2009-06

Stress effects on transfer from virtual environment flight training to stressful flight environments

McClernon, Christopher K.

Monterey, California: Naval Postgraduate School

http://hdl.handle.net/10945/10441

NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

DISSERTATION

STRESS EFFECTS ON TRANSFER FROM VIRTUAL ENVIRONMENT FLIGHT TRAINING TO STRESSFUL FLIGHT ENVIRONMENTS

by

Christopher K. McClernon

June 2009

Dissertation Supervisor: Michael McCauley

This Dissertation was done at the MOVES Institute
Approved for public release; distribution is unlimited

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 0704-0188*
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington DC 20503.			
1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** June 2009	**3. REPORT TYPE AND DATES COVERED** Dissertation	
4. TITLE AND SUBTITLE: Stress Effects on Transfer from Virtual Environment Flight Training to Stressful Flight Environments			**5. FUNDING NUMBERS**
6. AUTHOR(S): McClernon, C. K.			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) The Modeling, Virtual Environment, and Simulation Institute Naval Postgraduate School Monterey, CA 93943-5000			**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A			**10. SPONSORING / MONITORING AGENCY REPORT NUMBER**
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited			**12b. DISTRIBUTION CODE**

13. ABSTRACT (maximum 200 words)

The purpose of this research was to investigate the effects that stress training has on stressful flight operations to mitigate the human factors preconditions to aircraft accidents. In addition, stress training implementation strategies were investigated in order to develop pedagogy pertinent to stress training. A series of three empirical experiments were performed to test the transfer of both human emotional states and task skills from a virtual environment to subsequent test scenarios. Results indicated that stress training improved performance, decreased physiological responses to stress, and decreased subjective appraisals of stress in a simulator criterion session. A second experiment tested the generalization of these results to a novel, real-world stressor. In this study, stress training in a flight simulator was found to enhance performance and moderate the adverse effects of stress when piloting an aircraft in a stressful flight environment. A third empirical study tested the transfer of flight simulator skills to a real-world flying task. Flight simulator training improved the performance of a training group when compared to a no-training, control group. This line of research demonstrates stress training as a viable approach for preparing trainees for stressful flight environments and stress in general.

14. SUBJECT TERMS Stress, training, transfer of training, flight simulator, virtual environment, human physiology, human performance, strain, stress coping, stress exposure training			**15. NUMBER OF PAGES** 227
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UU

NSN 7540-01-280-5500 Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited

STRESS EFFECTS ON TRANSFER FROM VIRTUAL ENVIRONMENT FLIGHT TRAINING TO STRESSFUL FLIGHT ENVIRONMENTS

Christopher K. McClernon
Major, United States Air Force
B.S., Behavioral Sciences—Human Factors, United States Air Force Academy, 1999
M.S., Industrial Engineering—Ergonomics, North Carolina State University, 2003

Submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY IN
MODELING, VIRTUAL ENVIRONMENTS, AND SIMULATION

from the

NAVAL POSTGRADUATE SCHOOL
June 2009

Author: Christopher K. McClernon

Approved by:

Michael E. McCauley
Research Professor, Operations Research and MOVES
Supervisor and Committee Chair

William Becker
Research Professor, MOVES

Anthony Ciavarelli
Research Professor, MOVES

Rudolph Darken
Professor, Computer Science and MOVES

Nita Miller
Associate Professor, Operations Research and MOVES

Paul O'Connor, LCDR, USN
Assistant Professor, Operations Research

Approved by: Mathias Kölsch, Chair, MOVES Academic Committee

Approved by: Doug Moses, Vice Provost for Academic Affairs

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

The purpose of this research was to investigate the effects that stress training has on stressful flight operations to mitigate the human factors preconditions to aircraft accidents. In addition, stress training implementation strategies were investigated in order to develop pedagogy pertinent to stress training. A series of three empirical experiments were performed to test the transfer of both human emotional states and task skills from a virtual environment to subsequent test scenarios. Results indicated that stress training improved performance, decreased physiological responses to stress, and decreased subjective appraisals of stress in a simulator criterion session. A second experiment tested the generalization of these results to a novel, real-world stressor. In this study, stress training in a flight simulator was found to enhance performance and moderate the adverse effects of stress when piloting an aircraft in a stressful flight environment. A third empirical study tested the transfer of flight simulator skills to a real-world flying task. Flight simulator training improved the performance of a training group when compared to a no-training, control group. This line of research demonstrates stress training as a viable approach for preparing trainees for stressful flight environments and stress in general.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGEMENTS

