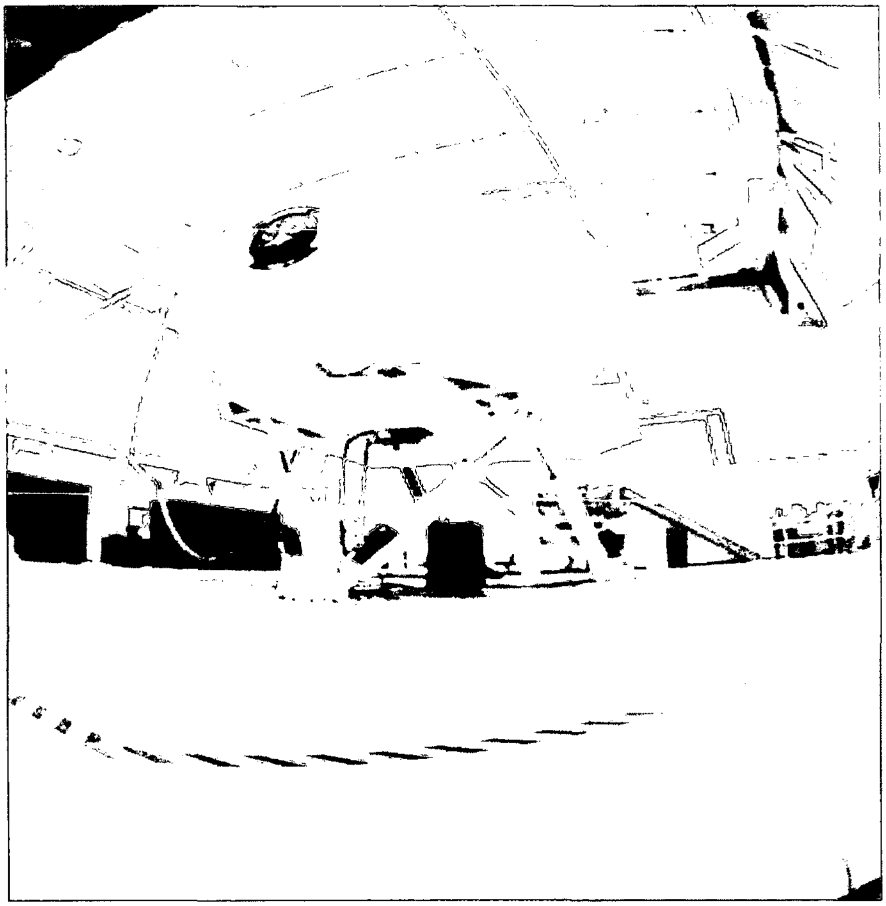

Figure 7: MH-60R Operational Flight Trainer (OFT)

Weapons Tactics Trainers replicate mission area system displays and high fidelity tactile control panels used by the ATO and AW to monitor visual and aural cues provided by simulated RADAR, SONAR, and other mission sensor systems.

An instructor suite is networked with OFT and WTTs, providing a simulated mission environment, including simulated sensors and synthetic natural environment that replicates various geographical areas of the world.

Connecting a WTT to an OFT creates a Weapons System Trainer (WST), as shown in Figure 8, that enables the crew to train as a team. The NAS Mayport site can accommodate two crews at a time. Currently at NAS Mayport, there are two instructor suites, one for each WST. Using pre-built databases, an instructor can choose from ten virtual world geographic locations in which to train. Specific examples include AUTEC range, the Mediterranean Sea, North Atlantic, Straits of Taiwan, and Straits of Hormuz. The Ocean Environment generator is a high fidelity system, but presents only one set of conditions per location. For instance, the AUTEC environment is hardcoded, based on readings taken on a single day in December. No other variations are available. The instructor can inject a variety of entities into the simulation environment to stimulate the simulated sensors and to present targets for the simulator weapons systems to engage. Based on crew proficiency, instructors can inject additional factors, such as background noise, to increase the intensity of the training and to present a greater challenge to the more proficient crews.

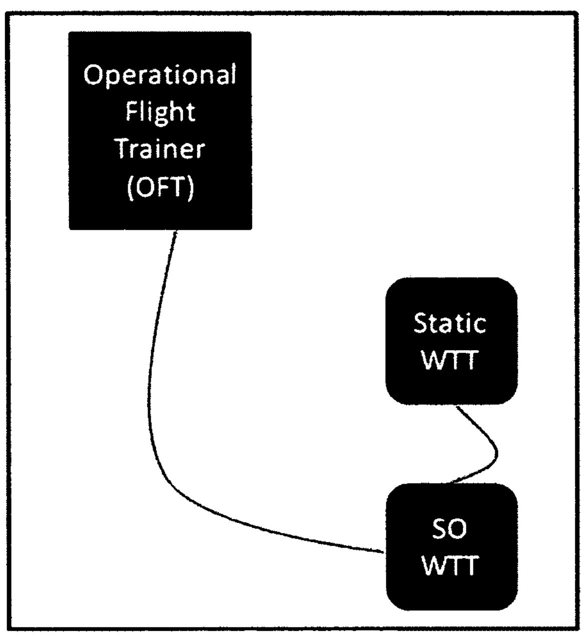

Figure 8: SH-60B Weapons Systems Trainer (WST) Configuration

Selective Fidelity Tactical Team Trainers

Prince and Jentsch (2001) postulate that low-fidelity PC-based training devices offer the potential for training CRM skills such as decision making, communication and workload management. The authors define low-cost, low-fidelity simulation along a continuum from no technology role-playing exercises in a classroom all the way to hardware / software solutions that enable crew members to perform many of the flight-related activities, they would perform either in a full fidelity simulator or in the real airplane. Citing the Brannick et al. (2000) study, they offer evidence of training transfer between a low-fidelity device and a high-fidelity simulator. They suggest some instances where low-fidelity devices can fit into existing CRM training, either to supplement, enhance or even replace some training elements. They do however; caveat their opinion by stating "there is no evidence to suggest that these systems can, or should, replace high-fidelity simulators in high-level training or evaluation" (p.156).

Mission Rehearsal Tactical Team Trainer (Gallo, et al, 2004)

Stick and Rudder Training for the Mind… that is how Gallo et al. (2004) described their experience with initial development and user testing of the ASW Virtual At Sea Training application, Mission Rehearsal Tactical Team Trainer (MRT3). The authors relate that making a unit's "at-sea" training days more effective was a driving factor in the system's concept and development. They note that shortfall in netted, integrated synthetic training ashore led to Center of Maritime Dominance issuing a Mission Needs Statement in 1997. Since existing high fidelity simulators were not designed to integrate or be interoperable with each other or other Navy simulations, Naval Aviation Warfare Command Training System Division with funding assistance from the Office of Naval Research, decided to leverage existing Virtual At Sea Training research and development efforts to create the MRT3. The initial Proof of Concept Demonstration conducted in October 2003, produced anecdotal evidence to the systems efficacy for a specific ASW mission scenario. A subsequent article, recounts the changes in system components and technological advances that led to the MRT3 becoming an integral part of the Fleet Forces Command – sponsored Fleet Synthetic Training events. Once again, anecdotal evidence indicates the training efficacy of MRT3, but falls well short of formal training effectiveness study.

The MRT3 is a hybrid system consisting of simulation and gaming technologies that provides networked crew-level training for crewmembers of aviation platforms engaged in Anti-Submarine Warfare (ASW). The hybrid MRT3 solution enables remote training, wherever and whenever ASW crew members want to use it. The MRT3 system does not provide the structure of a pure gaming system (i.e., mechanisms and instructional logic providing guidance and feedback tailored to situation and performance outcome and step-by-step progression through successive levels of difficulty). However, the system does provide repeated exposure to important cue patterns; facility to manipulate variables, view from different perspectives; enables observation of system over time; enables successive tasks that progress towards goals; and enhances motivation by providing a realistic sense of accomplishment, informative feedback, and the same sense of challenge associated with real-world task performance.

The MRT3 initiative began in 2003 as a part of the ONR-funded ASW Virtual At Sea Training (VAST) program. Designed as a fully integrated tactical team trainer, MRT3 allows the entire ASW Team to train together as they would while performing real-world missions. MRT3 is not a flight trainer; rather, the focus of this deployable system is on providing a capability to develop ASW tactical team expertise, specifically the cognitive aspects associated with performing a tactical mission.

There are two MRT3 system variants, the SH-60B and the SH-60F. As is the case with the real life referent platforms, the SH-60F variant has one more SENSO station than the SH-60B. Both variants can be networked together for coordinated tactical training, and have been certified to federate with NCTE to perform integrated training as a participant in Fleet Synthetic Training and Fleet Battle Experiment events.

The SH-60B MRT3 (depicted in Figure 9) is mounted on five laptop computers; an instructor / operator station (IOS), the Pilot workstation, the Airborne Tasking Officer (ATO) workstation, the Sensor Operator (SENSO) workstation and the acoustics generation modeling engine. The five computers are networked together, and the system includes transceiver headsets tied into a voice over internet protocol (VOIP) radio emulation that replicates intercom and external radio nets.

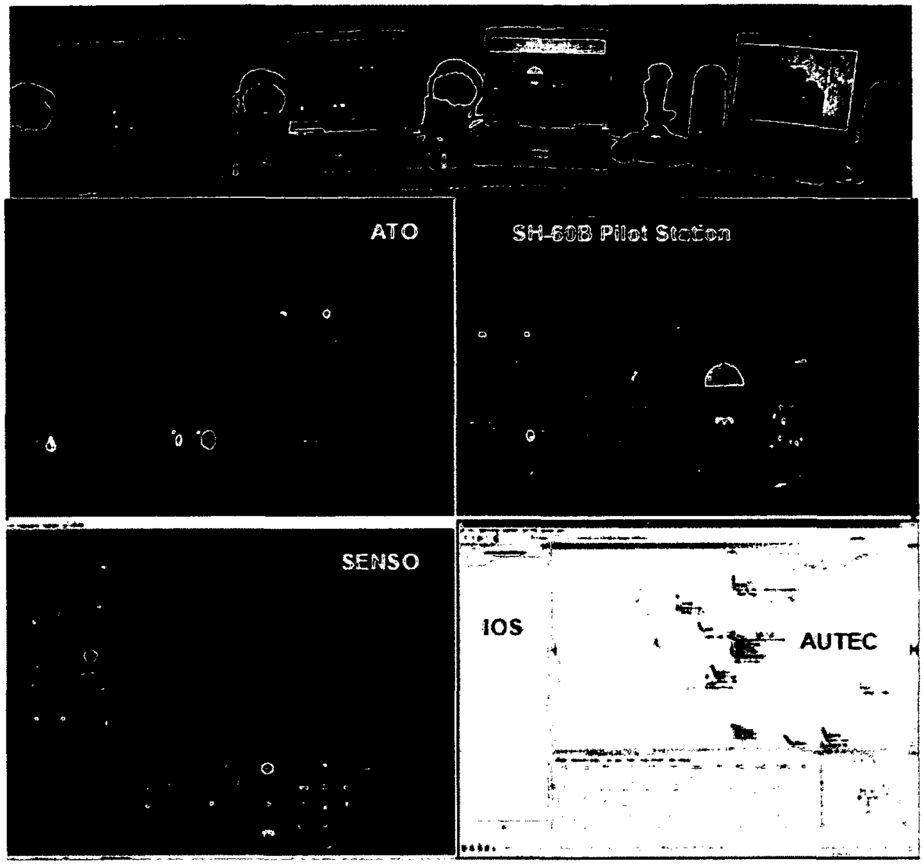

Figure 9: SH-60B Mission Rehearsal Tactical Team Trainer (MRT3) Suite and Station Displays

Pilot Station

Microsoft® Flight Simulator X provides basic aircraft functionality, enabling the pilot to exercise aircraft flight control while coordinating with the other crew members to accomplish their ASW mission. The pilot controls the virtual aircraft with a Commercial Off-The-Shelf (COTS) "joy stick" and the laptop keyboard. The pilot work station features an "outside the cockpit" visual display in addition to a functional instrument panel, providing necessary flight situational awareness (SA) for the pilot to navigate and maneuver the aircraft in response to crew mission interactions and the tactical situation.. The MRT3 development team modified Microsoft® Flight Simulator X to enable it to operate in a Distributed Interactive Simulation (DIS) environment. The MRT3 Pilot Station uses the Microsoft Windows® operating system.

Airborne Tactical Officer (ATO) Station

The ATO crew station includes a tactical display (including acoustic and non-acoustic sensor displays) along with a photo-realistic control panel featuring functional keys, switches and dials that are activated with a mouse click. The ATO crew station also offers a simulated Weapons Armament Panel that enables the ATO to engage targets within the virtual environment. The ATO workstation's photorealistic control panels and sensor display were developed using DiSTI GL Studio®.

Sensor Operator (SENSO) Station

As with the ATO workstation, the SENSO crew station features photorealistic sensor displays including radar, Magnetic Anomaly Detector (MAD), and acoustic (active and passive) sonobuoys, enabling the SENSO crew member to perform the tasks associated with localization, tracking and classification of ASW targets within the virtual environment.

Instructor Operator Station (IOS)

The Instructor Operator Station (IOS) provides the facility to configure, launch and control the MRT3 suite of computers, and consists of two major components; Joint Semi Automated Forces (JSAF) and the Common Distributed Mission Training Station (CDMTS). JSAF is a constructive simulation providing relevant synthetic entities representing friendly, neutral and opposing forces, as well as marine mammal models, in a common virtual environment, shared by the air crew member workstations and the acoustics generator simulation. Using JSAF models and environmental features, the Instructor / Operator can build realistic mission scenarios in virtually any area of the globe.

The other major component, CDMTS provides a common interface among the MRT3 suite of computers. Embedded in CDMTS is the Multi-host Automation Remote Command and Instrumentation (MARCI) application which provides the standard simulation management (SIMAN) functions; Launch, Start, Stop, Pause, Resume, along with the capability to load and manage sets of components (processes and applications) known collectively as a configuration. The application allows the user to create, edit, and store multiple, unique system configurations. Both JSAF and CDMTS run on the same laptop computer running the LINUX operating system.

Acoustics Generation Modeling Engine

To provide an accurate as possible representation of the acoustic environment, the SH-60B MRT3 uses an acoustics generation modeling engine leveraged from the Navy's Complete Acoustic Analysis Training System (CAATS). Since the high fidelity CAATS application provides ocean acoustics characteristics input for only passive sensors, the MRT3 development team modified the application to include accurate acoustics modeling for active sensors' input. The acoustics engine models various aspects of the maritime environment that impact sensor detection of underwater target. The acoustics model of water temperature, salinity and currents, when coupled with JSAF entity and environmental models and crewmember sensor displays, combine to provide a realistic and challenging tactical picture.

Preparing for the Empirical Study

As early as 1998, the Navy, recognizing the need to provide additional contextually appropriate opportunities to practice, commissioned the Networking for Naval Aviation Training study (Hougland et al., 1998). Championing adoption of the High Level Architecture standard to enable networked unit, Battle Group and Joint collective training, the study presented alternative strategies and an in-depth cost analysis for three networking options: 1) modify legacy high fidelity simulators, 2) develop selective fidelity tactical team trainers, and 3) employ a mixed legacy-Tactical Team Trainer (TTT) systems strategy. The study laid the groundwork for the development of the Navy Continuous Training Environment, a persistent networked training environment that integrates live, virtual and constructive simulations with live command, control and communications systems. The study also set the stage for research and development of three selective fidelity TTTs; two Mission Rehearsal Tactical Team Trainer (MRT3) variants of the SH-60 helicopter and the P3 Aircrew Tactical Team Trainer (PACT3). The P3 Orion is a multi-mission fixed wing aircraft; Anti-Submarine Warfare is one of its missions.

A conversation with the government's MRT3 and PACT3 Program Manager (Glass, 2010), revealed that beyond the 2003 MRT3 Proof of Concept Demonstration, no training effectiveness studies have been accomplished on any of the Tactical Team Trainers. Subsequent discussions with the MRT3 Program Manager as well as with Commander Naval Air Force U.S. Atlantic Fleet (O'Hanlon, 2010), revealed a need and desire to increase training effectiveness, especially for helicopter aircrew members. These conversations led to a series of discussions with the Helicopter Maritime Strike Weapons School Atlantic (HSMWSL) leadership during which several deficiencies were revealed.

Prior to deployment, a helicopter unit undergoes a thorough pre-deployment workup, including training in high fidelity simulators followed by an exercise on a live range. However, other than during these periods of increased focus on improving readiness, there is a significant shortfall in training opportunities due to limited number of high fidelity simulators and scheduling constraints. For example: 1) prior to pre-deployment workups, each aircrew gets only one three to four-hour block of simulator time every 8 -10 weeks, 2) aircrews usually can practice only one mission during their scheduled simulator block and 3) initial training and pre-deployment training take precedence over refresher training, thus further limiting simulator availability for any more than infrequent training. To further exacerbate the problem, even during pre-deployment workups, aircrews get only one three to four-hour block of simulator time for each mission area.

Instructor analysis of training performance revealed several performance deficiencies among aircrews during the simulator phase of pre-deployment training and also during training on the live range. Performance deficiencies include aircrew inability to accurately and consistently detect, acquire and track targets during the simulator phase and instructors' observations that approximately 30% of aircrews fail to place a weapon with an acceptable degree of accuracy during the live range phase. Instructors attribute these deficiencies to the limited number of contextually appropriate training repetitions available to the aircrews.

In an effort to provide additional opportunities to practice, Helicopter Weapons School Atlantic instituted a prototype Ground School, adding another week to the beginning of the Simulator Phase of the Helicopter Aircrew Readiness Program (HARP). The v1.0 ground school consists of one day of Anti-Submarine Warfare academics, one day of Surface Warfare academics and an increase in the number of HARP simulator sessions from one 3.0 hour simulation block per crew to three 2.0 hour simulation blocks per crew. For a HARP focused on training ten crews, the v1.0 ground school increases the total number of high fidelity simulator training events from 10 to 30, thus providing aircrews a three-fold increase in practice opportunities. Recognizing high fidelity simulator availability constraints, the v1.0 Ground School also features the use of low fidelity virtual environments (VE) to provide an additional 10 events for targeted practice in critical "Kill Chain" related skills. Practice focused on mission relevant Crew Resource Management (CRM) skills: specifically in the areas of aircrew member leadership/teamwork, communications and situational awareness (Danko, 2011) and (Danko, 2012).

Interested in validating the entire ground school concept, the Weapons School Commander saw value and was willing for the command to participate in an empirical study. To ensure the low physical fidelity simulation system was ready to provide value to the training program, instructors conducted an internal analysis of MRT3 capabilities to support training. Several shortfalls were noted and the Weapons School submitted a request for funding to upgrade the system to correct the shortfalls and match configurations currently fielded in the fleet. An additional motivation driving the analysis was the recognition that the SH-60B is being retired from the Navy's air fleet, to be replaced by the MH-60R by 2017. Within the last few years of an aircraft program's life-cycle, the only additional funding provided for modifications is for safety of flight issues; ergo no additional funding for training system upgrades. This fact focused the request not only on meeting current system requirements, but also to modularize the system so the system will be easily modified to match live aircraft system upgrades as they occur, thus maintaining concurrency with the fleet and avoiding negative transfer of training.

As discussed earlier in this chapter, since its introduction to the Fleet in 2006, MRT3 has primarily been used as a training aid to stimulate higher level Commanders and their staff as a surrogate for live aircraft in large-scale Fleet Synthetic Training events. Although exercise participants recognized the potential for MRT3 to provide mission-specific targeted training for aircrews, little progress had been made toward integrating the system into routine training. The Weapons School's desire to increase the number of contextually appropriate opportunities to practice provided the necessary impetus to investigate more thoroughly the MRT3 ability to support training as part of the expanded HARP curriculum.

CHAPTER 5

METHODS

This chapter details the experimental design, methodology and other information associated with this study, including statistical techniques.