I would like to first and foremost thank my advisor, mentor, and friend, Dr Michael McCauley. His support, insight, and guidance was always invaluable, timely, and comforting in an otherwise difficult process. I will forever be indebted to him.

I would also like to thank all of the members of my dissertation committee during my time at NPS: Drs. William Becker, Anthony Ciavarelli, Rudy Darken, Nita Miller, Paul O'Connor, and Brent Olde. The wealth of knowledge they shared was both immensely broad and deep.

There are many military, federal, and civilian organizations whose support made this project possible. I am grateful for the support from the Monterey Airport Control Tower, the Northern California Terminal Radar Approach Control Facility, the Monterey Navy Flying Club, the Travis Aero Club, and the Avidyne Corporation. In addition, this project was made possible by funding from the Office of Naval Research, the Air Force Air Education and Training Command, and the MOVES Institute. I am also thankful for my extended "family" in the Department of Behavioral Sciences and Leadership at the United States Air Force Academy—their support and sponsorship of this PhD degree is greatly appreciated.

Last, and most importantly, I would like to thank my loving wife, Stephanie, and my son, Jackson. They both are always a great source of support and joy, and they provided a safe haven from dissertation stressors. I am also thankful for the countless hours Stephanie contributed to scheduling all research participants—her contributions made the project professional and successful.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

Anxiety never disappears in a human being in an airplane — it merely remains dormant when there is no cause to arouse it. Our challenge is to keep it forever dormant.

— Harold Harris, Vice President, Pan American World Airways, 1950

A. THESIS STATEMENT

Piloting an aircraft is stressful. Exorbitant costs, highly critical missions, and human lives often depend on one or few pilots. Unlike operators of ground vehicles or sea vessels, aircraft pilots cannot stop in place, exhaustively analyze a situation, and methodically take action. In the event of an emergency situation, pilots must act and they must act immediately. This responsibility and time pressure places an enormous amount of stress on pilots, and it is imperative that pilots take correct and swift actions during these times of high stress.

As long as humans are responsible for piloting aircraft, "the human element will be crucial in the response to emergency and abnormal situations" (Burian, Dismukes, & Barshi, 2003, p. 7). Although moderating stress within the cockpit through the implementation of automation, procedures, technology, and system reliability is one approach to preventing aircraft accidents, another method for alleviating the adverse affects of stress is through training. Stress training has shown promise in a number of different disciplines, and now with the promulgation of virtual environments (VEs) for training, stress training can be tested and implemented in a much more reliable, controlled, and efficient manner.

This thesis addresses stress training as a method for moderating stress and enhancing performance in a stressful flight environment. In addition, the generalization of stress training in a VE to other real-world stressors is addressed. These two thesis statements constitute the empirical experiments in this research project and their supporting literature.

B. PROBLEM STATEMENT

Stress is prevalent in modern military training. When a new military recruit enters basic training, cadres quickly surround the trainee and apply high levels of stress through yelling, physical activity, and mental tests, to name a few. During field exercises for ground troops, ordinance, live rounds, and other simulations of combat's chaos are often implemented. During shipboard firefighting exercises, sounds, smells, and, often, actual fires that mimic the severity of a shipboard fire are used to train firefighters. During the start of each day, military student pilots often undergo "stand-up" exercises where students are chosen at random to recite emergency procedures in the presence of their instructors and classmates. Each of these training scenarios (among many other military examples) are performed for the purpose, at least partly, of introducing stress in training.

The need for stress training is not new. The first two years of World War II saw over 400,000 neuropsychiatric patients admitted to Army hospitals, with a rate of over 100 neuropsychiatric "casualties" per 1,000 troops drafted (Bourne, 1970). These stress-induced illnesses of World War II were not limited to ground troops; psychiatric casualties were common occurrences in fighter and bomber pilots (Hastings, Wright, & Clueck, 1944). As Bourne points out, "much of this attrition might have been avoided had the military been better prepared" (1970, p. 72).

Stress is often introduced in training so that real-world stress is more familiar and easily mitigated. Stress training relies on transfer-of-training research including Osgood's (1949) similar elements theory, indicating that training should share elements with the transfer task in order for the training to be effective, or as the military often refers to as "train how you fight." Therefore, to train for a stressful task, a stressful training environment often is utilized. Stress training also relies on Overton's (1964) state-dependent learning where retention and retrieval are dependent on a person's emotional, physiological, and mental states during both training and recall.

Although stress has been introduced into training since the beginning of a modern military, very little empirical evidence for its effectiveness, implementation, and pedagogy exists. For example, when should stress be introduced during training? How much stress should a trainee be exposed to? How should stress training be paired with skills training for a particular task? The current military training model often includes introducing as much stress as possible, so that a combat experience is mundane by comparison. However, as the United States Marine Corps recently found, inappropriate stress training can result in both physical injury and post-traumatic stress disorder before any combat time is experienced (Army Times, 2007).

Modern advancements in virtual reality (VR) provide a controlled medium for the introduction of stress training. One current application of VEs relevant to stress is the treatment of post-traumatic stress disorder patients (Rizzo, Morie, Williams, Pair, & Buckwalter, 2005). In this type of treatment, patients are presented with highly controlled and regimented levels of stress exposure. These exposure treatments coupled with methodical counseling sessions teach patients how to psychologically mediate the discomfort they have with past experiences. These treatments continue to show promise for such patients. If VEs are proven to be effective for treating patients with stress-related disorders, then they may also show promise for trainees that are preparing to undergo stressful experiences.

One particularly stressful domain that often incorporates VEs in training is aviation. The aviation domain also is inherently unforgiving to errors, attention deficits, and failures in standard procedures—all circumstance that stress can propagate. In Reason's (1990) *Human Error*, a theoretical framework for the aetiology of complex technological system accidents is presented in order to illuminate production system access points and hazard safeguards for accident mediation. This model (see Figure 1.1) defines the antecedents to an accident that begins with organizational deficiencies in the form of decision and management latent failures (see "Organizational Influences" and "Line Management" in Figure 1.1) followed by preconditions to an unsafe act. In order to alleviate unsafe acts, Reason recommends that, first, the preconditions to an unsafe act must be eliminated to the extent possible. In the form of stress, this approach suggests that excessive stress must be designed out of a system. However, Reason's second approach to alleviating accidents is accepting that unsafe acts and system failures will occur, and, therefore, defenses are needed to intervene between an unsafe act and its adverse consequences. One potential safeguard to prevent unsafe acts from propagating into accidents includes training that prepares operators to respond to stressful abnormal situations.