As mentioned before, this study's thesis is that a framework for delivering contextually appropriate opportunities for warfighter practice can be applied to discover whether modeling, simulation and game-based virtual environments have the potential to improve individual performance for learners beyond the Novice Stage (e.g., Competent Stage) of skills acquisition, and that the contextually appropriate practice (CAP) framework can be used to assess the potential of low fidelity virtual environments to provide targeted practice and to improve individual performance, not only during training in high-fidelity virtual environments (near transfer) but also in the live environment (far transfer).

This study investigates the relationship of technology and learning science, and seeks to determine whether a CAP framework can be applied to answer the following research questions:

1. Given that delivering targeted practice via Virtual Environments (VE) contributes to improved individual performance at the Novice Stage of skills acquisition (Hussain et al., 2009), does VE-delivered targeted practice also contribute to improved individual performance at more advanced (e.g., Competent) stages of skills acquisition?
2. Does low-fidelity VE-delivered targeted practice contribute to improved individual performance during training in high fidelity simulators? During Live training?

To evaluate the thesis and answer the research questions, the study features an empirical evaluation of training effectiveness afforded by delivering additional training repetitions, using both low-fidelity virtual environment simulator systems and high-fidelity aircraft simulators.

This research was granted an Institutional Review Board (IRB) exemption. The approval letter is shown at Appendix A, and the approved Application for Exempt Research is shown at Appendix B. Data were derived from instructor reports on human subjects participating in regularly-scheduled aircrew member training as part of a unit preparing for deployment.

Research Design

Nonequivalent Control Group Design is a quasi-experimental design used where the researcher lacks full control over the scheduling of experimental stimuli (i.e., who is exposed and when). The design allows researchers to introduce something like experimental design into data collection scheduling within a natural setting, using naturally assembled collectives, such as classrooms (Campbell & Stanley, 1963). Because participants in this study constituted a "naturally assembled collective", the data collection effort was divided into treatment and control group using the Nonequivalent Control Group Design (Table 10).

Table 10 Research Design

Design 10, Nonequivalent Control Group Design - Campbell and Stanley, 1963				
Treatment:	N	O	X	O
Control:	N	O		O

Methodology

This effort includes a two-phased empirical study. Phase I examines helicopter aircrew members performance using both low-fidelity and high-fidelity modeling and simulation-based virtual environments. Phase II investigates helicopter aircrew members performance in the live environment. In both phases, aircrew members' communications, leadership/teamwork and situation awareness are assessed in the context of performing two anti-submarine warfare (ASW) mission-related tasks; Detection / Localization and Tracking. I selected the Detection / Localization and Tracking mission area tasks for three reasons: 1) performance is quantifiable, 2) the tasks provide opportunity for intra-aircrew member interaction and 3) there is a consensus among Weapons School Instructors and Commanding Officer that air-crew performance of these tasks is critical to mission accomplishment.

Participants

Seven aircrews from three squadron detachments were engaged in Helicopter Aircrew Readiness Program (HARP); a five-week long pre-deployment ramp-up exercise designed to ensure the aircrew members are proficient in mission-related tasks in two primary warfare areas: Surface Warfare and Anti-Submarine Warfare. A total of 20 aircrew members were scheduled, but only 19 participated in the study. Participants included six (6) enlisted aerial warfare (AW) specialists and thirteen (13) officers who are rated SH-60 helicopter aviators. Due to absence of data, one participant's results were excluded. Participant demographics are depicted in Table 11. Participant ages ranged from 24 to 39 years with the average age being 29.3 years. Time in Service averaged 8.1

years (minimum 2.2 years; maximum 20.0 years). Time in aircraft type ranged from 60 to 2300 hours with an average of 834 hours. Only three aircrew members had more than 2 months experience with their current crew. Average time since last simulation training event was five (5) months and average time since last live mission training event was over 17 months.

Table 11 Empirical Study Participants

	n	Avg age	Time in Svc (yrs)	Time in A/C type (hrs)	Time since last ASW Sim event (mos)	Time since last ASW Live event (mos)
Treatment						
Officers	4	29.8	8.9	798.5	3.3	4.3*
Enlisted	1	24.0	4.5	876.0	0.5	1.5
Total	5	28.6	8.1	806.8	3.9	2.9
Control						
Officers	9	29.6	8.7	570**	1.3	14.6*
Enlisted	5	29.4	7.0	1495.3**	15.9	37.25*
Total	14	29.5	8.1	847.6	6.5	22.2

* One participant had not participated in a live event

** No data provided for two participants

Aircrew member competency ranged from Level I to Level III in accordance with the Navy's five-level rating system depicted in Table 12. Seven (35%) aircrew members had previously been assessed at Level III, eight (40%) at Level II and the remaining five (25%) at Level I. Simultaneous with the HARP, participants at Level I were being evaluated to advance to Level II. Two of the enlisted AW specialists (both in the Control Group) had previously instructed at the Weapons School (Level IV), but had been assigned to other duties for at least the last two years, so were not current and had been downgraded to Level II. One was slated for separation from the Navy within a year due to lack of advancement to the next rank.

Table 12 US Navy Helicopter Air Crew Competency Levels

Level	Description	Approximate Equivalent: Dreyfus Skills Acquisition Stage (1981; 1986; 2004)	Approximate Equivalent: Steinaker & Bell Experiential Taxonomy (1979)
Level I	Post Fleet Replacement Squadron (FRS)	Advanced Beginner	Exposure / Participation
Level II	Prior to / during first cruise	Competent	Participation / Identification
Level III	Post cruise	Proficient	Identification
Level IV	Dept Head/OIC/FRS Tactics instructors	Proficient / Expert	Internalization / Dissemination
Level V	Seahawk Weapons and Tactics Instructor (SWTI)	Expert	Dissemination

Treatment Group relative experience

The relative experience of the five aircrew members comprising the treatment group was dispersed over a broad range. While both HACs had significant longevity and experience, the H2P / ATOs were both relatively junior. One of the HACs (prior enlisted with over 16 years' time in service) had 715 hours in the SH-60B helicopter and was assessed at experience level II. Her most recent ASW simulation event was 12 months earlier, and most recent ASW live event was six months earlier. The other HAC had 13 years in the military, 2200 hours in aircraft type, and was assessed at experience level III. His most recent ASW practice in a simulator was five months previous and it had been two months since his last ASW live event. Both H2P / ATOs had three years in the military and were assessed at experience level I. One had 167 hours in the SH-60B aircraft type, while the other had only 67 hours. Both had recent (within one month) ASW practice in a simulator, while only the member with more flight time had participated in the live event (2 months prior). The same enlisted AW specialist performed SO duties for both crews. Assessed at experience level II, he had four and a half years in service and 876 flight hours in the SH-60B aircraft type. His most recent ASW simulator practice was two weeks prior and his most recent ASW live event was a month and a half before.

Systems

The low fidelity virtual environment used in the study was the Navy's Mission Rehearsal Tactical Team Trainer (MRT3). The MRT3 suite consisted of three laptop computer-based crew member workstations, a common instructor operator station (IOS) from which the instructor controlled the simulation scenario within the synthetic environment and role played appropriate shipboard and / or airborne command and control elements. Each crew-member workstation and the IOS were equipped with headsets enabling simulated radio and intercom voice communications via voice over internet protocol (VOIP). The Pilot workstation was equipped with a joystick to enable the Helicopter Aircraft Commander (HAC) to control the virtual helicopter. Both the Aircraft Tactics Officer (ATO) and the Sensor Operator (SENSO) workstations were equipped with ergonomically-correct representations of their respective mission equipment control key sets. All three crew-member workstations were in approximately correct relative physical locations, encouraging the crew members to communicate over their radio/intercom as they would while performing their missions either in a high fidelity simulator or in "live" aircraft.

The high fidelity virtual environment used in the study was the SH-60B Weapons Systems Trainer (WST). There are two SH-60B WSTs at Naval Air Station (NAS) Mayport, each consisting of SH-60B Operational Flight Trainer, a six degrees-of-freedom (6 DOF) full motion simulator that features exact replicas of aircraft cockpit, including pilot and co-pilot controls, instruments and systems monitoring panels. The study employed both WSTs. A physically separated Weapons Tactics Trainer (WTT), networked with the OFT, replicated mission area system displays and high fidelity tactile control panels used by the AW to monitor visual and aural cues provided by simulated RADAR, SONAR, and other mission sensor systems. A networked instructor suite controlled the combined OFT, WTT and their shared simulated mission environment. Based on an assessment of the crews overall skills, the instructor had the ability to increase mission complexity and difficulty.

Live systems used were U.S. Navy SH-60B aircraft at the Atlantic Undersea Test and Evaluation Center (AUTEC).

Virtual Environment Training Scenarios

Similar training scenarios used for both low-fidelity and high-fidelity VEs were geographically located in a synthetic representation of a body of water in the vicinity of the Middle East. Practice scenarios increased in complexity from first to second practice session, including the addition of neutral ships, several adversary vessels (besides the targeted submarine), a US submarine, several US surface ships, and one or two aircraft

(P3 Orion fixed wing sub hunter and / or MH-60R, both of which have more sophisticated sensors than an SH-60B). Prior to the first practice session (in both low-fidelity and high-fidelity VE), instructors conducted a mission brief, outlining the tactical situation and providing current mission-related intelligence, such as suspected adversary activity and suspected last known location. Prior to the second practice session, aircrew members being assessed for advancement to the next competency level performed a formal mission briefing, outlining the results of their mission analysis and detailing their mission plan.

Tasks

Scheduled training tasks were first to Detect / Localize a potential enemy submarine, based on information gathered from mission briefing, assistance from friendly ships/aircraft (role played by instructors) and their own aircraft on-board sensors. Once the contact was detected and localized, the aircrew members were to use extant rules of engagement (ROE) and interpretation of results portrayed by passive (Doppler signature) and active (SONAR) sensors to track and identify the contact. The aircrew had to positively identify a contact as an enemy submarine (based on its unique set of acoustics signatures) prior to requesting and gaining permission to engage with weapons systems.

Each aircrew member has a unique role in performing ASW tasks. For instance, the Helicopter Aircraft Commander (HAC), who usually is the senior aircrew member and also the mission commander, is responsible for maneuvering the aircraft and supervising / coordinating aircrew activities. As the name implies the Helicopter Second Pilot (H2P) can take the controls and maneuver the aircraft from the left seat of the helicopter. Generally the H2P also performs the role of Aircrew Tactics Officer (ATO), communicating externally with the "Mother" ship, other ships and aircraft as well as communicating / coordinating internally with the HAC and the Sensor Operator (SO). The SO is also often referred to as the Anti-Submarine Warfare specialist or AW. The ATO maintains the "big picture" of the mission, while the SO focuses on sensor-provided cues and indications to detect, localize and track a potential target. Inherent in aircrew coordination for ASW missions, are crucial CRM skills: communications, situation awareness and leadership-teamwork. Effective communications leads to shared situation awareness, which along with leadership / teamwork in turn leads to better performance. Presumably, the net effect is a higher probability of mission success.

Procedure

The Helicopter Weapons School's lead instructor randomly selected four aircrews to participate in the study as the Treatment Group, which were to practice using the low-fidelity VE prior to their normally scheduled training (Table 13). Due to equipment problems, only two aircrews (five participants) were afforded the opportunity to practice using the low-fidelity VE. All participants (both Treatment and Control) had two ASW mission practices using the high-fidelity VE. During Phase II, all participants performed ASW mission for record in live helicopters at the Atlantic Undersea Test and Evaluation Center (AUTEC).

Table 13 Empirical Study Procedure

	Low Fidelity VE Practice	High Fidelity VE Practice 1	High Fidelity VE Practice 2	Live
Treatment	X	X	X	X
Control		X	X	X

Instruments

Data collection instruments used in this study are shown in Table 14. Individual performance in the areas of leadership/teamwork, communications and situation awareness was captured using observational data collection tables adapted from Anti-Air Teamwork Observation Measure (ATOM) procedure (Smith-Jentsch et al., 1998), as well as a self-report Mission Awareness Rating Scale (MARS) Questionnaire adapted from (Matthews & Beal, 2002). ATOM-derived data collection areas focused on Supporting Behavior, Leadership and Initiative, Information Exchange, and Communication Behavior. Instructors observed participants during training events and graded each aircrew member on a numerical Likert-type scale (0 – 5). The MARS-derived questionnaire provides insight into overall perceived situation awareness.

Participants were asked to assess their ability to detect/understand cues (in terms of identifying, understanding, predicting and strategy to achieve goals) as well as the difficulty they experienced in detecting/understanding cues (in terms of identifying, understanding, predicting and deciding how best to achieve mission goals).

Table 14 Data Collection Instruments Used in Empirical Study

Instrument	Description
Mission Awareness Rating Scale (MARS) Questionnaire (adapted from Matthews & Beal, 2002)	Subjective participant self-assessment of individual situation awareness
Modified ATOM data collection tables adapted from Anti-Air Teamwork Observation Measure (ATOM) procedure (Smith-Jentsch et al., 1998)	Instructor assessment of crew-level Information Exchange, Communication Behavior, Supporting Behavior and Leadership & Initiative
Performance Outcome data collection tables (White, 2012)	Uses Behaviorally Anchored Rating Scales (BARS) adapted from the Master Flight Brief / Grade Sheet and Master Flight Brief / Grade Sheet Guidance to assess performance outcomes on two ASW tasks – Detection / Localization and Tracking

To facilitate quantifiable analysis, the MARS questionnaire was modified assigning a Likert-type numerical value to the otherwise qualitative scale as demonstrated in the following example:

Please rate your ability to identify mission-critical cues in this mission.

1 = VERY EASY – able to identify all cues

2 = FAIRLY EASY – could identify most cues

3 = SOMEWHAT DIFFICULT – many cues hard to identify

4 = VERY DIFFICULT – had substantial problems identifying most cues

To assess targeted ASW mission-related task performance, Helicopter Weapons School instructors used Performance Outcome data collection tables. The tables were adapted from the Weapons School's Master Flight Brief / Grade Sheet and Master Flight Brief / Grade Sheet Guidance and employ Behaviorally Anchored Rating Scales (BARS). The forms use a 0 - 5 Likert-type scale to assess performance outcomes on two Anti-Submarine Warfare tasks – ASW Detection / Localization and ASW Tracking.

Types of Data

Weapons School instructors provided the following types of data:

- Instructor numerically scored assessments of individual and crew performance of mission-related tasks
- Numerically scored reports of CRM (decision-making, communications) behavior based on instructor observation, including:
    - Supporting Behavior
    - Leadership and Initiative
    - Information Exchange
    - Communication Behavior.
- Participants' self-assessed situation awareness scores, including:
    - Ability to Detect/Understand Cues
    - Difficulty to Detect/Understand Cues.

Statistical Analysis

Statistical hypotheses make assertions about one or more samples or populations and form the basis for conducting statistical tests, enabling scientists and engineers to draw conclusions about systems, processes or phenomena of interest. The null hypothesis (H_0) postulates that there is no difference between the two sample groups of data while the alternative hypothesis (H_a) represents the question/theory under consideration. Generally, investigators seek either to reject H_0, based on sufficient evidence in the data or they fail to reject the null hypothesis because of insufficient evidence in the data, thus lending credence to H_a (Walpole et al., 2007).

This study adopts the traditional hypothesis testing paradigm, which seeks to determine whether two population samples differ from one another significantly and usually take the following form:

$$H_0 : |\overline{X} - \mu| = 0$$
$$H_a : |\overline{X} - \mu| > 0 \qquad \text{(Eq. 1)}$$

where

h_0 is the null hypothesis, h_a is the alternate hypothesis, $\overline{X}$ is the sample mean, and μ is the population mean.

Depending on the situation, certain test statistics are better suited than others for determining whether or not data are sufficiently similar to facilitate comparison. For instance, the *t*-test is suitable for two small sets of data (n_1 and/or n_2<30). However, one should choose the type of test depending on the similarity of the standard deviations of the two sets. If the standard deviations are sufficiently similar they can be "pooled" and the Student *t*-test can be used. When the standard deviations are not sufficiently similar, an alternative procedure for the t-test is Cochran's *t*-test. The *F*-test (or *Fisher's test*) provides the criterion to determine if the variances of the data sets significantly differ. Therefore, for small data sets, the *F*-test should be used prior to the *t*-test. The following discussion on Student *t*-test, Cochran's *t*-test and *F*-test was derived from van Reeuwijk and Houba (1998).

Student t-test

Generally, the Student *t*-test may be applied to small data sets (n1, n2 < 30) where σ_1 and σ_2 are similar according to the results of an *F*-test. When comparing two sets of data, the Student *t*-test equation can be written as:

$$t = \frac{|\overline{X}1 - \overline{X}2|}{\sigma_p} \times \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \quad \text{(Eq. 2)}$$

where

$\overline{X}_1$ = mean of data set 1

$\overline{X}_2$ = mean of data set 2

σ_p = "pooled" standard deviation of the sets

n_1 = number of data in set 1

n_2 = number of data in set 2

The pooled standard deviation σ_p is calculated by:

$$\sigma_p = \sqrt{\frac{(n_1-1)\sigma_1^2+(n_2-1)\sigma_2^2}{n_1+n_2-2}} \quad \text{(Eq. 3)}$$

where

σ_1 = standard deviation of data set 1

σ_2 = standard deviation of data set 2

n_1 = number of data in set 1

n_2 = number of data in set 2.

To perform the t-test, the critical value of t *may* be found in the table (Appendix K); the applicable number of degrees of freedom *df* is here calculated by:

$$df = n_1 + n_2 - 2 \tag{Eq. 4}$$

Cochran's t-test

Where σ_1 and σ_2 are dissimilar according to the results of an F-test and size of data sets (n_1, n_2) < 30, Cochran's t-test may be a more appropriate statistical approach. When comparing two data sets with dissimilar standard deviations, calculate Cochran's t as follows:

$$t = \frac{|\overline{X}1 - \overline{X}2|}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \tag{Eq. 5}$$

Then determine an "alternative" critical t-value using the following calculation:

$$t_{tab} = \frac{t_1 \frac{\sigma_1^2}{n_1} + t_2 \frac{\sigma_2^2}{n_2}}{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \tag{Eq. 6}$$

where

t_{tab} = critical t from table (Appendix K)

t_1 = t_{tab} at n_1-1 degrees of freedom

t_2 = t_{tab} at n_2-1 degrees of freedom

Once the alternative critical t is determined, the t-test can be performed as usual: if $t < t_{tab}$ the null hypothesis (no significant difference in the means) is accepted.

***F*-test**

The *F*-test compares the variance of two data sets to determine if they belong to the same population. The test compares the ratio of the two variances as indicated in the following equation:

$$F = \frac{\sigma_1^2}{\sigma_2^2} \qquad \text{(Eq. 7)}$$

where the larger σ^2 is the numerator by convention. If the values differ no more than slightly, the estimates σ_1, and σ_2, are similar and their ratio (and that of their squares) should not deviate much from unity. Practitioners normally compare *F* with the applicable critical *F* value found in the critical *F*-value table (Appendix L). To read the table it is necessary to know the applicable number of degrees of freedom for σ_1, and σ_2 which are calculated by:

$$df_1 = n_1 - 1 \qquad \text{(Eq. 8)}$$

$$df_2 = n_2 - 1 \qquad \text{(Eq. 9)}$$

If $F \leq F_{tab}$ (critical value of *F* from the table) one can accept the null hypothesis ($\sigma_1 = \sigma$) at the 95% confidence level.

CHAPTER 6

RESULTS

Preliminary analysis of data associated with the Detect / Localize task (depicted in Table 15) appears to indicate an overall 0.25 average score (15%) improvement over reported historical averages during both Practice 1 and Practice 2 (Historic HARP simulator average scores = 3.75; historic HARP Live flight average scores = 4.0). Although Practice 1 for the Track task (depicted in Table 16) exceeds historical averages, Practice 2 shows a slight decrease when compared to historical averages. Performance for both tasks during live task performance shows an overall decrease in performance (2.75% and 14.5%, respectively) over historical averages. A more comprehensive review of both mission task performance areas follows, providing some insight into these apparent anomalies.