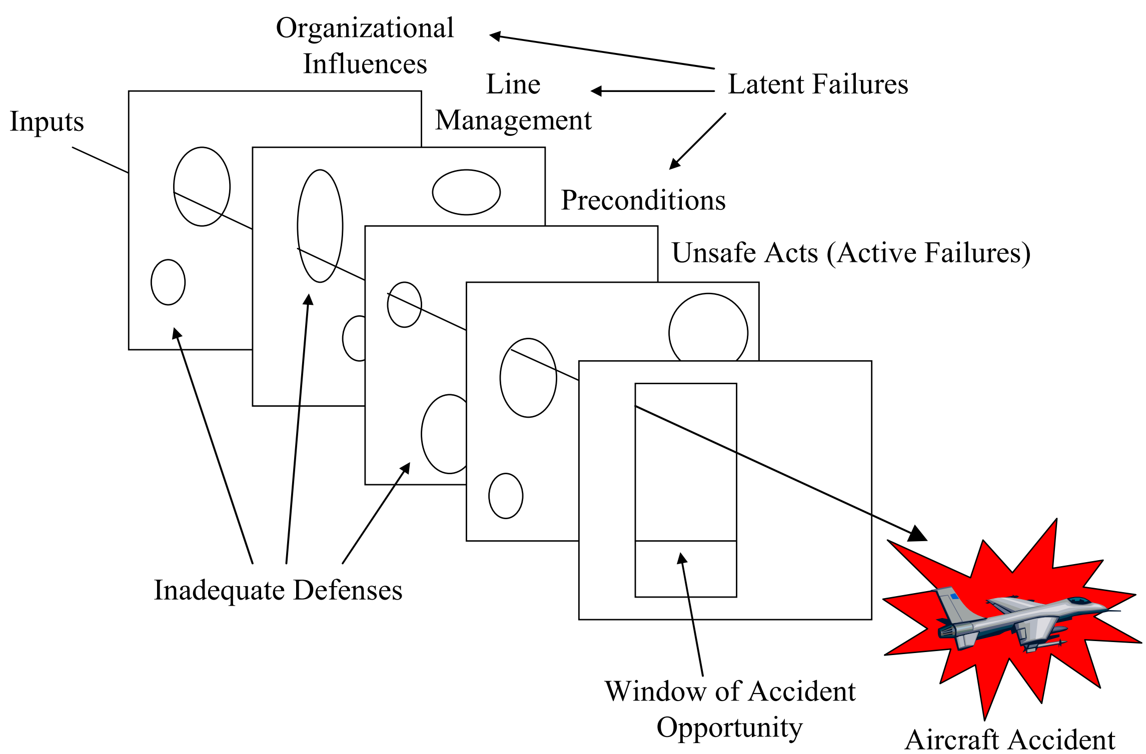


Figure 1.1. Accident causal chain for production systems (adapted from Reason, 1990)

In a new DoD-directed reorganization of aircraft accident analyses, the Human Factors Analysis and Classification System (HFACS) was instituted to help mitigate human factors hazards in the aviation community (see Figure 1.2). This hierarchical approach to aircraft accident investigations expands on Reason's (1990) model by identifying specific Human Factors antecedents to aircraft accidents. Although organizational influences and supervision can certainly impart undue stress on flight operations, the human factors pertaining to accidents (see "Preconditions" in Figure 1.2) are most heavily associated with stress. Environmental factors, personnel factors, and conditions of individuals are all subject to stress and often are preconditions to aircraft accidents (Gibb & Olson, 2005).