Table 15 Detect/Localize Task Performance Average Scores Summary

Detect/Localize Task Performance	Practice 1	Practice 2	Live
Treatment Group	4.50	4.00	2.80
Control Group	3.79	4.00	4.29
Overall	4.00	4.00	3.89

Table 16 Track Task Performance Average Scores Summary

Track Task Performance	Practice 1	Practice 2	Live
Treatment Group	4.50	3.20	2.80
Control Group	3.79	3.93	3.64
Overall	4.00	3.74	3.42

Preliminary analysis of the CRM areas of leadership/teamwork, communications and situation awareness reveals a similar pattern for Practice 1 but less coherent results achieved during Practice 2 and the Live Event. These areas too, will be addressed in greater depth in following sections.

Table 17 Leadership/Teamwork Average Scores Summary

Leadership/Teamwork	Practice 1	Practice 2	Live
Treatment Group	3.88	3.90	3.27
Control Group	3.35	3.56	3.79
Overall	3.56	3.65	3.59

Table 18 Communications Average Scores Summary

Communications	Practice 1	Practice 2	Live
Treatment Group	4.25	3.85	3.30
Control Group	4.03	3.49	3.77
Overall	4.09	3.59	3.61

Table 19 Situation Awareness Average Scores Summary

Situation Awareness	Practice 1	Practice 2	Live
Treatment Group	3.20	N/A	2.63
Control Group	2.98	N/A	3.39
Overall	3.04	N/A	3.08

Detailed Analysis – Practice 1

Overall, the data suggest a potential transient effect following a training intervention using the MRT3 low fidelity trainer for the Treatment Group in terms of initial performance during Practice 1. In every case, the Treatment Group's average Practice 1 scores were higher than the Control Group average Practice 1 scores for all five areas assessed. Table 19 provides basic statistical data on all five areas of interest for both Treatment and Control Groups during Practice 1. For the following discussions, n_1 = number of Treatment Group participants and n_2 = number of Control Group participants.

Table 20 Practice 1 Descriptive Statistics

	Treatment Group			
	Min	**Max**	**Mean**	**Std Dev**
Detect/Localize	4.00	5.00	4.50	0.55
Track	4.00	5.00	4.50	0.55
Leadership/Teamwork	1.50	5.00	3.76	0.77
Communications	2.00	5.00	4.04	0.84
Situation Awareness	2.00	5.00	3.20	0.21
	Control Group			
Detect/Localize	3.00	5.00	3.79	0.77
Track	3.00	5.00	3.79	0.71
Leadership/Teamwork	0.00	5.00	3.35	1.12
Communications	1.00	5.00	4.03	0.55
Situation Awareness	1.00	4.00	2.98	0.59

Examining results of the Detect / Localize task performance scores observed during Practice 1, a two-tailed *t*-test (n_1= 6, n_2 = 14, F = 0.411, t = 0.037) supports rejection of the null hypothesis at the 95% confidence level, suggesting a potential effect associated with low fidelity simulator practice. Figure 10, a scatter plot illustrating the relationship between Treatment and Control Group scores for the Detect / Localize task during Practice 1 also appears to corroborate the statistical evaluation of a significant difference in the means.

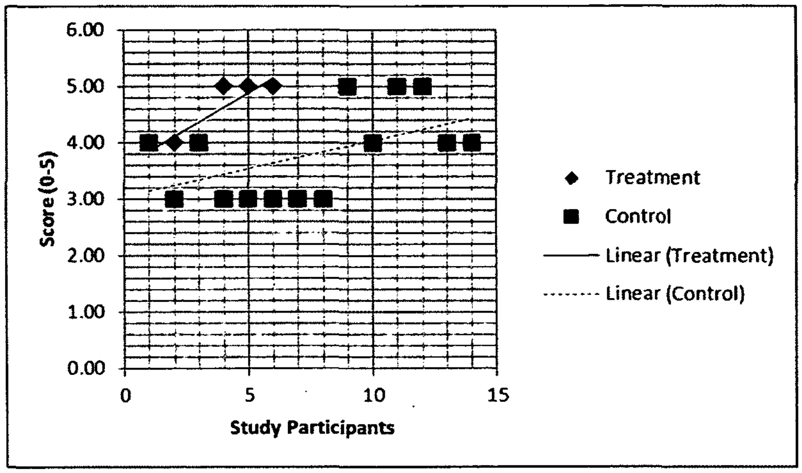

Figure 10: Scatter Plot of Practice 1 Data for Detect/Localize Task

Similarly, for the Track task performance scores observed during Practice 1, a two-tailed *t*-test ($n_1 = 6$, $n_2 = 14$, $F = 0.6164$, $t = 0.0303$) supports rejection of the null hypothesis at the 95% confidence level, suggesting a potential effect associated with low fidelity simulator practice. Figure 11, a scatter plot illustrating the relationship between Treatment and Control Group scores for the Track task during Practice 1 also appears to corroborate the statistical evaluation of a significant difference in the means.

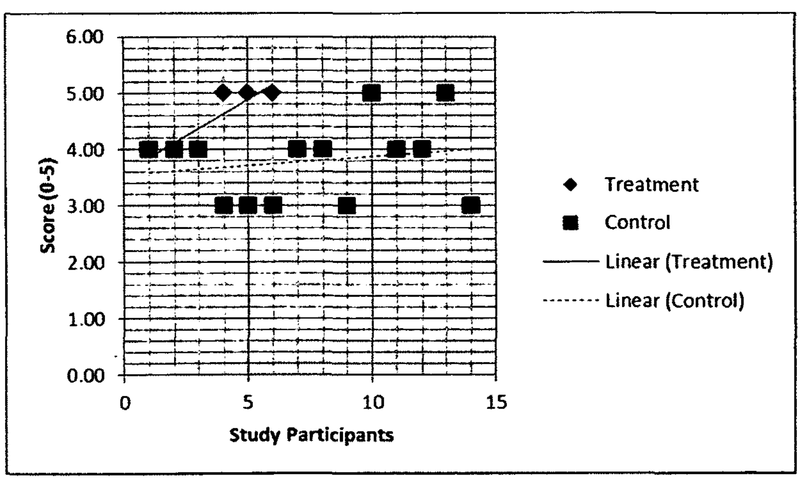

Figure 11: Scatter Plot of Practice 1 Data for Track Task

Assessment of the Leadership / Teamwork data captured during Practice 1, a two-tailed *t*-test ($n_1 = 6$, $n_2 = 14$, $F = 0.9274$, $t = 0.4033$) failed to reject the null hypothesis at the 95% confidence level. There were two extreme outliers in this category. In both cases, the Instructor noted deficient performance, especially in supporting behavior and leadership and initiative components. Running an excursion that excludes the most extreme outliers in each group results in ($n_1 = 5$, $n_2 = 13$, $F = 0.2233$, $t = 0.0230$), which would result in rejection of the null hypothesis. Figure 12 depicts the scatter plot with the outliers and Figure 13 illustrates the scatter plot of the data with the outliers removed.

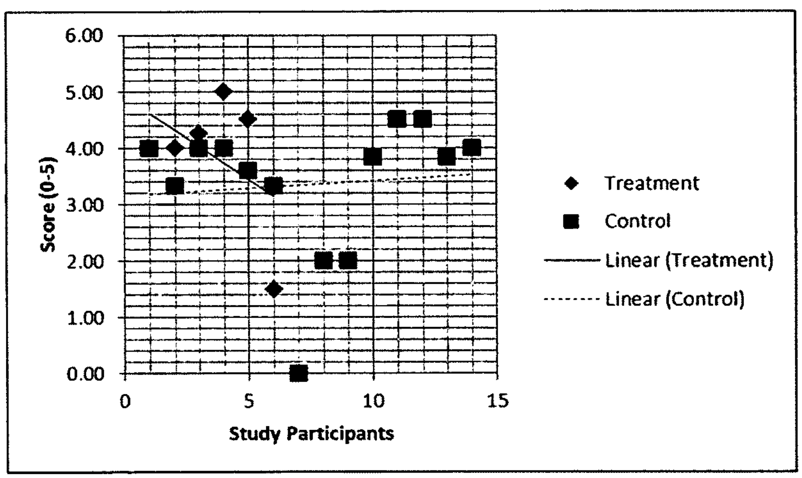

Figure 12: Scatter Plot of Practice 1 Leadership/Teamwork Data (with Outliers)

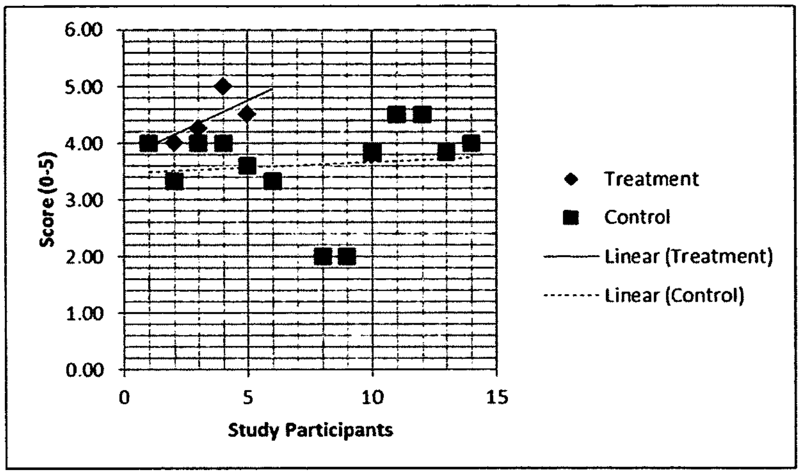

Figure 13: Scatter Plot of Practice 1 Leadership/Teamwork Data (Omitting Outliers)

Analysis of the Communication data from Practice 1 using a two-tailed *t*-test (n_1= 6, n_2 = 14, F = 0.5632, t = 0.4733) failed to reject the null hypothesis at the 95% confidence level. The components of Communications that appeared to be most troublesome to Control Group aircrew members were those associated with consulting all available sources (28.5% with score 2), passing information (21.4% with score 2), clarify/acknowledge receipt (14.3% had scores 2) and provide situation updates (14.3% scored <2). Only one aircrew member in the Treatment Group scored less than 3, and that was in the area of providing situation updates. Figure 14 illustrates the relationship between Treatment and Control Group data.

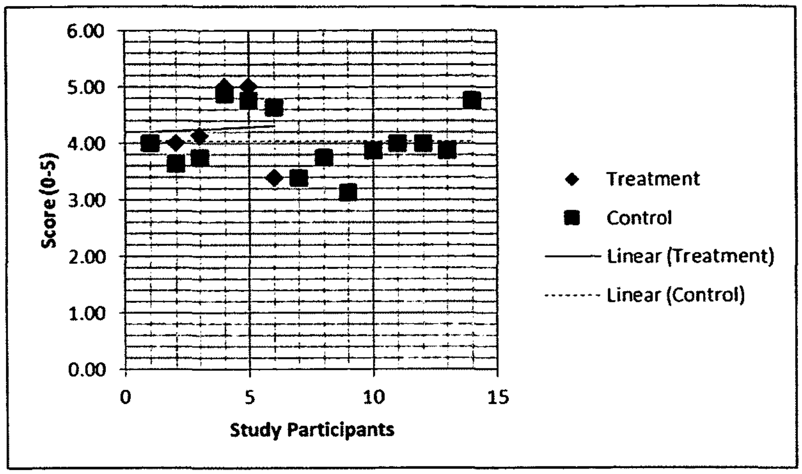

Figure 14: Scatter Plot of Practice 1 Communications Data

Situation Awareness is the only area in which scores are based on participants' self-assessment. Analysis of the Situation Awareness data from Practice 1 using a two-tailed *t*-test ($n_1 = 5$, $n_2 = 14$, $F = 0.0328$, $t = 0.3086$) failed to reject the null hypothesis at the 95% confidence level. In fact, if the two most extreme outliers were omitted, the *t*-test outcome approaches unity, i.e., $t = 0.9565$. Figure 15 visualizes the relationship between Situation Awareness scores based on Treatment Group and Control Group self-assessment.

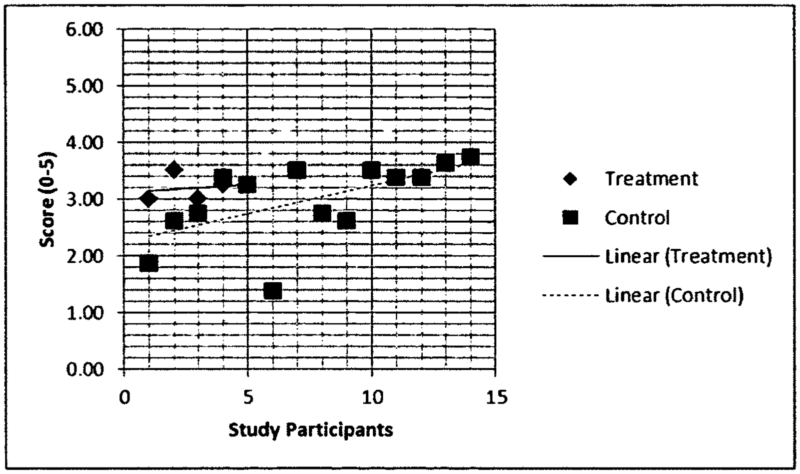

Figure 15: Scatter Plot of Practice 1 Situation Awareness Data

Detailed Analysis – Practice 2

Practice 2 results appear to be somewhat ambiguous compared to those obtained during Practice 1 (Table 21). Control Group Practice 2 performance showed overall improvement in every area over Practice 1. The following section investigates in greater depth.

Table 21: Practice 2 Descriptive Statistics

	Treatment Group			
	Min	**Max**	**Mean**	**Std Dev**
Detect/Localize	4.00	4.00	4.00	0.00
Track	3.00	4.00	3.20	0.45
Leadership/Teamwork	2.00	5.00	3.90	0.62
Communications	3.00	4.00	3.85	0.50
Situation Awareness	N/A	N/A	N/A	N/A
	Control Group			
Detect/Localize	3.00	5.00	4.00	0.53
Track	3.00	5.00	3.93	0.55
Leadership/Teamwork	2.00	5.00	3.56	0.62
Communications	2.00	5.00	3.49	0.65
Situation Awareness	N/A	N/A	N/A	N/A

Both groups' mean score for Detect / Localize task were similar. Analysis of the Practice 2 data using a two-tailed *t*-test (n_1= 5, n_2 = 14, *F* = n/a, *t* = 1) failed to reject the null hypothesis at the 95% confidence level. The two-tailed t-test reveals no significant difference between the groups' outcomes, despite a 0.53 difference in standard deviations of the samples. Figure 16 depicts a visual representation of the relationship of the scores.

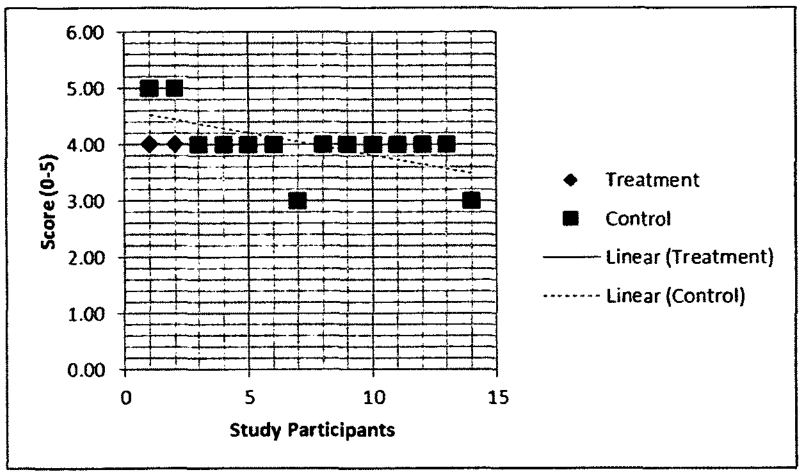

Figure 16: Scatter Plot of Practice 2 Detect/Localize Data

Analysis of the Practice 2 Track task data using a two-tailed *t*-test (n_1= 5, n_2 = 14, *F* = 0.9966, *t* = 0.0165) supports rejection of the null hypothesis, but in the opposite direction of that expected. As shown in Figure 17, the Control Group Scores were consistently better than those of the Treatment Group during Practice 2. A search of the basic data appears to rule out the possibility of instructor grading bias, since the same instructors evaluated crew member performance on the Detect / Localize task. A possible explanation may be related to intra-crew leadership / teamwork and communications, which is addressed in the Practice 2 Leadership / Teamwork and Communications discussions.

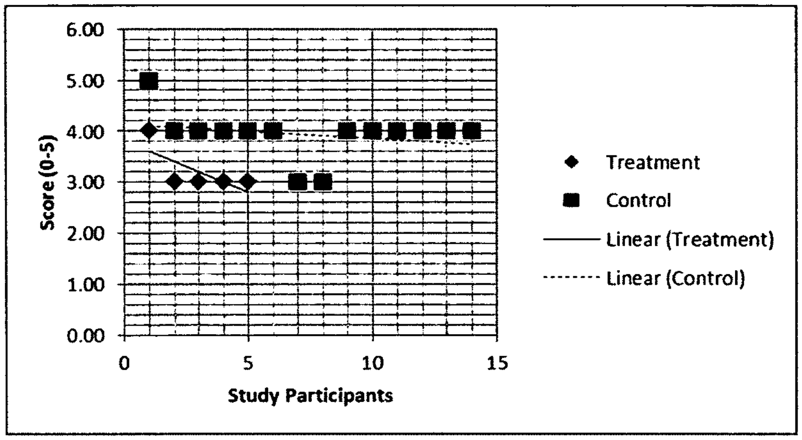

Figure 17: Scatter Plot of Practice 2 Track Task Data

Analysis of the Leadership / Teamwork data from Practice 2 using a two-tailed *t*-test ($n_1 = 5$, $n_2 = 14$, $F = 0.9448$, $t = 0.3287$) failed to reject the null hypothesis at the 95% confidence level (see Figure 18). Although the data from both groups exhibit similar variability and standard deviations, several instructor observations as they relate to the Track task bear further investigation. During Practice 2, Treatment Group aircrew members' performance were inconsistent with those observed during Practice 1, especially in the areas of monitoring for and communicating errors, requesting / offering backup, leadership guidance and statement of clear and appropriate priorities. As observed in the Practice 1 analysis, if outliers were omitted, the difference for both groups is significant. The seemingly lackadaisical approach to leadership / teamwork, combined with a similar drop-off in the area of intra-crew communications may offer some explanation for the decrease in Track task performance experienced by the

Treatment Group from Practice 1 to Practice 2, but offers little explanation for the apparent increase in Control Group Track task performance over the same interval.

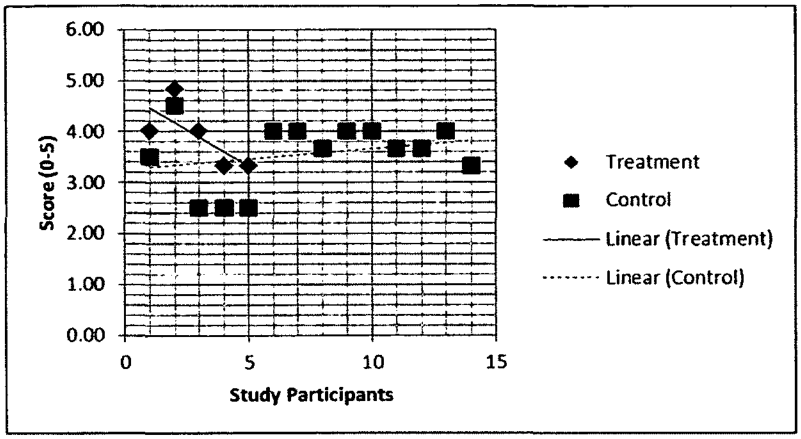

Figure 18: Scatter Plot of Practice 2 Leadership / Teamwork Data

Use of a two-tailed *t*-test ($n_1 = 5$, $n_2 = 14$, $F = 0.5161$, $t = 0.2450$) failed to reject the null hypothesis at the 95% confidence level for Practice 2 Communications (see Figure 19 for illustration). Both groups experienced a decrease in scores compared to Practice 1. Treatment Group average Communications score decreased 9.4% while the Control Group had a 13.4% decrease. The decrease in intra-crew communications scores may have a direct bearing on the relative decrease in Leadership / Teamwork scores but offers no explanation for the lack of decreased Control Group Track task performance between Practice 1 and Practice 2.

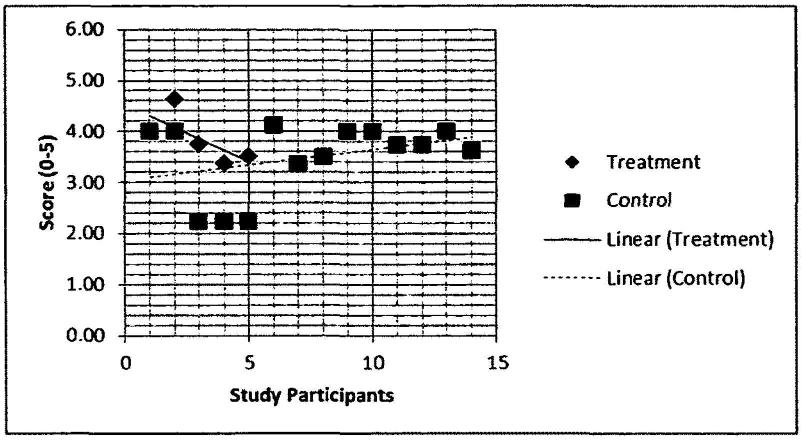

Figure 19: Scatter Plot of Practice 2 Communications Data

Data were not collected on Situation Awareness during Practice 2.

Detailed Analysis – Live Event

Treatment Group scores in the Live Event trended downward from those reported for Practice 1 and Practice 2, while Control Group showed overall improvement compared to Practice 2 results (Table 22).

Table 22: Live Event Descriptive Statistics

	Treatment Group			
	Min	Max	Mean	Std Dev
Detect/Localize	2.00	4.00	2.80	1.10
Track	2.00	4.00	2.80	1.10
Leadership/Teamwork	1.00	5.00	3.27	0.69
Communications	1.00	4.00	3.30	0.75
Situation Awareness	1.00	4.00	2.63	0.56
	Control Group			
Detect/Localize	3.00	5.00	4.29	0.80
Track	3.00	5.00	3.64	0.61
Leadership/Teamwork	3.00	5.00	3.70	0.66
Communications	2.00	5.00	3.71	0.79
Situation Awareness	3.00	4.00	3.36	0.29

Analysis of the Live Event Detect / Localize task data using a two-tailed *t*-test ($n_1 = 5$, $n_2 = 14$, $F = 0.3936$, $t = 0.0343$) supports rejection of the null hypothesis, but in the opposite direction of that expected. Figure 20 illustrates the relationship of Treatment Group versus Control Group scores.

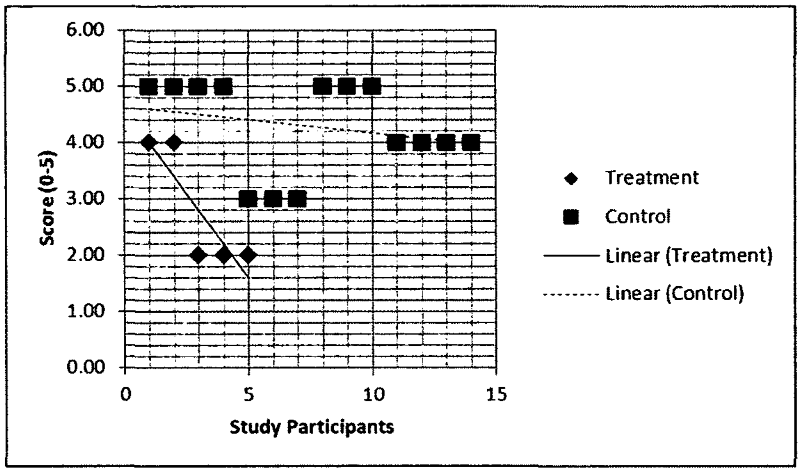

Figure 20: Scatter Plot of Live Event Detect/Localize Task Data

Track task scores during the Live Event were lower on average than Detect / Localize task scores. A two-tailed *t*-test ($n_1 = 5$, $n_2 = 14$, $F = 0.1184$, $t = 0.1650$) failed to reject the null hypothesis at the 95% confidence level for Live Event Track task. Figure 21 depicts relationship of Treatment and Control Group scores.

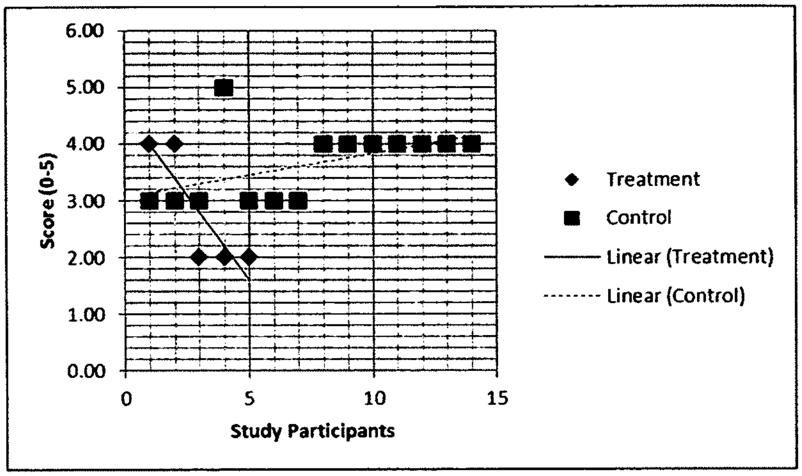

Figure 21: Scatter Plot of Live Event Track Task Data

Assessment of Live Event Leadership / Teamwork using a two-tailed *t*-test ($n_1 = 5$, $n_2 = 14$, $F = 0.8481$, $t = 0.2648$) failed to reject the null hypothesis at the 95% confidence level. Figure 22 represents the relationship between Treatment and Control Group data.

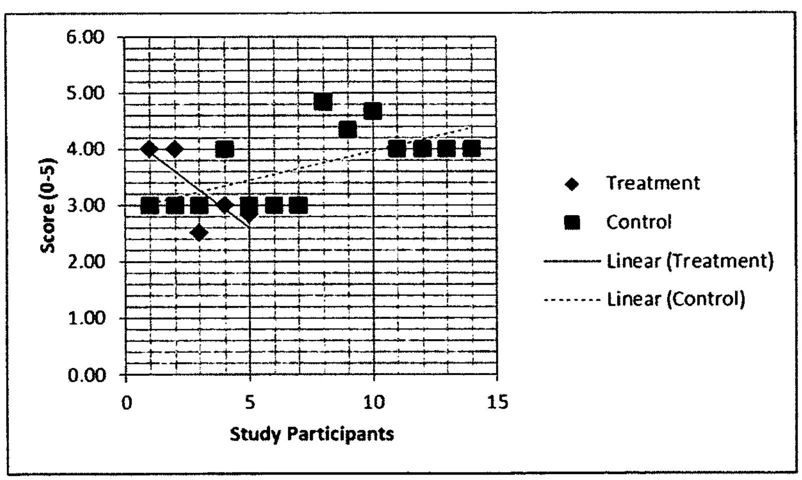

Figure 22: Scatter Plot of Live Event Leadership / Teamwork Data

Analysis of Communications scores captured during the Live Event using a two-tailed *t*-test ($n_1 = 5$, $n_2 = 14$, $F = 0.9503$, $t = 0.3342$) failed to reject the null hypothesis at the 95% confidence level. Figure 23 depicts the relationship between Treatment and Control Group data.

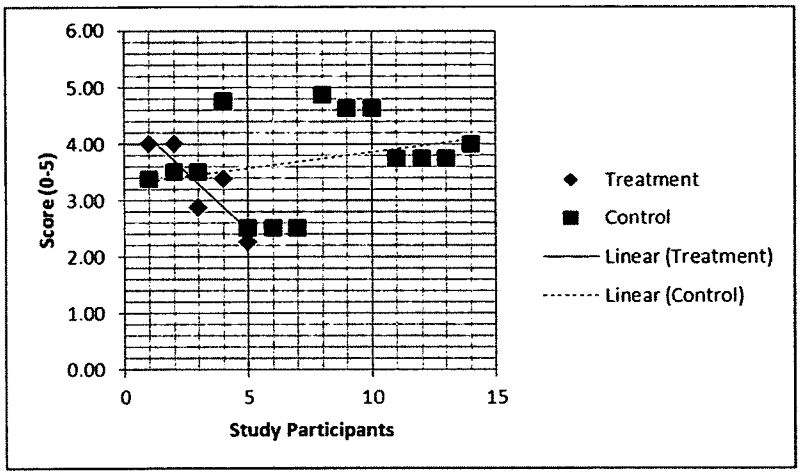

Figure 23: Scatter Plot of Live Event Communications Data

Self-reported Situation Awareness captured during the Live Event, assessed using a two-tailed *t*-test ($n_1 = 5$, $n_2 = 8$, $F = 0.1646$, $t = 0.0388$) supports rejection of the null hypothesis, but in the opposite direction of that expected. Figure 24 shows the relationship between Treatment and Control Group data.

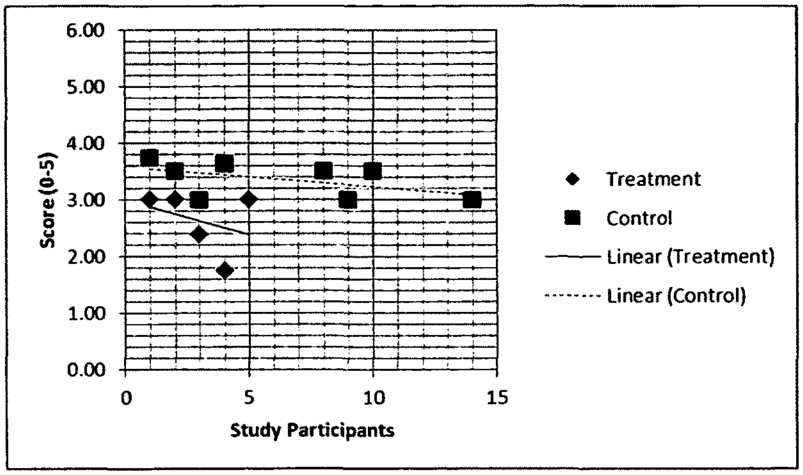

Figure 24: Scatter Plot of Live Event Situation Awareness Data

Treatment Group Trends

Although Treatment Group Crew 1 outperformed Treatment Group Crew 2 during Practice 1, the performance of the enlisted AW specialist was significantly degraded as part of the second crew. Supporting Behavior, Leadership and Initiative, Information Exchange, and Communication Behavior all exhibited reduced performance.

During Practice 2, the enlisted AW participated only with Treatment Group Crew 2. Both crews performed similarly in the mission tasks, but Crew 1 out-performed Crew 2 during the second practice by a significant margin in the areas of Leadership / Teamwork and Communications (18% and 19% respectively), while Situation Awareness was not assessed during Practice 2.

The difference between the two Treatment Group crews was even more marked during the live event. Mission task scores for Crew 1 were double Crew 2 mission task scores. Similarly, Crew 1 scores exceed Crew 2 scores by 68% in Communications, by 70% in Leadership / Teamwork, and by 80% in Situation Awareness.

These trends imply some sort of dysfunction amongst Crew 2 members. Unfortunately the data collection instruments used in this study capture only the impact of individual attitudes and dysfunctional behavior on training performance. The instruments are not designed to interpret results.

Applying the CAP Frameworks to Competent Stage Learners

As was shown in the previous discussion on the Recruit Training Command Novice Stage training, this investigation asserts that CAP-D and CAP-A frameworks are also applicable to Competent Stage training. Figure 25 applies the CAP-D framework to providing targeted practice to Navy helicopter crews who are at the Competent Stage of skills acquisition. As discussed in Chapter 4, the Weapons School did not modify the training content, context, targeted skills or skills acquisition stage being addressed. However, the Weapons School Commander modified the HARP training schedule to provide additional opportunities to practice and agreed to the addition of the low-fidelity MRT3 to support an additional training iteration for randomly selected participants.

Applying the CAP-D framework to this study, we see that the Content consisted of procedural knowledge. Similar to the RTC effort, this study focused on procedural knowledge since Conceptual and Declarative Knowledge had been addressed during previous training.

The Delivery Mechanism was the Mission Rehearsal Tactical Team Trainer (MRT3), a low physical fidelity (e.g., minimal pilot / co-pilot controls, no motion), high mission fidelity (e.g., sensor displays, accurate synthetic environment - acoustics generation, geographic locale/features) personal computer-based virtual environment.

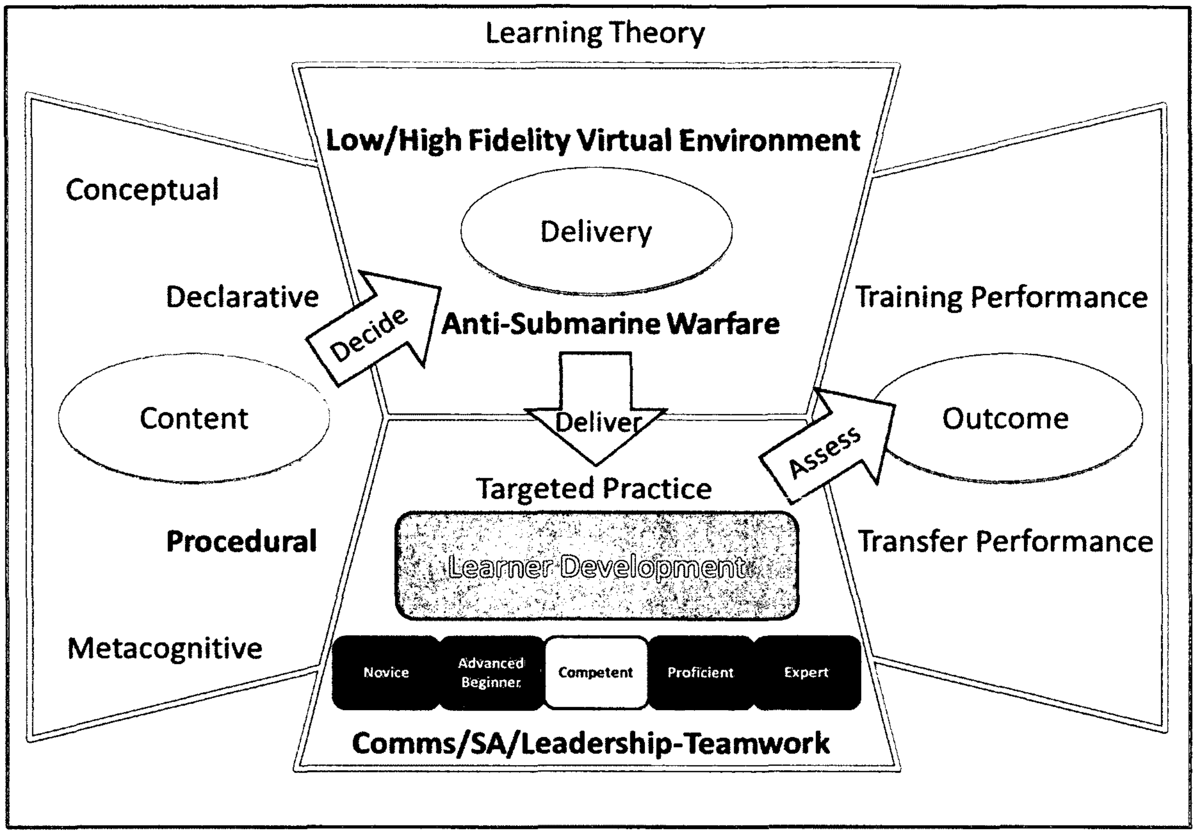

Figure 25: CAP-D Framework Applied to Learners at the Competent Stage of Skills Acquisition

The Context was an Anti-Submarine Warfare scenario, a critical but under-practiced mission and Targeted Practice included a coordinated search to detect and localize a potential enemy submarine; and once detected and localized, to maintain contact, to positively identify by acoustic signatures and to engage, when authorized by the appropriate authority.

Besides the mission tasks, Targeted Skills were Leadership / Teamwork, Communications and Situation Awareness. Both mission tasks require accurate, timely information sharing. Leadership is not just a function of rank / seniority; each aircrew member is expected to take the lead as appropriate, when their expertise and situation awareness make him / her the best qualified to do so. All crewmembers need to understand situation, so sharing information is paramount.

For this study, the Learner Development Stage was on-average, Competent, although, as pointed out in Chapter 4, the Treatment Group had only one member at Experience Level III, while the others were either at Level II or were tested and achieved Level II during the HARP that hosted this study.

The Outcome was Training and Transfer performance based on instructor assessments using Performance Outcome Data Collection Form (Appendix H) for the two mission tasks and the Modified ATOM Data Collection Forms (Appendix I) to capture Leadership/Teamwork and Communications Behavior. Additionally, Outcome was based on participant self-reported situation awareness based on the adapted Mission Awareness Rating Scale (MARS) Questionnaire (Appendix J).

As shown in Figure 26, this study applied the CAP-A framework to assess the suitability of low-fidelity VE to deliver contextually appropriate practice for learners at the Competent Stage of skills acquisition. In contrast with the RTC studies, assessment of system performance yielded inconclusive results. Although the disparity between the studies could point to a potential weakness in the framework, it is more likely that several other factors (some internal and some external) may have influenced the results.

During Practice 1 in high-fidelity VEs, Treatment Group mission task performance scores exceeded both Control Group and historical averages, suggesting a transient "warm up" effect following a training intervention using the low fidelity trainer. However, the CAP-A framework exposed a gap during Practice 2 and Live Event, in that the data do not provide sufficient evidence of treatment effect persistence for the Anti-

Submarine Warfare mission tasks. Indeed, both Treatment and Control Group Track mission task performance declined between Practice 2 and the Live Event, indicating a potential delta in environmental cues in both low-fidelity and high-fidelity simulators. Additionally, performance in the CRM areas of Leadership / Teamwork and Communications may not be directly attributable to the delivery mechanism, since both low-fidelity and high fidelity systems offered a similar level of opportunity to practice the skills. Although there was no statistically significant difference between Treatment and Control Group Situation Awareness scores during Practice 1 in the high-fidelity VE, Control Group performance trended better in the Live Event. Situation Awareness data were not collected during Practice 2, thus reducing the amount of comparable data and reinforcing the system assessment grade of Inconclusive.