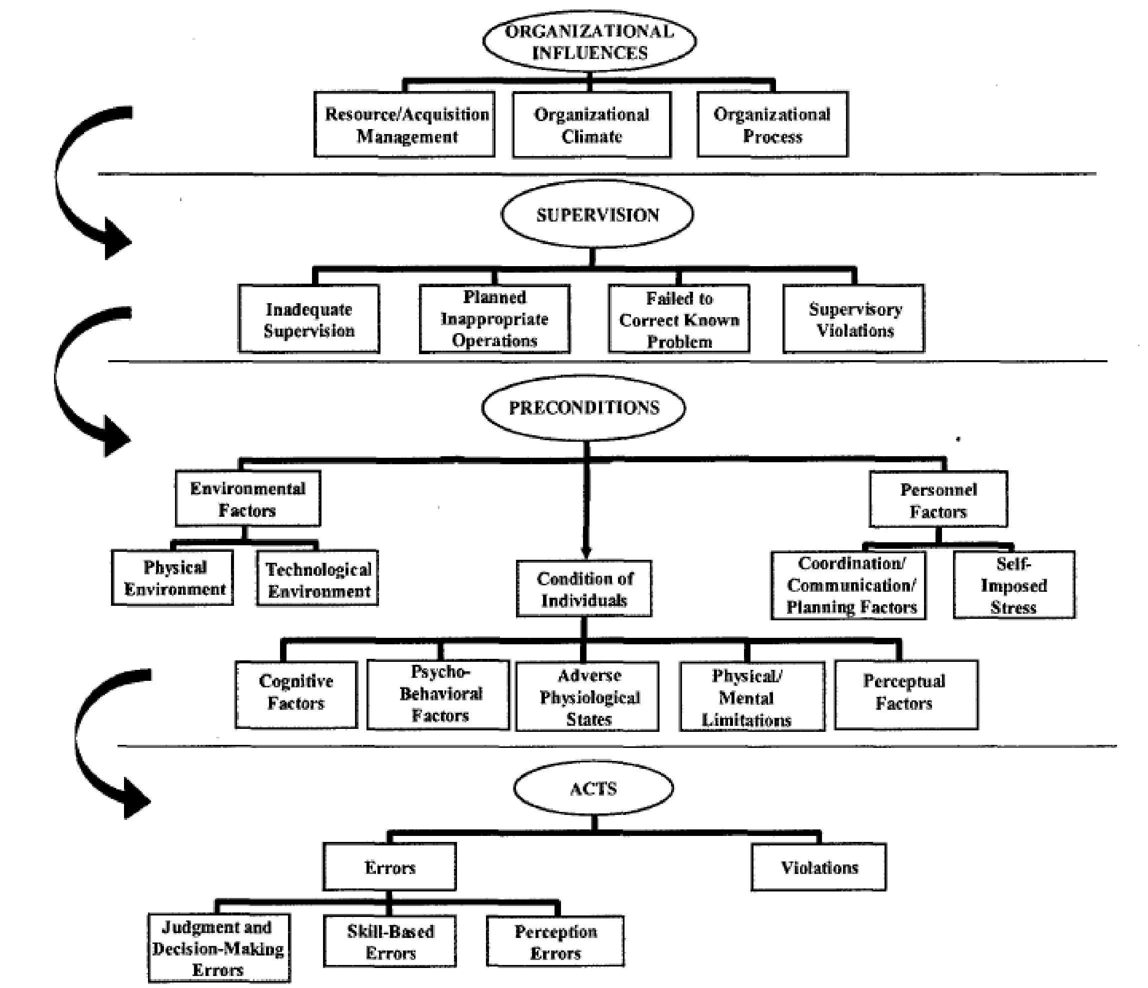


Figure 1.2 DoD HFACS (In Gibb & Olson, 2005)

The human ability to perform under overly stressful situations is often hindered. However, "the aviation industry lacks substantive human performance guidelines for designing, validating, certifying, and training procedures for emergency and abnormal situations" where stress is most prevalent (Burian, Dismukes, & Barshi, 2003, p. 1). In an attempt to address these concerns, the NASA emergency and abnormal situations project has set an overriding goal to "develop guidance for procedure and checklist development and certification, training, crew coordination, and situation management based on knowledge of the operational environment, human performance limitations and capabilities, and cognitive vulnerabilities in real-world emergency and abnormal situations" (Burian, Dismukes, & Barshi, 2003, p. 1). Their work suggest that stress and time pressure has undeniable effects on attention, memory retrieval, and problem solving, although the specific nature of these effects in the aviation domain is still under investigation.

Stress training is one approach to optimizing performance in adverse situations. Empirical evidence presented in Chapter II indicates that stress training is both effective and efficient when properly paired with task training (Driskell, Johnston, & Salas, 2001; Friedland & Keinan, 1992; Johnston & Cannon-Bowers, 1996; Keinan, Friedland, & Sarig-Naor, 1990; Meichenbaum, 1993). Other research presented indicates the potential for stress training to generalize to novel tasks and novel stressors (Driskell, Copper, & Moran, 1994; Driskell, Johnston, & Salas, 2001; Saunders, Driskell, Hall, & Salas, 1996; Thorndike & Woodworth, 1901).

The present research project expands on these findings to further empirically define stress training's efficacy. In addition, this research project examines the efficacy of VEs for introducing stress training and imparting flying skills. This project tests the following three research questions:

(1) Does stress training in VE transfer to a stressful simulator criterion session?

(2) Does VE training generalize to a novel task and a novel stressor in the real physical world?

(3) Do task skills learned in a VE transfer to a real world task?