Contextually Appropriate Practice - Assessment (CAP-A) Framework					
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Mission Task Performance	Detect/Localize potential target	Behaviorally Anchored Rating Scales (BARS) adapted from the Weapons School's Master Flight Brief/Grade Sheet	Instructor assessed numerical score (Q-O)	Performance Outcome Data Collection Form	I
	Track potential target	Behaviorally Anchored Rating Scales (BARS) adapted from the Weapons School's Master Flight Brief/Grade Sheet	Instructor assessed numerical score (Q-O)	Performance Outcome Data Collection Form	I
Leadership/ Teamwork	Leadership and Initiative	Offers guidance	Instructor assessed numerical score (Q-O)	Modified ATOM Data Collection Form	I
		Shared Leadership			I
		States Priorities			I
	Supporting Behavior	Monitors for errors	Instructor assessed numerical score (Q-O)	Modified ATOM Data Collection Form	I
		Corrects Errors			I
		Requests/offers backup			I
Communications	Information Exchange	Uses all available sources	Instructor assessed numerical score (Q-O)	Modified ATOM Data Collection Form	I
		Passes information			I
		Provides situation updates			I
		Acknowledges or asks for clarification			I
	Communications Behavior	Use proper phraseology	Instructor assessed numerical score (Q-O)	Modified ATOM Data Collection Form	I
		Communications are audible			I
		Does not engage in excess chatter			I
		Provides complete reports			I
Situation Awareness	Ability to detect and understand important cues	Ability to identify mission-critical cues	Self-assessed numerical score (q-s)	Modified MARS Questionnaire	I
		Ability to understand what was going on			I
		Ability to predict what was about to occur next			I
		Ability to develop strategy (how best to achieve goals)			I
	Difficulty to detect and understand important cues	Mental effort to identify cues	Self-assessed numerical score (q-s)	Modified MARS Questionnaire	I
		Mental effort to understand what was going on			I
		Mental effort to predict was about to occur next			I
		Mental effort to decide how best to achieve goals			I

Q-O: Quantitative data collected objectively
q-s: Qualitative data collected subjectively

Figure 26: CAP-A Framework Applied to Learners at the Competent Stage of Skills Acquisition

Discussion

By way of review, this study's thesis is that a framework for delivering contextually appropriate opportunities for warfighter practice can be applied to discover whether modeling, simulation and game-based virtual environments have the potential to improve individual performance for learners beyond the Novice Stage (e.g., Competent Stage) of skills acquisition, and that the conceptually appropriate practice (CAP) framework can be used to assess the potential of low fidelity virtual environments to provide targeted practice and to improve individual performance, not only during training in high-fidelity virtual environments (near transfer) but also in the live environment (far transfer).

Since the study investigates the relationship of technology and learning science, it seeks to determine whether a CAP framework can be applied to answer the following research questions:

1. Given that delivering targeted practice via Virtual Environments (VE) contributes to improved individual performance at the Novice Stage of skills acquisition (Hussain et al., 2009), does VE-delivered targeted practice also contribute to improved individual performance at more advanced (e.g., Competent) stages of skills acquisition?
2. Does low-fidelity VE-delivered targeted practice contribute to improved individual performance during training in high fidelity simulators? During Live training?

Evaluating the Thesis Statement and Research Questions

This dissertation takes a multi-discipline approach, combining modeling and simulation and educational psychology aspects. Although I would have preferred better statistical results from the empirical study, the focus of my thesis is on the contextually appropriate practice (CAP) framework. The mixed results from my empirical study serve to illustrate how the CAP frameworks can work under less than ideal circumstances.

In essence the contextually appropriate practice-delivery (CAP-D) framework is descriptive, visualizing the various elements selected to accomplish a training objective. The contextually appropriate practice-assessment (CAP-A) framework is also descriptive, capturing the diagnostic tools used to determine, for instance, whether a low-fidelity simulation provides appropriate opportunities to practice within a particular context or far a particular stage of skills acquisition. If, as was the case in my study, results are negative or inconclusive, a practitioner can cycle back through the CAP-D framework to investigate whether the content, delivery mechanism, targeted skills or learner development stages were matched up correctly.

If the results are statistically significant, as was the case with the Recruit Training Command studies; Great! Everything matched up appropriately. However, with negative or inconclusive data (which is very likely in many real-world situations), the CAP-D and CAP-A can provide an opportunity to uncover gaps, adapt to the situation and make appropriate changes.

The CAP framework was applied, first during a pilot study using the results from the Recruit Training Command studies (Bowers et al., 2009; Hussain et al., 2009; Hussain et al., 2011) to establish a baseline and subsequently during the empirical study associated with this effort to test my thesis statement. The RTC studies suggest that VE-delivered targeted practice contributes to improved individual performance at the Novice Stage of skills acquisition. Unlike the RTC studies, results from applying the CAP framework during my empirical study were inconclusive. A brief comparison of the two studies may shed some light on the differences.

Similar to the RTC studies, procedural knowledge was the primary content of interest during this study. However, while the RTC studies focused on game-based delivery, this investigation employed a low-fidelity / high-fidelity virtual environment as the delivery mechanism.

Another contrast between the studies; like the RTC research, the skills in this study included communications and situation awareness, but this study focused on leadership and teamwork vice decision-making as critical Crew Resource Management (CRM) factors.

While the context of the Novice Stage training was shipboard damage control procedures, this study addressed individual aircrew member performance in the context of conducting helicopter Anti-Submarine Warfare team training. Where the Novice Stage study used a game-based VE to deliver opportunities to practice individual knowledge and skills the learners had not previously applied, this study used both low-fidelity and high fidelity VEs to deliver additional opportunities to refresh skills previously obtained, but which either had not been performed recently or had not been performed heretofore at the Competent Stage of skills acquisition.

In addition to the differences in content, context and delivery methods, assessment efforts employed different data collection instruments. This study employed a Behaviorally Anchored Rating Scales (BARS) data collection instrument adapted from the Master Flight Brief / Grade Sheet and Master Flight Brief / Grade Sheet Guidance to assess performance outcomes on two ASW tasks – Detection / Localization and Tracking. Instructors in this study used modified ATOM data collection tables adapted from Anti-Air Teamwork Observation Measure (ATOM) procedure (Smith-Jentsch et al., 1998) to observe and record results of six behaviors and attitudes associated with leadership / teamwork, communications and situation awareness.

Finally, this study administered the Mission Awareness Rating Scale (MARS) Questionnaire (adapted from Matthews & Beal, 2002) to elicit subjective participant self-assessment of their individual situation awareness. The RTC research also employed several instructor-graded forms that took a BARS approach to assessing Communications and Situation Awareness in addition to a couple that captured Decision-Making performance. No participant self-assessment forms were used during the RTC investigation.

In answer to Research Question #1, based on results of my empirical study, the CAP framework revealed insufficient evidence to support the assertion that VE-delivered targeted practice contributes to improved individual performance at the Competent Stage of skills acquisition. Two key differences between the studies were the difference in relative experience and the difference in rank. The military experience level and rank of the RTC study participants was essentially the same across the board; they were all at the Novice Stage and all were recruits. While experience levels amongst the RTC participants were relatively uniform, the experience differential of helicopter aircrew members involved in this study was broad, ranging from less than three years' time in service for junior officers and enlisted personnel considered to be Advanced Beginners all the way to 20 years' time in service for the HAC considered to be level III (Proficient Stage). Additionally for the aircrew members, time in aircraft type spanned from a low of 60 hours to a high of 2300. Another potential key difference might be attributed to individual motivational factors presented in Hersey and Blanchard's (1977) Situational Leadership® model. Where Novice attributes are described as low confidence-high commitment (enthusiastic learner), Advanced Beginners and Competent Stage learners (developmental levels D2 and D3 in the Situational Leadership® model) exhibit some competence – low commitment (disillusioned learner) and high competence-variable commitment, respectively. Finally, there was a difference in the sample size and group distribution between the two studies. While my study consisted of 19 respondents, there were 31 RTC study participants. More noteworthy is the split between treatment group participants and control group participants. Where my study had a 5 / 14 treatment – control split, the RTC study was more evenly divided (15 treatment and 16 control participants).

In answer to Research Question #2, the CAP framework revealed that results from Practice 1 suggest a transient "warm up" effect following a training intervention using the low fidelity trainer. As Alvarez et al. (2004) observe, the first outcome from training is the learner's reaction. Responses from the Treatment Group after their first high-fidelity simulator event indicated that they thought the low-fidelity VE provided a good "warm up" prior to training in the high-fidelity simulator. Treatment Group members shared their opinion that the practice in the low fidelity simulator has utility, relevance, and value. In addition, the self-reported diminished sense of situation awareness during the live event could be an indication of the need for improving simulator cues and for additional practice especially in the skills associated with the Track task, in accordance with Bowers et al.'s (2009) assertion that to elicit targeted behaviors, the practice environment must provide appropriate cues and adequate fidelity.

Based on overall results from the empirical study, the CAP framework revealed indications of a gap in training performance during Practice 2 and a gap in transfer performance during the Live Event, thus the validity of the assertion that low fidelity virtual environments have the potential to support improved performance during training in the high-fidelity VE (near transfer) or in the live environment (far transfer) is not supported. Although the Treatment Group's scores were decidedly higher during Practice 1, the Control Group displayed steady improvement throughout the study, leading to an inconclusive evaluation of the low-fidelity simulator.

The decline in Treatment Group performance over the three training events can be partially explained by tracing the scores of one particular crew, Treatment Group Crew 2. During Practice 1, only one crewmember scored low in the areas of Leadership / Team-work and Communications. During Practice 2, the entire crew began to display lower scores across all performance areas and during the Live event, that particular crew consistently displayed the lowest scores of all study participants. One can only conjecture dysfunctional behavior among the crew members since none of the instruments used in the study captured specific affective interactions between aircrew members. All we have to go on is the results obtained through instructor assessments using the data collection instruments and self-assessed situation awareness. Furthermore, examination of instructor comments revealed no additional insight.

In light of the results of the data gathered during this investigation, it might be worthwhile to re-visit one or more of the study elements. For instance, modifying the experimental design by adding a low-fidelity training intervention prior to Practice 2 and the Live Events respectively might lead to increased training and transfer performance. The additional opportunities to practice would be consistent with Wetzel, Konoske and Montague's (1983) findings on the relationship between performance degradation and time elapsed since previous training. Taking this approach might also lead to increased skills internalization and automaticity, as described in the works of Kraiger et al. (1993), Rogers et al. (1997), and Moors & De Houwer (2006).

Other Factors Impacting the Study

Learner performance may have been influenced by individual responses to events beyond the control of the participants, their command and the empirical study. Undoubtedly the most emotionally significant event during the study was the suicide of a senior enlisted member of one of the participating squadrons. Although the deceased was not involved in the training, the affected unit cancelled all training and all flights for the following day. Based on leader comments, the event added even more stress to a unit that was already severely understaffed and overtaxed. When the aircrew members from that unit returned to HARP training, the strain was evident in their nearly complete inability to concentrate and perform complex coordinated technical tasks.

Another round of data collection might have revealed a different pattern, but due to military scheduling, this was not possible. It took nearly a year to coordinate schedules to conduct the empirical study. The sponsoring Commodore and the Helicopter Weapons School Commander both changed command and neither incoming commander was familiar with the effort. Moreover, the HARP content is changing as a result of the Fleet transition from the SH-60B helicopter (for which we currently have an existing low-fidelity tactical team trainer) to the MH-60R (for which no low-fidelity trainer yet exists).

CHAPTER 7

CONCLUSIONS AND RECOMMENDATIONS

The motivation for this research emerged from an abiding interest in the use of modeling, simulation and serious games to enhance human training, education, learning and development, especially in the military domain. Although research abounds in the use of simulation and games for education and training, little research has specifically focused on transforming conceptual, declarative, procedural and metacognitive knowledge into cognitive and physical skills through simulation-supported contextually appropriate targeted practice. Moreover, the literature revealed no existing correlated framework for delivering and assessing contextually appropriate opportunities for warfighter practice.

The primary value of this investigation is development of the CAP-D and CAP-A frameworks for delivering and assessing contextually appropriate opportunities for targeted warfighter practice. This study demonstrated the CAP-D and CAP-A frameworks' potential utility; first by showing in a pilot study how they apply to the use of a game-based delivery system within the context of U.S. Navy recruit (Novice Stage) training. Subsequently, in an empirical study involving Competent Stage learners, the frameworks were applied to discover gaps in training pertaining to the use of low-fidelity VE training systems prior to conducting training in high-fidelity simulators and live helicopters on an ocean range within the context of U.S. Navy helicopter Anti-Submarine Warfare mission task performance.

This study contributes to filling a gap in the training research and development body of knowledge concerning the use of modeling, simulation and gaming technologies to deliver contextually appropriate opportunities to practice. Toward that end, this effort developed a multi-discipline Contextually Appropriate Practice – Delivery (CAP-D) framework, which leverages learning theory, principles of instructional design and instructional system design, along with concepts of training and transfer effectiveness. The framework simultaneously captures the key elements of learner development from a theoretical standpoint while offering the flexibility to modify elements of the model to accommodate changes in modeling and simulation content delivery mechanisms, context and learner skills acquisition stage.

To augment the CAP-D model, a complementary Contextually Appropriate Practice – Assessment (CAP-A) framework provides a description of the means to evaluate a system's suitability to deliver contextually appropriate targeted practice. Within the CAP-A construct, practitioners can use data collection instruments to capture quantitative and/or qualitative measures that describe performance in selected mission task areas along with other targeted learning objectives of interest to determine if the combination of content, delivery mechanism, context and targeted practice achieve the desired learning outcomes. More detailed evaluation of assessment scores pointing to strengths, weaknesses and potential areas for remediation require the use of additional tools outside of the CAP-A framework.

Recommendations for Future Work

This investigation is foundational, providing a springboard for future research. The CAP framework is by no means complete. At present, the CAP-D framework is merely descriptive, i.e., it describes knowledge content type, delivery mechanism, targeted skills, and learner skill level the practitioner chose to use. In its current state, the CAP-D framework cannot confirm which delivery mechanism works best for a given type of training content under particular contexts. Similarly, the CAP-A framework describes assessed areas, tasks, measures, data types, and collection instruments; but does not diagnose performance of or prescribe corrections to the CAP-D framework elements. The CAP-A does not predict future performance, identify strengths, weaknesses or potential areas for remediation. Finally, neither the CAP-D nor the CAP-A framework prescribes corrective measures or suggests alternative approaches; they only provide the framework to describe those activities.

Additional research is needed to make the frameworks prescriptive. A helpful first step would be to conduct studies identifying potential weaknesses or gaps to assist in making the frameworks more robust and useful. Applying a rigorous engineering approach, the CAP framework could be automated, with executable modules associated with each of the CAP-D and CAP-A components. The baseline CAP-D framework could be used as a "front-end" or graphical user interface (GUI) linking it to a decision-tool "back-end" (probably employing a system dynamics modeling approach, due to the model's complexity). Using the GUI, a practitioner could select desired knowledge type (content) and learner skills acquisition level and based on results of the research described in the following paragraphs, the system could diagnose problems, explore possibilities, identify constraints, and provide recommendations for modifying components such as knowledge content, delivery mechanism, learner skills acquisition stage or assessment elements. Another potential refinement would be to enable the automated CAP-D to ingest and process results gleaned from assessments using the CAP-A framework, and thereby produce recommendations for the next cycle.

More research is needed to discover and map where modeling, simulation and serious games best fit in the learning continuum. Additional research is needed to develop means, processes, tools and tests to identify optimal areas for delivering targeted practice using modeling, simulation and game-based learning delivery systems. We need to discover if certain delivery mechanisms are better suited for particular knowledge types (content), stages in learner development, learning scenarios or tasks. Further investigation is needed to evaluate what skills are best trained at each acquisition stage.

If we understand where modeling, simulation and serious games best fit in the learning continuum; what skills are best trained at each acquisition stage; and whether certain delivery mechanisms are better suited for particular content and / or skills, the CAP-D and CAP-A frameworks can provide a consistent approach to support future research in other training domains and contexts.

A suggested roadmap for future investigation is provided at Figure 27. Several candidates for future research include using the CAP-D and CAP-A frameworks to investigate other knowledge content, delivery methods, context and sequence and at different stages of skills acquisition.

For instance, the proposed automated frameworks could accommodate team-based training for Advanced Beginners to corroborate Hussain et al.'s (2011) assertion that team performance improves after practicing as individuals. An automated CAP-D framework would certainly support this type of investigation and CAP-A could provide the framework for evaluating the results.

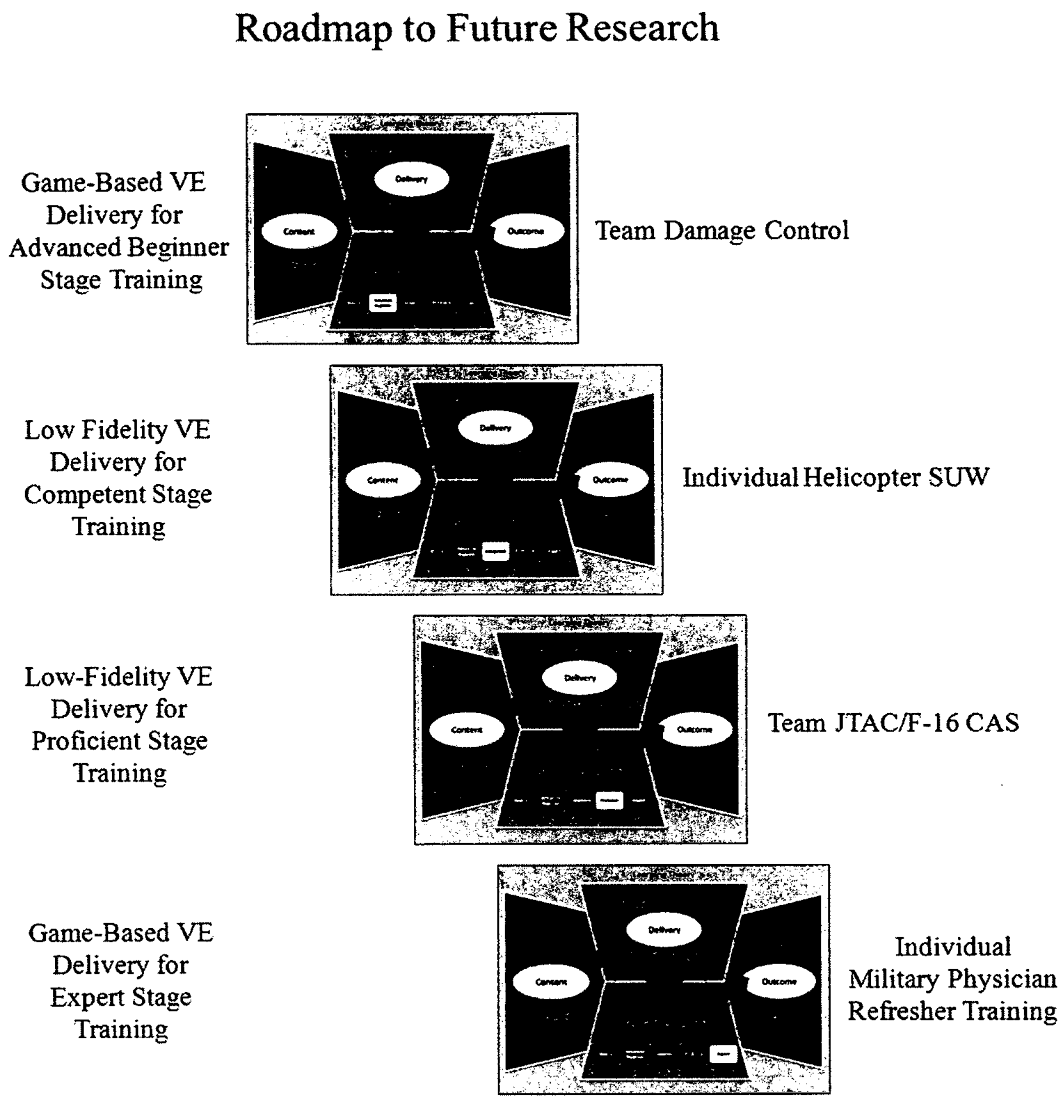

Figure 27: Suggested Roadmap for Future Research

Another candidate for future research could extend this work to investigating individual helicopter aircrew members' performance in context of helicopter Surface Warfare (SUW) training. As with this study, focus could be on the use of low-fidelity VE-delivered contextually appropriate practice for learners at the Competent Stage of skills acquisition.

Since both the RTC and helicopter Anti-Submarine Warfare use cases were Navy-centric, consider conducting a team study involving US Air Force F-16 pilots and Joint Tactical Air Controllers at the Proficient Stage of skills acquisition practicing Close Air Support (CAS) skills via low-fidelity VE.