These three research questions were tested (1) in an aviation context, (2) using relatively long-term acute stressors, (3) using an operational task, (4) using operational performance measures, (5) including expert performance evaluations, (6) measuring human physiological stress responses, (7) evaluating subjective appraisals of stress, and (8) presenting training in a low fidelity VE.

C. CONTRIBUTIONS

One potential general finding from these experiments is the gained knowledge and emphasis on training pedagogy. Most researchers and learning theorists agree that the learning pedagogy, training development and implementation, and accomplishment of learning objectives are of utmost importance when considering a training system's efficacy (Caird, 1996; Clark, 2001). For many years, researchers and engineers have attempted to enhance training by incorporating technology and, more specifically, increasing simulator fidelity. Many current military and commercial flight simulators have convincing graphics, physics, and motion platforms with little empirical evidence for these technologies' training efficacy. The National Research Counsel (Druckman & Swets, ed., 1988) states that "effective instruction is the result of such factors as the quality of instruction, practice of study time, motivation of the learner, and the matching of the training regiment to the job demand." This thesis addresses some of these components of effective instruction.

This research also provides insight into stress in general. Very little is known regarding the effects that stressors have on human capabilities, moderation of stress, causes of stress, etc. The generalization portion of the proposed research (see Chapter IV) sheds light onto the cognitive constituents of stress. This research supports the notion that stress has structural characteristics separate from the surface characteristics (i.e., domain, individual differences, etc.) for which they interact. More concerning this topic can be found in Chapter II.

Most importantly, this research introduces new approaches to preparing combat troops for the inherently stressful environments they will surely face. Machines are immune to emotional stressors, but as long as humans are asked to participate in combat, the human entity will be a primary factor contributing to combat outcomes. If combatants can focus all of their cognitive abilities to the tasks at hand while simultaneously imposing stress on the adversary, then training can impart an advantage to our troops before they even step on the battlefield.

1. The "Sam-to-Sully" Paradigm

On 15 January 2009, Captain Chesley "Sully" Sullenberger flawlessly piloted an Airbus 320, US Airways Flight 1549, carrying 155 people to a water ditching in the Hudson River. The aircraft "Sully" was piloting struck a flock of large birds shortly after departure from New York's LaGuardia airport, resulting in the first ever commercial airliner water ditching with no fatalities. His calm and controlling demeanor in the face of a highly stressful flight environment presents a case study for how to prepare someone for a stressful situation. First, a review is needed of what skills, attributes, and training led to such a momentous accomplishment. The laborious path between a young, novice (called “Sam”) to a highly skilled practitioner ("Sully") is then investigated and future considerations for streamlining this path are considered.

Captain Sullenberger was possibly the most qualified person on the planet to pilot an Airbus 320 without engine power. In an interview with Katie Couric, Captain Sullenberger stated, "in many ways as it turned out my entire life up to that moment had been a preparation for that moment" (Sullenberger, 2009, February 5). He is a graduate of the United States Air Force Academy where he was a glider instructor and received the coveted Outstanding Cadet in Airmanship Award. Following the Academy he excelled in pilot training, graduating first in his class and becoming an F-4 Phantom pilot. After the Air Force, Captain Sullenberger was hired as a pilot for US Airways where he remained for 29 years. In total, Captain Sullenberger has accumulated over 19,000 hours and over 40 years of flying experience.

He also has contributed significantly to the field of Aviation Safety as a safety chairman for the Air Line Pilots Association and through participation on numerous Air Force and FAA aircraft accident investigations. Captain Sullenberger was instrumental in the creation of the FAA Advisory Circular, developed the cockpit resource management course at U.S. Airways and has co-authored NASA technical papers on error.

There is no dispute that Captain Sullenberger was well prepared for the situation encountered on 15 January 2009—in fact, there is little else the passengers of that flight could have requested in a pilot that day. Captain Sullenberger achieved a level of expertise throughout his career that can be described as "experienced-based expertise."

In contrast, "Sam" is a fictitious Senior in High School. Sam has never flown an aircraft, but is capable, willing, and eager to learn. As an aviation community, the best way to prepare Sam for a stressful situation like a water ditching of an Airbus 320 with 155 people on-board is through experience-based expertise. However, experience-based expertise is incredibly inefficient and impracticable. It is costly and logistically impossible to train all pilots to this level of proficiency. Therefore, the Sam-to-Sully paradigm encourages us to explore other alternatives to experience-based expertise. This research project considers the feasibility of compartmentalizing the skill and emotional constituents of a stressful situation, followed by carefully and methodically merging these two elements in a total training system. We explore the efficacy of training pedagogy that considers these constituents to more efficiently and effectively prepare "Sam" for the stressful situation in which Captain "Sully" found himself.

THIS PAGE INTENTIONALLY LEFT BLANK

II. BACKGROUND AND RELATED WORK

The alleviation of human error, whether design or intrinsically human, continues to be the most important problem facing aerospace safety.

— Jerome Lederer

Training in VEs is not a new concept. The military has trained pilots in VEs for almost eighty years. Surprisingly, VE training pedagogy is still misunderstood and the current VE training practices are mostly warranted by rule-of-thumb and status quo. The primary reason for a dearth in VE and transfer research in general is the required cost and resources to perform such studies. However, VEs have the potential to provide extensive long-term cost savings, more efficient training, and safer training, among many other advantages. The empirical experiments presented in this thesis test the feasibility of VEs for introducing stress into training. Therefore, in the first section of this chapter (Section A), VEs are defined and further advantages for using VEs in training are provided. Given the aviation context of this thesis, a brief history of the use of VEs in flight training is then provided. These sections provide the support necessary to warrant VE flight training research and suggest benefits of such research.

The experiments conducted in this thesis explore the transfer of training from a VE to a stressful transfer task. Transfer of training is addressed in Section B of this chapter beginning with the definition and measurement of transfer of training. Although training transfer is critical to the development and testing of any training system, such research is not always performed; the following sub-sections address why transfer research is scarce and suggests alternatives to transfer research for determining training efficacy. The final subsection, "VE Factors Affecting Transfer of Training," assimilates Sections A and B to determine what features and components of VEs are pivotal to proper transfer of training research. This final sub-section will shed light on the areas of VE transfer of training that are lacking empirical evidence, one of which is stress training in VEs.

The experiments presented in Chapters III and IV explore the efficacy and feasibility of incorporating stress in training. Therefore, in an attempt to better understand the human responses to stress, Section C first defines stress. A historical review of stress examines early pioneering work in stress up to and including recent stress discoveries. Section D then explores various approaches to preparing trainees for stressful situations by introducing stress into training. Finally, similar to the previous section, a model of stress training in VEs is presented and an analysis of factors that can affect stress within VEs is conducted.

A. TRAINING IN VIRTUAL ENVIRONMENTS

Virtual reality is defined as "a high-end user-computer interface that involves real-time simulation and interactions through multiple sensorial channels" (Burdea & Coiffet, 2003). The synthetic environment in which VR takes place is referred to as a *virtual environment*, or an organization of sensory information that leads to perceptions of a synthetic environment and its contents as being non-synthetic (Bailenson, Yee, Blascovich, Lundblad, & Jin, 2007; Blascovich, 2002). Blascovich (2002) further defines an Immersive Virtual Environment (IVE) as a VE that is arranged in such as a way as to "create a psychological state in which the individual perceives himself or herself as existing within the VE" (p. 129). An example of an IVE is a flight simulator that enables a trainee to feel as if they are actually flying an aircraft (independent of how accurately it represents an aircraft). The perception of immersion can be experienced in a full-motion, highly realistic flight simulator, or an IVE can also occur in mental rehearsal without any software or hardware implementation at all. These two examples of immersion points out the reliance of immersion within a VE on the psychological state of the user, as opposed to the training system itself.

There are two defining features unique to VEs. First, a VE can be explored in real time, similar to real-world exploration. Real-time exploration requires that as control inputs are made, the virtual system quickly translates inputs and the environment reacts appropriately. The second characteristic of VEs is that a user may interact with objects and events in the simulation (Cobb & Fraser, 2005). These objects may be other people and objects, or it may simply be navigation within the environment. According to these two features, a cinema experience is not a VE as the user passively watches scripted events unfold. However, Blascovich (2002) may disagree with this assertion, again placing emphasis on the user's psychological state—if a participant feels immersed, then that alone qualifies any experience (e.g., books, dreams, movies, etc.) as an IVE experience. Blascovich's definition further identifies the existence of a VE not within the technology, but rather the psyche of an individual.