Delivering contextually appropriate practice to military health care providers is another extremely important warfighter concern that until recently has been somewhat overlooked. Although the use of life-like mannequins has become more commonplace, studies aimed at investigating the potential of game-based VE's to provide refresher training to military personnel at the Expert Stage of skills acquisition are still needed.

BIBLIOGRAPHY

Alexander, A., Brunye, T., Sidman, J., & Weil, S.A. (2005). *From Gaming to Training: A Review of Studies on Fidelity, Immersion, Presence, and Buy-in and Their Effects on Transfer in PC-Based Simulations and Games*. Woburn, MA: DARWARS Training Impact Group.

Alvarez, K., Salas, E. and Garofano, C. (2004). An Integrated Model of Training Evaluation and Effectiveness, *Human Resource Development Review*, 3 (4), December, 2004, 385-416.

Anderson, L. W. and David R. Krathwohl, D. R., et al (Eds.) (2001). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. Allyn & Bacon. Boston, MA (Pearson Education Group); cited in Wilson, L. O., Beyond Bloom - A new Version of the Cognitive Taxonomy, retrieved 10 Jul 2011 from: http://www.uwsp.edu/education/lwilson/curric/newtaxonomy.htm

Bell, B. & McNamara, J. (2005). How to Avoid Using Stupid Agents to Train Intelligent People, *Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC) 2005*. Orlando FL.

Benner, P. (1984). *From Novice to Expert: Promoting Excellence and Power in Clinical Nursing Practice*. Menlo Park, CA: Addison-Wesley.

Benner, P. (2004). Using The Dreyfus Model of Skill Acquisition to Describe and Interpret Skill Acquisition. *Bulletin of Science Technology Society*, 24, 188-199. Retrieved 23 Jun 2012 from http://bst.sagepub.com.proxy.lib.odu.edu/content/24/3/188.full.pdf+html

Bernatovich, D. (1999). The Effect of Presence on the Ability to Acquire Spatial Knowledge in Virtual Environments, Master's Thesis, Monterey, CA: Naval Postgraduate School.

Bowers, C., Hussain, T., Roberts, B., Cannon-Bowers, J., & Blair, L. (2009 unpublished manuscript). Preparing to Practice: The Use of a Game-Based Simulation as a Pre-Training Intervention for Recruits. University of Central Florida and BBN Technologies.

Brannick, M., Prince, C. and Salas, E. (2005). Can PC-Based Systems Enhance Teamwork in the Cockpit? *The International Journal of Aviation Psychology*, 15 (2), 173–187.

Brenner, T., Sheehan, K., Arthur, W., and Bennett, W. (1998). Behavioral and Cognitive Task Analysis Integration for Assessing Individual and Team Work Activities, presented during *Evaluating Innovations in Training Assessment and Occupational Modeling Technologies Symposium,* San Antonio, TX. Retrieved 29 June 2012 from http://www.ijoa.org/currentwork/98symposium/brenneretal.html

Campbell, D.T. & Stanley, J.C. (1963). Experimental and Quasi-Experimental Designs for Research. Reprinted from *Handbook of Research on Teaching*, Houghton Mifflin Company, Dallas, TX.

Clark, D. R. (2004). Instructional System Design. Retrieved January 25, 2011 from http://www.nwlink.com/~donclark/hrd/sat.html

Cohn, J., Stanney, K.M., Milham, L., Bell Carroll, M., Jones, D., Sullivan, J., & Darken, R. (2008). Training Effectiveness Evaluation: From Theory to Practice. In D. Schmorrow, J. Cohn, & D. Nicholson (Eds.), *The Handbook of Virtual Environment Training: Understanding, Predicting and Implementing Effective Training Solutions for Accelerated and Experiential Learning,* 3 (2), 157-172). Aldershot, Hampshire, UK: Ashgate Publishing.

Danko, Andrew (2010, Oct). Executive Officer (XO), Helicopter Maritime Strike Weapons School Atlantic (HSMWSL), (M.A. White, Interviewer).

Danko, Andrew (2011, Nov). Commanding Officer (CO), Helicopter Maritime Strike Weapons School Atlantic (HSMWSL), (M.A. White, Interviewer).

Danko, Andrew (2012, Jan). Commanding Officer (CO), Helicopter Maritime Strike Weapons School Atlantic (HSMWSL), (M.A. White, Interviewer).

Davis, Dave (2011, Dec). Surface Warfare (SUW) Curriculum Development Officer, Helicopter Maritime Strike Weapons School Atlantic (HSMWSL), (M.A. White, Interviewer).

Dennis, K. A. & Harris, D. (1998). Computer-Based Simulation as an Adjunct to Ab Initio Flight Training. *The International Journal of Aviation Psychology, 8,* 277–292.

Dreyfus H.L. & Dreyfus, S.E. (1986). *Mind Over Machine: The Power of Human Intuition and Expertise in the Era of the Computer.* Oxford: Blackwell.

Dreyfus, S.E. (1981). Four Models v Human Situational Understanding: Inherent Limitations on the Modelling of Business Expertise. US Air Force Office of Scientific Research, Contract no. F49620-79-C-0063.

Dreyfus, S.E. (2004). The Five-Stage Model of Adult Skill Acquisition, *Bulletin of Science Technology & Society* 2004, 24, 177. Retrieved 24 June 2012 from http://bst.sagepub.com.proxy.lib.odu.edu/content/24/3/177.full.pdf+html

Driskell, J.E., Willis, R.P., & Cooper, C., (1992). Effect of Overlearning on Retention. *Journal of Applied Pyschology*, 77, 615-22.

Ellis, Steven R. (1994). What are Virtual Environments? *IEEE Computer Graphics*, January/February 1994. 14 (1), 17-22, retrieved 23 Jun 2012 from: http://www.computer.org/portal/web/csdl/doi/10.1109/38.250914

Engel, K., Langkamer, K.,Estock, J., Orvis, K., Salas, E., Bedwell, W. & Conkey, C., (2009). Investigating the Effectiveness of Game-Based Approaches to Training, *Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC)2009*. Orlando, FL.

Eraut, M. (1994). *Developing Professional Knowledge and Competence*. London: Falmer

Ewell, P.T. (2001). *Accreditation and Student Learning Outcomes: A Proposed Point of Departure*. Washington, DC: Council for Higher Education Accreditation (CHEA). Retrieved 29 June 2012 from: http://www.chea.org/award/StudentLearningOutcomes2001.pdf

Fishwick, P. (2010). What Is Simulation? Thu Oct 19 10:30:41 EDT 1995; retrieved 22 Aug 2010 from: http://www.cise.ufl.edu/~fishwick/introsim/node1.html

Flin, R. & O'Connor, P. (2001). Applying Crew Resource Management on Offshore Oil Platforms, *Improving Teamwork in Organizations: Applications of Resource Management Training*, Salas, E., Bowers, C. and Edens, E., editors. (217-234). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Gallo, A., Glass, J., Frye, C., Matthews, C. & Kotick, D. (2004). Stick and Rudder Training for the Mind, *Interservice/Industry Training, Simulation and Education Conference (I/ITSEC)* 2004, Orlando, FL.

Gallo, A. W., Glass, J. P., Frye, C., Douglass, R., Velez, C., Buckley, J., & Cipolla, L. (2006). ASW VAST MRT3: The Tip of the Virtual Spear. *Interservice/Industry Training, Simulation and Education Conference (I/ITSEC) 2006*, Orlando, FL.

Gamor, K. (2001). Moving Virtuality into Reality: A Comparison Study of the Effectiveness of Traditional and Alternative Assessments of Learning in a Multisensory, Fully Immersive Physics Program, PhD Dissertation, Fairfax, VA: George Mason University.

Gately, M., Watts, S., and Pleban, R. (2002). Assessing Decision-Making Skills in Virtual Environments, Technical Report 1130, U.S. Army Research Institute for Behavioral and Social Sciences, Alexandria, VA.

Giovanni, D. de, Roberts, T. and Norman, D. (2009). Relative Effectiveness of High-Versus Low-Fidelity Simulation in Learning Heart Sounds, *Medical Education*, Jul 2009, 43(7), 661-668. Retrieved from http://proxy.lib.odu.edu/.

Glass, J. (2010, July). MRT3 Program Manager. (M.A. White, Interviewer).

Global Security (2010). MH-60R Seahawk [ex Strikehawk], retrieved 3 Oct 2010 from: http://www.globalsecurity.org/military/systems/aircraft/sh-60r.htm

Greitzer, F., Kuchar, O. and Huston, K. (2007). Cognitive Science Implications for Enhancing Training Effectiveness in a Serious Gaming Context. *ACM Journal of Educational Resources in Computing*, 7 (3), Article 2 (August 2007). DOI=10.1145/1281320.1281322 http://doi.acm.org/10.1145/1281320.1281322.

Gruchalla, K. (2004). Immersive Well-Path Editing: Investigating the Added Value of Immersion, *Proceedings of the 2004 Virtual Reality (VR'04), IEEE Virtual Reality 2004* March 27-31, Chicago, IL. Retrieved from http://proxy.lib.odu.edu/.

Hale, K. S. (2006). Enhancing Situational Awareness Through Haptics Interaction In Virtual Environment Training Systems, PhD Dissertation, University of Central Florida, Orlando, FL.

Hall, Richard (1998). Virtual Statistician, *Psychology World*, retrieved 2 November 2010 from: http://web.mst.edu/~psyworld/virtualstat.htm

Hersey, P. and Blanchard, K. H. (1977). Management of Organizational Behavior 3rd Edition– Utilizing Human Resources. New Jersey/Prentice Hall.

Hill, J. R., & Hanaffin, M. J. (2001). Teaching and Learning in Digital Environments: The Resurgence of Resource-Based Learning. *Educational Technology, Research and Development*, 49 (3), 37-52.

Holt, R. W., Boehm-Davis, D. A. & Beaubien, J. M. (2001). Evaluating Resource Management Training, *Improving Teamwork in Organizations: Applications of Resource Management Training*, Salas, E., Bowers, C. and Edens, E., editors. (165-188). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers, 2001.

Hougland, E., Paterson, D., Owens, J., Oser, R. and Bergondy, M. (1998). Networking for Naval Aviation Training: High Level Architecture Implementation Cost Analysis, Special Report 98-005B, Naval Air Warfare Center Training Systems Division, Orlando, FL.

Hussain, T., Roberts, B., Bowers, C., Cannon-Bowers, J., Menaker, E., Murphy, C., Pounds, K., Koenig, A., Wainess, R., Lee, J. (2009). Designing and Developing Effective Training Games for the US Navy, *Interservice/Industry Training, Simulation and Education Conference (I/ITSEC)* 2009. Orlando, FL.

Hussain, T., Bowers, C., Blasko-Drabik, H. (2011). Validating Cognitive Readiness on Team Performance Following Individual Game-Based Training (in press).

Jones, L. & Tukey, J. (2000). A Sensible Formulation of the Significance Test; Retrieved from: http://www.psych.illinois.edu/~broberts/Jones%20&%20Tukey%20on%20significance%20testing,%202000.pdf.

Kanki, B. G. & Smith, G. M. (2001). Training Aviation Communication Skills, *Improving Teamwork in Organizations: Applications of Resource Management Training*, Salas, E., Bowers, C. & Edens, E., editors. (pp 95-127). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

King, W. (2009). Virtual Battle Space 2 Army Gaming System Debuts, www.army.mil, *Official Homepage of the United States Army*, 2/26/2009. Retrieved from http://www.army.mil/-news/2009/02/26/17502-virtual-battle-space-2-army-gaming-system-debuts.

Koonce, J. M., Moore, S. L., & Benton, C. J. (1995). Initial Validation of a Basic Flight Instruction Tutoring System (BFITS). In R. S. Jensen & L. A. Rakovan (Eds.), *Proceedings of the Eighth International Symposium on Aviation Psychology* (1037–1040). Columbus: Ohio State University.

Lane, D., Lu, J., Peres, C. and Zitec, E. (2008). Online Statistics: An Interactive Multimedia Course of Study, Rice University, University of Houston, Clear Lake, and University of Houston, Downtown, retrieved from http://onlinestatbook.com

Lintern, G., Roscoe, S. N., Koonce, J. M., & Segal, L. (1990). Transfer of Landing Skills in Beginning Flight Training. *Human Factors, 32,* 319–327.

Lintern, G., Taylor, H. L., Koonce, J. M., Kaiser, R. H., & Morrison, G. A. (1997). Transfer and Quasi-Transfer Effects of Scene Detail and Visual Augmentation in Landing Training. *The International Journal of Aviation Psychology,* 7, 149–169.

Mania, K., and Chalmers, A. (2001). The Effects of Levels of Immersion on Memory and Presence in Virtual Environments: A Reality Centered Approach, *CyberPsychology & Behavior*, 4 (2), 247-264.

Morris, C., Hancock, P. & Shirkey, E. (2004). Motivational Effects of Adding Context Relevant Stress in PC-Based Game Training. *Military Psychology*, 135-147.

Morris, C., and Tarr, R. (2002). Templates for Selecting PC-Based Synthetic Environments for Application to Human Performance Enhancement and Training, *Proceedings of the IEEE Virtual Reality 2002 (VR '02)*, 109-115, Orlando, FL.

MSCO, (2010). DoD Modeling And Simulation (M&S) Glossary, The Department of Defense (DoD) Modeling and Simulation Coordination Office (M&S CO); retrieved 22 Aug 2010 from: http://www.msco.mil/files/Draft_MS_Glossary_March_B_version.pdf

MSCO, (2012). DoD Modeling And Simulation (M&S) Glossary, The Department of Defense (DoD) Modeling and Simulation Coordination Office (M&S CO); retrieved 23 Jun 2012 from http://www.msco.mil/MSGlossary_TRM_EG.html

Norlander, K. (2001). Emergent Leadership on Collaborative Tasks in Distributed Virtual Environments, Master's Thesis, Naval Post-Graduate School, Monterey, CA.

O'Hanlon, R., (2010, July). RADM, Commander Naval Air Force U.S. Atlantic Fleet (COMNAVAIRLANT). (G.R. Jones and M.A. White, Interviewers).

Parsons, Tracy (2011, Dec). LCDR, MRT3 Project Officer, Naval Air Warfare Command Training Systems Division (NAWC TSD). (M.A. White, Interviewer).

Perkins, D.N. & Salomon, G. (1992). Transfer of Learning. *International Encyclopedia of Education, Second Edition*. Oxford, England: Pergamon Press. Retrieved from http://learnweb.harvard.edu/alps/thinking/docs/traencyn.htm.

Porter, C. O., Hollenbeck, J. R., Ilgen, D. R., Ellis, A. P., West, B. J., & Moon, H. (2003). Backing up Behaviors in Teams: The Role of Personality and Legitimacy of Need. *Journal of Applied Psychology*, 88 (3), 391-403.

Prince & Jentsch (2001). Aviation Crew Resource Management Training With Low-Fidelity Devices. *Improving Teamwork in Organizations: Applications of Resource Management Training*, Salas, E., Bowers, C. and Edens, E., editors. (147-164). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers, 2001.

Proctor, M., Panko, M. and Donovan, S. (2004). Considerations for Training Team Situation Awareness and Task Performance Through PC-Gamer Simulated Multi-Ship Helicopter Operations, *The International Journal of Aviation Psychology*, Vol. 14, Num. 2, 2004, 191 -205.

Rantanen, E. and Talleur, D. (2005). Incremental Transfer and Cost Effectiveness of Ground-based Flight Trainers in University Aviation Programs, *Proceedings Of The Human Factors And Ergonomics Society 49th Annual Meeting, 2005.*

Ricci, K., Salas, E. & Cannon-Bowers, J. (1996). Do Computer-Based Games Facilitate Knowledge Acquisition And Retention? *Military Psychology*, 295-307.

Robinson, A. & Mania, K. (2007). Technological Research Challenges of Flight Simulation and Flight Instructor Assessments of Perceived Fidelity, *Simulation & Gaming*, 38 (1), 112-135.

Roessingh, J. (2003). Transfer of Manual Flying Skills from PC-Based Simulation to Actual Flight—Comparison of In-Flight Measured Data and Instructor Ratings, *The International Journal Of Aviation Psychology*, 15 (1), 67–90.

Royal College of Obstetricians and Gynaecologists (2006). Basic Log Book. Retrieved April 26, 2009 from http://www.rcog.org.uk/files/rcog-corp/uploaded-files/ED-Basic-logbook.pdf, as cited in UMNDJ (2011)

Salas, E. & Cannon-Bowers, J. (2001). The Science of Training: A Decade of Progress. *Annual Review of Psychology*, 471-499.

Salas, E., Rosen, M., Held, J., and Weissmuller, (2009). Performance Measurement in Simulation-Based Training A Review and Best Practices, *Simulation Gaming 2009*, 40, 328 originally published online Dec 22, 2008; retrieved May 4, 2010 from http://sag.sagepub.com.

Seamster, T. & Kaempf, G., (2001). Identifying Resource Management Skills for Airline Pilots, *Improving Teamwork in Organizations: Applications of Resource Management Training*, Salas, E., Bowers, C. and Edens, E., editors (9-30). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers, 2001

Sikorsky SH-60 Seahawk: Encyclopedia II - Sikorsky SH-60 Seahawk - Models & Missions (2010). retrieved 3 Oct 2010 from: http://www.experiencefestival.com/a/Sikorsky_SH-60_Seahawk_-_Models_amp_Missions/id/4724204

Smed, J., Kaukoranta, T. & Hakonen, H. (2002). A Review on Networking and Multiplayer Computer Games, Technical Paper TR454 from the Turku Centre for Computer Science, Turku, Finland: University of Turku. Retrieved 8 August 2010 from http://staff.cs.utu.fi/~jounsmed/papers/TR454.pdf

Smith-Jentsch, K., Baker, D. & Salas, E. (2001). Uncovering Differences in Team Competency Requirements: The Case of Air Traffic Control Teams, *Improving Teamwork in Organizations: Applications of Resource Management Training*, Salas, E., Bowers, C. and Edens, E., editors. (31-54). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers, 2001.

Smith, R.D., (2008). Investigating the Disruptive Effect of Computer Game Technologies on Medical Education and Training, PhD Dissertation. Adelphi, MD: University of Maryland University College.

Steinaker, N. & Bell, M. (1979). The Experiential Taxonomy: A New Approach to Teaching and Learning. New York: Academic Press.

Sterling, G., Magee, L., and Wallace, P., (2000). Virtual Reality Training - A Consideration for Australian Helicopter Training Needs? *Proceedings SimTecT 2000*, Sydney, Australia.

Stoudenmire, E., White, M. & Roy, K. (2006). Assessment and Validation of Gaming Technology as Applied to Training: Current State and the Way Ahead, *Simulation Interoperability Workshop 2006*, Orlando, FL.

Taylor, H. L., Lintern, G., Hulin, C. L., Talleur, D. A., Emanuel, T. W.,&Phillips, S. I. (1999). Transfer of Training Effectiveness of a Personal Computer Aviation Training Device. *The International Journal of Aviation Psychology, 9,* 319–335.

Testa, B.M. (2009). The Army's Training Weapon: Serious Games, *Workforce Management Online*, January 2009. Retrieved from http://www.workforce.com/section/11/feature/26/08/34/index.html.

Therrien, Emile, and Cooper, Justin P. (2010, September). Instructors, Helicopter Maritime Strike Weapons School Atlantic (HSMWSL), (M.A. White, Interviewer)

Turpin, D. and Welles, R. (2006). Analysis of Simulator-based Training Effectiveness through Driver Performance Measurement, *Interservice/Industry Training, Simulation and Education Conference (I/ITSEC)* 2006. Orlando, FL.

UMDNJ (2011). Staged Models of Skills Acquisition, *University of Medicine and Dentistry of New Jersey web page*. Retrieved 22 Mar 2011 from http://www.umdnj.edu/idsweb/idst5340/models_skills_acquisition.htm

van Reeuwijk, L.P. and Houba, V.J.G. (1998). Guidelines for Quality Management in Soil and Plant Laboratories. (FAO Soils Bulletin - 74), Chapter 6, retrieved 5 July 2012 from http://www.fao.org/docrep/W7295E/w7295e08.htm#6 basic statistical tools

Vozenilek J., Huff J.S., Reznek M. and Gordon J.A. (2004). See One, Do One, Teach One: advanced technology in medical education. *Academic Emergency Medicine: Official Journal for the Society of Academic Emergency Medicine 2004 Nov;11(11):1149-54*. Retrieved 15 Sep 2012 from http://www.ncbi.nlm.nih.gov/pubmed/15528578

Wallace, P.R. (1997). Training System Definition. *ACT: Proceedings of SimTecT97*, Canberra, Australia

Walker, A., Carpenter, T., Moss, J., Switzer, F., Hoover, A., and Muth, E. (2009). The Evaluation of Virtual Environment Training for a Building Clearing Task, *Human Factors and Ergonomics Society Annual Meeting Proceedings,* 53 (18), 1206-1209.

Walpole, R., Myers, R., Myers, S., and Ye, K. (2007). Probability & Statistics for Engineers & Scientists, 8th Edition, Prentice Hall, Upper Saddle, NJ.

Webber, S. S. (2002). Leadership and Trust: Facilitating Cross-Functional Team Success, *Journal of Management Development*, 21, 201-214.

Wilson, K., Bedwell, W., Lazzara, E., Salas, E., Burke, S., Estock, J., Orvis, K., and Conkey, C., (2009). Relationships between Game Attributes and Learning Outcomes, Review and Research Proposals, *Simulation & Gaming,* 40 (2), 217-266.

Witmer, B.G. & Singer, M.J. (1998). Measuring Presence in Virtual Environments: A Presence Questionnaire, *Presence*, 7 (3), 225-240.

Wright, G. (2000). Helicopter Urban Navigation Training Using Virtual Environments, Master's Thesis, Naval Postgraduate School: Monterey, CA.

Youngblut, C. and Huie, O. (2003). The Relationship between Presence and Performance in Virtual Environments: Results of a VERTS Study, *Proceedings of the IEEE Virtual Reality 2003* (VR'03), Los Angeles, CA.

Yousoff, A., Crowder, R., Gilbert, L., and Wills, G. (2009). A Conceptual Framework for Serious Games, *2009 Ninth IEEE International Conference on Advanced Learning Technologies,* 21-23.

APPENDICES

APPENDIX A INSTITUTIONAL REVIEW BOARD (IRB) EXEMPTION APPROVAL LETTER

V M A S C
Virginia Modeling, Analysis and Simulation Center

OLD DOMINION UNIVERSITY

June 7, 2012

Frederic D. McKenzie, Ph.D.
1303 ECSB
4700 Elkhorn Avenue
Norfolk, VA 23529

Re: Proposal 12-6

Dear Dr. McKenzie:

The Frank Batten College of Engineering and Technology IRB College Committee has approved your request for an IRB exemption for the attached College Committee Proposal Number 12-6.

Sincerely,

John A. Sokolowski, Ph.D.
Executive Director
Virginia Modeling, Analysis & Simulation Center
1030 University Blvd.
Suffolk, VA 23435
757-686-6232 office
757-686-6214 fax
jsokolow@odu.edu

The *Discipline* of M&S™

1030 University Blvd., Suffolk, VA 23435 · 757.686.6200 · www.vmasc.odu.edu

APPENDIX B APPROVED APPLICATION FOR EXEMPT RESEARCH

Proposal Number: 12-06
(To Be Assigned by the College Committee or IRB)

APPENDIX B
OLD DOMINION UNIVERSITY
APPLICATION FOR EXEMPT RESEARCH

Note: For research projects regulated by or supported by the Federal Government, submit 10 copies of this application to the Institutional Review Board. Otherwise, submit to your college human subjects committee.

Responsible Project Investigator (RPI) **The RPI must be a member of ODU faculty or staff who will serve as the project supervisor and be held accountable for all aspects of the project. Students cannot be listed as RPIs.**		
First Name: Frederic	**Middle Initial:** D.	**Last Name:** McKenzie
Telephone: (757) 683-5590	**Fax Number:** (757) 683-3200	**E-mail:** rdmckenz@odu.edu
Office Address: 1303 ECSB, 4700 Elkhorn Avenue		
City: Norfolk	**State: VA**	**Zip: 23529**
Department: Modeling, Simulation, & Visualization Engineering Dept.; Joint Appointment, Electrical & Computer Engineering Dept.	**College: Engineering**	
Complete Title of Research Project: Delivering Contextually Appropriate Opportunities to Practice		**Code Name (One word):**
Investigators **Individuals who are directly responsible for any of the following: the project's design, implementation, consent process, data collection, and data analysis. If more investigators exist than lines provided, please attach a separate list.**		
First Name: Michael	**Middle Initial: A.**	**Last Name: White**
Telephone: 757-651-5619	**Fax Number:**	**Email: mwhite@alionscience.com**
Office Address: 5365 Robin Hood Road Suite 100		
City: Norfolk	**State: VA**	**Zip: 23513**
Affiliation: __Faculty _X_Graduate Student __ Undergraduate Student __Staff __Other		
First Name:	**Middle Initial:**	**Last Name:**
Telephone:	**Fax Number:**	**Email:**
Office Address:		
City:	**State:**	**Zip:**
Affiliation: __Faculty __Graduate Student __ Undergraduate Student __Staff __Other		
List additional investigators on attachment and check here: __		
Type of Research		

Proposal Number: _12-06_____
(To Be Assigned by the College Committee or IRB)

1. This study is being conduced as part of (check all that apply):

_ Faculty Research
_X Doctoral Dissertation
_ Masters Thesis
_ Non-Thesis Graduate Student Research
_ Honors or Individual Problems Project
_ Other______________________

Funding

2. Is this research project externally funded or contracted for by an agency or institution which is independent of the university? Remember, if the project receives ANY federal support, then the project CANNOT be reviewed by a College Committee and MUST be reviewed by the University's Institutional Review Board (IRB).

___Yes **(If yes, indicate the granting or contracting agency and provide identifying information.)**
__X_No

Agency Name:
Mailing Address:
Point of Contact:
Telephone:

Research Dates

3a. Date you wish to start research (MM/DD/YY) __5__/__21__/__2012__
3b. Date you wish to end research (MM/DD/YY) __5__/__20__/__2013__

Human Subjects Review

4. Has this project been reviewed by any other committee (university, governmental, private sector) for the protection of human research participants?
___Yes
_X__No

4a. If yes, is ODU conducting the primary review?
__Yes
__No (If no go to 4b)

4b. Who is conducting the primary review?

Proposal Number: 12-06
(To Be Assigned by the College Committee or IRB)

5. Attach a description of the following items:

_X_Description of the Proposed Study
__Research Protocol
__References
__ Any Letters, Flyers, Questionnaires, etc. which will be distributed to the study subjects or other study participants
__If the research is part of a research proposal submitted for federal, state or external funding, submit a copy of the FULL proposal

Note: The description should be in sufficient detail to allow the Human Subjects Review Committee to determine if the study can be classified as EXEMPT under Federal Regulations 45CFR46.101(b).

Exemption categories

6. Identify which of the 6 federal exemption categories below applies to your research proposal and explain why the proposed research meets the category. Federal law 45 CFR 46.101(b) identifies the following EXEMPT categories. Check all that apply and provide comments.

SPECIAL NOTE: The exemptions at 45 CFR 46.101(b) do not apply to research involving prisoners, fetuses, pregnant women, or human in vitro fertilization. The exemption at 45 CFR 46.101(b)(2), for research involving survey or interview procedures or observation of public behavior, does not apply to research with children, except for research involving observations of public behavior when the investigator(s) do not participate in the activities being observed.

____(6.1) Research conducted in established or commonly accepted educational settings, involving normal educational practices, such as (i) research on regular and special education instructional strategies, or (ii) research on the effectiveness of or the comparison among instructional techniques, curricula, or classroom management methods.
Comments:

____(6.2) Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures or observation of public behavior, unless: (i) Information obtained is recorded in such a manner that human subjects can be identified, directly or through identifiers linked to the subjects; AND (ii) any disclosure of the human subjects' responses outside the research could reasonably place the subjects at risk of criminal or civil liability or be damaging to the subjects' financial standing, employability, or reputation.
Comments:

____(6.3) Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures, or observation of public behavior that is not exempt under paragraph (b)(2) of this section, if:
(i) The human subjects are elected or appointed public officials or candidates for public office; or (ii) federal statute(s) require(s) without exception that the confidentiality of the personally identifiable information will be maintained throughout the research and thereafter.
Comments:

Proposal Number: 12-06
(To Be Assigned by the College Committee or IRB)

__X__(6.4) Research, involving the collection or study of existing data, documents, records, pathological specimens, or diagnostic specimens, if these sources are publicly available or if the information is recorded by the investigator in such a manner that subjects cannot be identified, directly or through identifiers linked to the subjects.
Comments:

Instructors at Helicopter Strike Maritime Weapons School Atlantic (HSMWSL) regularly conduct Helicopter Aircrew Readiness Program training in preparation of fleet squadron aircrew members' deployment. Type of data includes information collected from performance in a flight simulator and in live aircraft.

Weapons School Instructors will mail the PI data via USPS Priority Mail. To ensure participant anonymity, the PI will request photocopy of data collected with names whited out, but with an identifier to ensure proper association of each participant's data with the appropriately sequenced training event; i.e., Practice 1, Practice 2, or Live event. If data provided inadvertently contains information through which subjects can be identified, PI will white out the identifying information, assign participant training event sequence identifier, make a photocopy of the sanitized document, and destroy the original. Raw data will be stored in a locked cabinet. All published data will be anonymous and aggregated. PI will ensure all raw data associated with this study is destroyed not later than five years after project completion.

___ (6.5) Does not apply to the university setting; do not use it

____(6.6) Taste and food quality evaluation and consumer acceptance studies, (i) if wholesome foods without additives are consumed or (ii) if a food is consumed that contains a food ingredient at or below the level and for a use found to be safe, or agricultural chemical or environmental contaminant at or below the level found to be safe, by the Food and Drug Administration or approved by the Environmental Protection Agency or the Food Safety and Inspection Service of the U.S. Department of Agriculture.
Comments:

PLEASE NOTE:

1. You may begin research when the College Committee or Institutional Review Board gives notice of its approval.
2. You MUST inform the College Committee or Institutional Review Board of ANY changes in method or procedure that may conceivably alter the exempt status of the project.

Responsible Project Investigator (Must be original signature) **Date**

APPENDIX C POST-TRAINING QUESTIONNAIRE

Crew#_________ Participant#_________ Date_________

DO YOU THINK TRAINING WITH MRT3 HELPED IMPROVE YOUR CREW'S READINESS TO PERFORM ON THE LIVE RANGE? WHY OR WHY NOT?

FOR WHICH TASKS DO YOU THINK MRT3 COULD BE USED FOR TRAINING AND READINESS CREDIT?

WHAT DID YOU LIKE LEAST ABOUT TRAINING WITH MRT3?

WHAT DID YOU LIKE MOST ABOUT TRAINING WITH MRT3?

WHAT CHANGES WOULD YOU SUGGEST TO IMPROVE MRT3?

APPENDIX D TRAINING TRANSFER/EFFECTIVENESS LITERATURE

Transfer/ Effectiveness Study	Approach	Device	Objective
Alexander et al. (2005)	Survey	Massive multiplayer games (MMPGs)	Fidelity, immersion and transfer
Bernatovich (1999)	Experiment	Audio-visual Virtual Environment (VE)	Audio-visual cues and immersion
Brannick et al. (2005)	Experiment	PC-based flight simulator	Crew resource management
Dennis & Harris (1998)	Experiment	PC-based flight simulator	Basic flying skills
Engel et al. (2009)	Experiment	Game-based test bed	Feature control and learning outcomes
Gamor (2001)	Experiment	Game-based test bed	Virtual versus traditional assessment methods
Gately et al. (2002)	Experiment	Constructive simulation test bed	Decision making skills
Giovanni et al. (2009)	Experiment	High fidelity sound simulator vs low fidelity (CD)	Heart sound recognition
Greitzer et al. (2007)	Experiment	Serious game	Cyber security
Gruchalla (2004)	Experiment	Immersive virtual environment (IVE)	Oil well-path planning
Hussain et al. (2009)	Experiment	Serious game	Situational awareness, communications, navi-gation, decision making
Koonce et al. (1995)	Experiment	PC-based flight simulator	Landing skills
Lintern et al. (1990, 1997)	Experiment	PC-based flight simulator	Landing skills
Mania & Chalmers (2001)	Experiment	Desktop PC VE vs Desktop VE w/HMD	Memory recall and awareness
Morris et al (2004)	Experiment	PC-based game	Stress and transfer
Morris & Tarr (2002)	Experiment	PC-based game	Cognitive skills
Norlander (2001)	Experiment	Collaborative virtual environment (CVE)	Small unit leadership
Prince & Jentsch (2001)	Experiment	Low-fidelity PC-based simulation	Decision making, coordination, workload management
Proctor et al. (2004)	Experiment	PC-based flight simulator	Multi-crew team training
Rantenan & Talleur (2005)	Survey	Flight simulators	Transfer quantification methods
Ricci et al. (1996)	Experiment	PC-based game	Knowledge acquisition and retention
Robinson & Mania (2007)	Experiment	Full fidelity simulator	Visual fidelity
Roessingh (2003)	Experiment	PC-based flight	Aerobatic flight

		simulator	maneuvers
Smed et al. (2002)	Survey	Military simulations, networked VEs, and multiplayer computer games	Training system architecture and software resource management
Sterling et al. (2000)	Experiment	VR-based part task trainer vs full fidelity simulator	Takeoffs, Landings, day/night flight
Taylor et al. (1999)	Experiment	PC-based aviation training device (PCATD)	Instrument flight
Turpin & Wells (2006)	Experiment	PC-based flight simulator	Driving skills
Walker et al. (2009)	Experiment	HMD VE	Room clearing
Wilson et al. (2009)	Survey	PC-based games	Game attributes vs learning objectives
Witmer & Singer (1998)	Survey	Virtual Environments	Immersion and Presence
Wright (2000)	Experiment	PC-based VE	Helicopter navigation

APPENDIX E DEMOGRAPHICS QUESTIONNAIRE

Date________

All information will be kept in strict confidence. Only aggregate data will be reported.

Please provide the following information.

Name (optional)	
Rank	
Age	
Unit	
Duty Position	
Time in Service	
Time w/current crew	
Time in A/C Type	
Time since last simulator ASW event	
Time since last live ASW event	
Time since last simulator SUW event	
Time since last live SUW event	

APPENDIX F EXAMPLE MASTER FLIGHT BRIEF/GRADE SHEET

CHSMWPINST 3500.6A/
CHSMWPINST 3500.5A

Classification when filled in: UNCLASS/ CLASSIFIED /SECRET (circle one)

MASTER FLIGHT BRIEF/GRADE SHEET

ASW 101 (Crew)	**Title:**	PRE TORPEX	**Hours:** 2.0(S) 60B/R
	Objective:	Successfully complete proper employment of air launched torpedoes. Commencing with a direct path contact, track, classify, obtain attack criteria using available sensors, conduct an attack within weapons acquisition ranger and employ proper re-attack procedures.	

DATE:	**SQUADRON/DET:**	**HARP:**	

Evaluee(s): Pilot – ATO- AW/SO-	**Current ACWT Level:**	**Other Crew:** Pilot – ATO – AW/SO -	**Evaluator(s):**

REQUIRED ITEMS	KEY POINTS	MOP
1. DISCUSSION ITEMS	• MAD System (SH-60B) • Alpha/Kilo index • Mk-46/50/54 characteristics, presets, placement • Tactical oceanography • Simulated targets: MK-30, MK-39 • Submarine tactics/counter measures • Waterspace management • Clearing ping (MH-60R) • ASW 203 (TORPEX qual event)	0 1 2 3 4 5
2. BRIEFING	• NATOPS/mission brief • Environmentals/PC-IMAT • Current OPTASK direction/ROE/Air- + Water-space mgmt • Current intelligence • Ordnance loadout (buoys/smokes/SUS/torpedoes/etc.)	0 1 2 3 4 5
3. START UP/SYSTEM INIT/ DEPART	• Initialize mission systems (non LINK environment) • ESM configuration verified • Crypto verified • Flight path/threat sector consideration/EMCON • Weapons/ tactical checklists	0 1 2 3 4 5
4. LOCALIZATION	• Datum creation/pattern determination/time late/LOB search pattern/screen delta/AMP optimized pattern • Pattern orientation WRT approach path/force protection • Other platform data utilization • Counter-detection considered • Acoustic integration times/dome depth • Aural information utilization • Verniering technique and procedures • Non-acoustic sensor employment (ESM/RADAR/IFF/MAD/FLIR/LINK-16/visual • Contact/source recognition	0 1 2 3 4 5
5. TRACKING a. min time 20 mins	• Sonobuoy utilization/expansion tactics • ALFS employment (MH-60R) • Actual sensor range determination(s) • Doppler tracking techniques • ATT employment • Lost contact	0 1 2 3 4 5
6. CLASSIFICATION	• Aural/acoustic signature recognition • Sensor correlation • Tactical reports (contact class, hull type, DATALINK track, LINK 16 track)	0 1 2 3 4 5

Classification when filled in: UNCLASS/ CLASSIFIED /SECRET (circle one)

7. ATTACK	• Attack authorization • Attack criteria • Set-up/flight profile/early alert • Weapon preset/delivery/placement • Utilization of other assets (if applicable) • Post attack TTPs/DOGBOX/etc • Torpedo hit assessment	0 2 4 6 8 10
8. RE-ATTACK	• Contact maintained/re-gained • Attack criteria • Mutual interference considerations • Weapon preset/delivery/placement • Post attack TTPs/DOGBOX/etc • Torpedo hit assessment	0 1 2 3 4 5
9. FLIGHT SAFETY	• Plot stabilization/MOT technique • Fuel management/ops normal reports • Emergency procedures	0 1 2 3 4 5
5. CRM	• Utilized mission analysis to develop and coordinate crew and aircraft resources to effectively manage ASW scenario • Utilized decision making to choose a course of action using logical and sound judgment based on available information to effectively execute successful ASW prosecution • Performed effective communication by clearly and accurately sending information, instructions or commands, and providing useful feedback to crew members and coordinated ASW assets as applicable • Maintained situational awareness with a high degree of accuracy between the perception of the current tactical ASW environment and the reality of the situation • Performed effective tactical communications and reporting	0 1 2 3 4 5
6. DEBRIEF	• BAW/CRM • Admin/Tac-admin • Alibis/Safety-of-flight issues • Mission reconstruction • Goods and others (self-evaluation) • Post mission paperwork completion (purple, shot validation)	COMP/INCOMP

Sim Duration:	**Qual**		**No Qual**		**Incomplete**		**MOE – Total Pts___ ÷ 11 = ___**
	ACWT QUAL			**N/A**			

Scenario: (Consider Evaluee(s) Op Area)	**Weather:**
General Comments:	
Items Requiring More Practice:	

Evaluator(s) Signature:	**Evaluee(s) Signature:**

Miscellaneous Remarks:

APPENDIX G EXAMPLE MASTER FLIGHT BRIEF/GRADE SHEET GUIDANCE

Score of 5, 10 or 15 – All performance standards met/no Safety-of-flight issues, appropriate corrections of a minor nature.
- **Discussion items** – All (100%) items answered correctly.
- **Flight planning/Briefing** – MPS/JMPS/Flight planning complete. No rework required.
- **System utilization** – All a/c systems utilized effectively to maximum extent from beginning to end of event. All CSG/Ship restrictions/considerations taken into account.
- **Communication** – All safety-of-flight and tactical communications were completed correctly. 100% of all linked symbols verified and correlated with ship.
- **Navigation** – All required chart work is completed prior to brief. For low level routes and NVG routes, all checkpoints are found and recognized within 1 ½ min.
- **Detection/Localization**
  ASW – Contact established including course/speed within 20 min of receiving datum.
  SUW – Track established on 100% of contacts in threat sector vital area within 15 min of launch or 10 min of tasking if already airborne.
- **Counter detection** – Counter detection tactics successfully incorporated into mission brief and executed IAW mission.
- **Search**
  ASW – HVU location incorporated into search pattern, search pattern utilizes all applicable environmental data and threat data.
  SUW – HVU location incorporated into tactical solution. Threat sector(s) established IAAW mission objectives.
- **Classification**
  ASW – Contact is classified correctly with correct source.
  SUW – Contact is classified correctly using all available sensors.
- **ROE** – Action is taken appropriate to ROE.
- **Tracking**
  ASW – Track accurate within 10°/5 kts when compared to known target track.
  SUW – Maintained track on the contacts in a 60 nm^2 area (95%) in moderate shipping density, 100% in light shipping density.
- **Attack/Target engagement/Re-attack** – If no worksheet is utilized then use the following:
  ASW – Utilize applicable TDT guidance. Attach within 2 min of attack authorization.
  SUW – Attack conducted +/- 10 sec of TOT (if given), attack within 2 min of attack authorization TOT not given or if "Immediate TOT" given. Correct procedures utilized, mission objectives met.
- **SACT** – Correct threat identified 100% of event, correct counter procedure initiate for each threat. Maneuvers completed with no corrections by the instructor.
- **SCAR/MAS/CAS – SARC/MAC/JTAC/FAC(A)** – 9-line given correctly with no deviations, alert attacking a/c of applicable threats, request BHA (if applicable).
If in attacking a/c – See Attack/Target engagement/Re-attack guidelines
- **Intercepts** – Maneuvers were completed with no NATOPS limits broken and no corrections by the instructor.
- **SAR Procedures/Swimmer deployment** – Correct search pattern utilized, search pattern flown with no corrections from instructor.
- **Shipboard approaches/procedures** – No corrections required from instructor. Student is within 20' /5 kts or approach profile >80% of approaches.

- **Section procedures** – Appropriate tactics utilized for GUNEX/CATMEX/HELLFIREX as applicable to include mutual support considerations and remote designation assignments and effectiveness. Maneuvers completed without instructor guidance/corrections.
- **AUF/Counter piracy procedures** – SNO granted/ROE utilized, as appropriate. Warning shots/disabling fire requested/granted prior to execution. Procedures effective >80% of the event.
- **CAL/LZ** – Procedures conducted correctly without deviation. SWEEP conducted without assistance.

APPENDIX H PERFORMANCE OUTCOME DATA COLLECTION FORMS

Performance Outcome: ASW TASKS

ASW DETECTION/LOCALIZATION						
Very Poor Superior	0	1	2	3	4	5
5 Contact established including course/speed within 20 min of receiving datum						
4 Contact established including course/speed within 30 min of receiving datum						
3 Contact established including course/speed within 35 min of receiving datum						
2 Contact established including course/speed within 40 min of receiving datum						
1 Contact established including course/speed within 45 min of receiving datum						
0 Contact established including course/speed > 45 min of receiving datum						
Comments:						

ASW TRACKING						
Very Poor Superior	0	1	2	3	4	5
5 Track accurate within 10°/5 kts when compared to known target track						
4 Track accurate within 15°/5 kts when compared to known target track						
3 Track accurate within 15°/10 kts when compared to known target track						
2 Track accurate within 20°/5 kts when compared to known target track						
1 Track accurate within 20°/10 kts when compared to known target track						
0 Track accurate within >20°/10 kts when compared to known target track						
Comments:						

APPENDIX I MODIFIED ATOM DATA COLLECTION FORMS

Information Exchange				Information Exchange Rating
Does not use information from all available sources	0	1-3	>3	0 1 2 3 4 5 5 Uses information from all available sources 0 Consistently does not use information from all available sources
Does not pass information to the appropriate persons before having to be asked	0	1-3	>3	0 1 2 3 4 5 5 Always passes information to the appropriate persons before having to be asked 0 Consistently does not pass information to the appropriate persons before having to be asked
Does not provide situation updates	0	1-3	>3	0 1 2 3 4 5 5 Always provides situation updates 0 Consistently does not provide situation updates
Does not clarify or acknowledge receipt of information`	0	1-3	>3	0 1 2 3 4 5 5 Always clarifies or acknowledges receipt of information 0 Consistently does not clarify or acknowledge receipt of information
Comments:				

Communication Behavior				Communication Behavior Rating
Improper Phraseology	0	1-3	>3	0 1 2 3 4 5 5 Uses proper phraseology 0 Consistently uses improper phraseology
Inaudible Communication	0	1-3	>3	0 1 2 3 4 5 5 Uses clear, audible tone of voice 0 Consistently uses inaudible communication
Excess Chatter	0	1-3	>3	0 1 2 3 4 5 5 Avoids excess/unnecessary chatter (brevity) 0 Consistently engages in excess/unnecessary chatter
Incomplete Reports	0	1-3	>3	0 1 2 3 4 5 5 Provides complete reports 0 Consistently provides incomplete reports
Comments:				

Supporting Behavior				Supporting Behavior Rating
Monitors for and communicates errors	0	1-3	>3	0 1 2 3 4 5 5 Monitors for errors 0 Consistently does not monitor for errors
Corrects errors when discovered	0	1-3	>3	0 1 2 3 4 5 5 Corrects errors when discovered 0 Consistently does not correct errors when discovered
Requests or offers backup or assistance to adjust workload among team members	0	1-3	>3	0 1 2 3 4 5 5 Requests and offers backup or assistance to adjust workload among team members 0 Consistently does not request or offer backup or assistance to adjust workload among team members
Comments:				

Leadership and Initiative				Leadership and Initiative Rating
Offers guidance or suggestions to others	0	1-3	>3	0 1 2 3 4 5 5 Offers guidance or suggestions to others 0 Consistently does not offer guidance or suggestions to others
Shared Leadership: shifts leadership to take advantage of team member strengths				0 1 2 3 4 5 5 Shifts leadership when appropriate 0 Consistently relies on formal/static roles
States clear and appropriate priorities	0	1-3	>3	0 1 2 3 4 5 5 States clear and appropriate priorities 0 Consistently does not state clear and appropriate priorities
Comments:				

APPENDIX J ADAPTED MISSION AWARENESS RATING SCALE (MARS) QUESTIONNAIRE

Name (optional) _____________ Date _____________

Position (check one) ____ HAC _____ H2P _____AW

Air Crew Identifier: _____

Mission Awareness Rating Scale (MARS)

Instructions: Please answer the following questions about the mission you just completed. Your answers to these questions are important in helping us evaluate the effectiveness of this training exercise. Check the response that best applies to your experience.

The first four questions deal with your ability to detect and understand important cues present during the mission.

1. Please rate your ability to **identify** mission-critical cues in this mission.
___ VERY EASY – able to identify all cues
___ FAIRLY EASY – could identify most cues
___ SOMEWHAT DIFFICULT – many cues hard to identify
___ VERY DIFFICULT – had substantial problems identifying most cues

2. How well did you **understand** what was going on during the mission?
___ VERY WELL – fully understood the situation as it unfolded
___ FAIRLY WELL - understood most aspects of the situation
___ SOMEWHAT POORLY – had difficulty understanding much of the situation
___ VERY POORLY – the situation did not make sense to me

3. How well could you **predict** what was about to occur next in the mission?
___ VERY WELL – could predict with accuracy what was about to occur
___ FAIRLY WELL – could make accurate predictions most of the time
___ SOMEWHAT POORLY – misunderstood the situation much of the time
___ VERY POORLY – unable to predict what was about to occur

4. How aware were you of **how to best achieve** your goals during this mission?
___ VERY AWARE – knew how to achieve goals at all times
___ FAIRLY AWARE – knew most of the time how to achieve mission goals
___ SOMEWHAT UNAWARE – was not aware of how to achieve some goals
___ VERY UNAWARE – generally unaware of how to achieve goals

The last four questions ask how difficult it was for you to detect and understand important cues present during the mission.

5. How difficult – in terms of mental effort required - was it for you to **identify** or detect mission-critical cues in the mission?
___ VERY EASY – could identify relevant cues with little effort
___ FAIRLY EASY – could identify relevant cues, but some effort required
___ SOMEWHAT DIFFICULT - some effort was required to identify most cues
___ VERY DIFFICULT – substantial effort required to identify relevant cues

6. How difficult – in terms of mental effort – was it to **understand** what was going on during the mission?
___ VERY EASY – understood what was going on with little effort
___ FAIRLY EASY – understood events with only moderate effort
___ SOMEWHAT DIFFICULT – hard to comprehend some aspects of situation
___ VERY DIFFICULT – hard to understand most or all aspects of situation

7. How difficult – in terms of mental effort – was it to **predict** what was about to happen during the mission?
___ VERY EASY – little or no effort needed
___ FAIRLY EASY – moderate effort required
___ SOMEWHAT DIFFICULT – many projections required substantial effort
___ VERY DIFFICULT – substantial effort required on most or all projections

8. How difficult – in terms of mental effort – was it to decide on **how to best achieve** mission goals during this mission?
___ VERY EASY – little or no effort needed
___ FAIRLY EASY – moderate effort required
___ SOMEWHAT DIFFICULT – substantial effort needed on some decisions
___ VERY DIFFICULT – most or all decisions required substantial effort

APPENDIX K CRITICAL *t* VALUES

df	One-sided	Two-sided
1	6.31	12.71
2	2.92	4.30
3	2.35	3.18
4	2.13	2.78
5	2.02	2.57
6	1.94	2.44
7	1.90	2.36
8	1.86	2.30
9	1.83	2.26
10	1.81	2.22
11	1.80	2.20
12	1.78	2.17
13	1.77	2.16
14	1.76	2.14
15	1.75	2.13
16	1.75	2.12
17	1.74	2.11
18	1.73	2.10
19	1.73	2.09
20	1.73	2.08
21	1.72	2.08
22	1.71	2.07
23	1.71	2.06
24	1.71	2.06
25	1.71	2.06
26	1.71	2.05
27	1.70	2.05
28	1.70	2.04
29	1.70	2.04
30	1.70	2.04

Adapted from van Reeuwijk and Houba (1998)

APPENDIX L CRITICAL *F* VALUES

Bold-face figures: one-sided. Normal-face figures: two-sided.

df_2	df_1 (of numerator)																		
	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	**161**	**200**	**216**	**225**	**230**	**234**	**237**	**239**	**241**	**242**	**244**	**246**	**248**	**249**	**250**	**251**	**252**	**253**	**254**
	648	800	864	900	922	937	948	957	963	969	977	985	993	997	1001	1006	1010	1014	1018
2	**18.5**	**19.0**	**19.2**	**19.3**	**19.3**	**19.3**	**19.4**	**19.4**	**19.4**	**19.4**	**19.4**	**19.4**	**19.4**	**19.4**	**19.5**	**19.5**	**19.5**	**19.5**	**19.5**
	38.5	39.0	39.2	39.3	39.3	39.3	39.4	39.4	39.4	39.4	39.4	39.4	39.5	39.5	39.5	39.5	39.5	39.5	39.5
3	**10.1**	**9.55**	**9.28**	**9.12**	**9.01**	**8.94**	**8.89**	**8.85**	**8.81**	**8.78**	**8.74**	**8.70**	**8.66**	**8.64**	**8.62**	**8.60**	**8.58**	**8.56**	**8.53**
	17.4	16.0	15.1	15.1	14.9	14.7	14.6	14.5	14.5	14.4	14.3	14.3	14.2	14.1	14.1	14.0	14.0	14.0	13.9
4	**7.71**	**6.94**	**6.59**	**6.39**	**6.26**	**6.16**	**6.09**	**6.04**	**6.00**	**5.96**	**5.91**	**5.85**	**5.80**	**5.77**	**5.75**	**5.71**	**5.69**	**5.66**	**5.63**
	12.2	10.7	9.98	9.60	9.36	9.20	9.07	8.98	8.90	8.84	8.57	8.66	8.56	8.51	8.46	8.41	8.36	8.31	8.26
5	**6.61**	**5.79**	**5.41**	**5.19**	**5.05**	**4.95**	**4.88**	**4.82**	**4.78**	**4.74**	**4.68**	**4.62**	**4.56**	**4.53**	**4.50**	**4.46**	**4.43**	**4.40**	**4.36**
	10.0	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.62	6.52	6.43	6.33	6.28	6.23	6.18	6.12	6.07	6.02
6	**5.99**	**5.14**	**4.76**	**4.53**	**4.39**	**4.28**	**4.21**	**4.15**	**4.10**	**4.06**	**4.00**	**3.94**	**3.87**	**3.84**	**3.81**	**3.77**	**3.74**	**3.71**	**3.67**
	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52	5.46	5.37	5.27	5.17	5.12	5.07	5.01	4.96	4.90	4.85
7	**5.59**	**4.74**	**4.35**	**4.12**	**3.97**	**3.87**	**3.79**	**3.73**	**3.68**	**3.63**	**3.57**	**3.50**	**3.44**	**3.41**	**3.38**	**3.34**	**3.31**	**3.28**	**3.23**
	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82	4.76	4.67	5.47	4.47	4.42	4.36	4.31	4.25	4.20	4.14
8	**5.32**	**4.46**	**4.07**	**3.84**	**3.69**	**3.58**	**3.50**	**3.44**	**3.39**	**3.35**	**3.28**	**3.21**	**3.15**	**3.12**	**3.08**	**3.04**	**3.02**	**2.98**	**2.93**
	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36	4.30	4.20	4.10	4.00	3.95	3.89	3.84	3.78	3.73	3.67
9	**5.12**	**4.26**	**3.86**	**3.63**	**3.48**	**3.37**	**3.29**	**3.23**	**3.18**	**3.13**	**3.07**	**3.00**	**2.93**	**2.90**	**2.86**	**2.82**	**2.79**	**2.76**	**2.71**
	7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03	3.96	3.87	3.77	3.67	3.61	3.56	3.51	3.45	3.39	3.33
10	**4.96**	**4.10**	**3.71**	**3.48**	**3.33**	**3.22**	**3.14**	**3.07**	**3.02**	**2.97**	**2.91**	**2.85**	**2.77**	**2.74**	**2.70**	**2.67**	**2.63**	**2.59**	**2.54**
	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78	3.72	3.62	3.53	3.42	3.37	3.31	3.26	3.20	3.14	3.08
11	**4.84**	**3.98**	**3.59**	**3.36**	**3.20**	**3.09**	**3.01**	**2.95**	**2.90**	**2.86**	**2.79**	**2.72**	**2.65**	**2.61**	**2.57**	**2.53**	**2.49**	**2.45**	**2.40**
	6.72	5.26	4.63	4.28	4.04	3.88	3.76	3.66	3.59	3.53	3.43	3.33	3.23	3.17	3.12	3.06	3.00	2.94	2.88
12	**4.75**	**3.88**	**3.49**	**3.26**	**3.11**	**3.00**	**2.92**	**2.85**	**2.80**	**2.76**	**2.69**	**2.62**	**2.54**	**2.50**	**2.46**	**2.42**	**2.38**	**2.34**	**2.30**
	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44	3.37	3.28	3.18	3.07	3.02	2.96	2.91	2.85	2.79	2.72
13	**4.67**	**3.80**	**3.41**	**3.18**	**3.02**	**2.92**	**2.84**	**2.77**	**2.72**	**2.67**	**2.60**	**2.53**	**2.46**	**2.42**	**2.38**	**2.34**	**2.31**	**2.26**	**2.21**
	6.41	4.97	4.35	4.00	3.77	3.60	3.48	3.39	3.31	3.25	3.15	3.05	2.95	2.89	2.84	2.78	2.72	2.66	2.60
14	**4.60**	**3.74**	**3.34**	**3.11**	**2.96**	**2.85**	**2.77**	**2.70**	**2.65**	**2.60**	**2.53**	**2.46**	**2.39**	**2.35**	**2.31**	**2.27**	**2.23**	**2.19**	**2.13**
	6.30	4.86	4.24	3.89	3.66	3.50	3.38	3.29	3.21	3.15	3.05	2.95	2.84	2.79	2.73	2.67	2.61	2.55	2.49
15	**4.54**	**3.68**	**3.29**	**3.06**	**2.90**	**2.79**	**2.70**	**2.64**	**2.59**	**2.54**	**2.48**	**2.41**	**2.33**	**2.29**	**2.25**	**2.21**	**2.17**	**2.12**	**2.07**
	6.20	4.76	4.15	3.80	3.58	3.41	3.29	3.20	3.12	3.06	2.96	2.86	2.76	2.70	2.64	2.58	2.52	2.46	2.40
16	**4.49**	**3.63**	**3.24**	**3.01**	**2.85**	**2.74**	**2.66**	**2.59**	**2.54**	**2.49**	**2.42**	**2.35**	**2.28**	**2.24**	**2.20**	**2.16**	**2.12**	**2.07**	**2.01**
	6.12	4.69	4.08	3.73	3.50	3.34	3.22	3.12	3.05	2.99	2.89	2.79	2.68	2.63	2.57	2.51	2.45	2.38	2.32

Adapted from van Reeuwijk and Houba (1998)

APPENDIX M ACRONYMS

ALFS – Airborne Low Frequency Sonar

ASW – Anti-Submarine Warfare

ATO – Air Tasking Officer

ESM – Electronic Support Measures

FLIR – Forward Looking Infrared

FRAC – Fleet Replacement Aircrew

FRS – Fleet Readiness Squadron

H2P – Helicopter Second Pilot

HAC – Helicopter Aircraft Commander

HARP – Helicopter Advanced Readiness Program

HSM – Helicopter Maritime Strike

ISAR – Inverse Synthetic Aperture Radar

LAMPS III – Light Airborne Multi-Purpose System

MAD – Magnetic Anomaly Detector

MMR – Multi-Mode Radar

NAS – Naval Air Station

OFT – Operational Flight Trainer

SENSO or SO – Sensor Operator

WST – Weapons System Trainer

WTT – Weapons/Tactics Trainer

VITA

for
MICHAEL ALLAN WHITE

EDUCATION

Doctor of Philosophy (Engineering with a Concentration in Modeling and Simulation), Old Dominion University, Norfolk, VA, December 2012

Master of Business Administration/Aviation Embry-Riddle Aeronautical University, Daytona Beach, FL, August 1985

Bachelor of Science, Professional Aeronautics, Embry-Riddle Aeronautical University, Daytona Beach, FL, April 1983

PROFESSIONAL QUALIFICATIONS

Certified Modeling and Simulation Professional

PUBLICATIONS

M. White (2010). Serious Games for UAS Training. In T. Marek, W. Karwowski, & V. Rice (Eds.), *Advances in Understanding Human Performance: Neuroergonomics, Human Factors, and Special Populations* (Chap 38, pp 371-380). Boca Raton, FL, USA: CRC Press, Taylor& Francis Group.

M. White & J. Moss (2009). Lessons Learned from Integration and Transition of Research Prototypes into Joint Interoperable Training Systems. In D. Schmorrow, J. Cohn, & D. Nicholson (Eds.), *The Handbook of Virtual Environment Training: Understanding, Predicting and Implementing Effective Training Solutions for Accelerated and Experiential Learning* (Vol 3, Sec 1, pp 117-124). Aldershot, Hampshire, UK: Ashgate Publishing.

E. Stoudenmire, M. White & K. Roy, (2007), Joint Interaction Validation, *Simulation Interoperability Workshop Fall 2007*, Orlando, FL.

PROFESSIONAL MEMBERSHIPS

National Training and Simulation Association, Association United States Army, Army Aviation Association of America, National Defense Transportation Association

INTERESTS

Modeling, Simulation and Serious Games for Training, Education, Learning & Development

Integration of Live, Virtual and Constructive Simulations with live Command, Control and Communications Systems

Verification and Validation of Modeling, Simulation and Gaming Technologies