Presence is another concept that is relevant to VR. Presence is described as the "[psychological] sense of being in the virtual, as opposed to the real environment" (Meehan, 2002, p. 645). Presence is the primary measure for development of VR systems. For example, a new, technologically advanced, and expensive VR training system has little value if it provides little presence to the user. Blascovich (2002) defines social presence as a "psychological state in which the individual perceives himself or herself as existing within an interpersonal environment" (p. 130), further removing emphasis from technology to the individual user. Despite the apparent importance of presence on VR design, there is still controversy surrounding the definition and measurement of presence. Nash, Edwards, Thompson, and Barfield (2000) provide a review of both subjective and objective measures used to evaluate presence within VEs.

Another important component of VEs is system fidelity, or the "extent to which the VE and interactions with it are indistinguishable from the participant's observations with a real environment" (Waller, Hunt, & Knapp, 1998, p. 130). Fidelity can be parsed into three different types: equipment, environment, and psychological. Equipment fidelity is the extent to which a training system physically represents a real world system. In the case of flight simulators, equipment fidelity is how realistic the physical cockpit, controls, and displays represent a real aircraft. Environment fidelity refers to the situational cues a trainee receives from a training system. In a flight simulator, the realism of external air traffic control communications in addition to other aircraft, weather, flight physics, and specific training scenarios can all be considered environmental fidelity. Finally, psychological fidelity is the cognitive appraisal of a training system's realism, and is theoretically independent of the equipment and environmental fidelity. For example, a static aircraft mock-up may be very low in equipment and environmental fidelity; however, if a trainee considers it very realistic and/or past experience allows missing information to be cognitively "filled in," then psychological fidelity can still be relatively high.

Fidelity does not have to refer to an entire training system. In fact, designing virtual training systems that are nearly identical to the real world, especially in the case of complex systems, can over saturate a trainee, quickly overloading human information-processing capabilities. In these situations, part-task training can be utilized to train specific elements of a system separate from other components. Only once the subcomponents are learned and practiced to a certain level of proficiency, they are combined. Other approaches increase total system fidelity as a trainee progresses in order to reduce overall complexity (Regian, Shebilske, & Monk, 1992). Simplification, guiding, and adaptive training are other examples of altering fidelity to enhance training efficacy (Wickens, Lee, Liu, & Becker, 2004).

1. Advantages and Applications of Virtual Environments

Virtual Environments were chosen as a training medium in this thesis due to the number of advantages VEs have over traditional training systems. The following section explores some of these advantages and applications of VEs.

Virtual environments are particularly attractive for real world situations that are impractical because they are inherently dangerous (e.g., firefighting, surgery, etc.), too remote (e.g., planetary exploration/training, remote military regions, etc.), too small to enter (e.g., atomic structures, molecular studies, etc.), logistically difficult (e.g., network centric warfare, collaborative operations, etc.), unduly expensive (e.g., flight training, space operations, etc.), too difficult to control (e.g., social research) or non-existent (e.g., future systems; Caird, 1996; Péruch, Belingard, & Thinus-Blanc, 2000; Rose, Attree, Brooks, Parslow, Penn, & Ambihaipahan, 2000; Witmer & Singer, 1998). In addition, VR is particularly attractive for training, because VR preserves "(a) visual-spatial characteristics of the simulated world, and (b) the linkage between motor actions of the student and resulting effects in the simulated world" (Regian, Shebilske, & Monk, 1992).

Related to learning, Saba Software, Inc. (2001) provides a review of e-Learning for training within a management domain. They define e-Learning as "the creation, delivery, and management of learning using Internet tools and technology" (p. 2) including simulations and VEs. They provide a number of advantages for implementing e-Learning including: e-Learning can take place anytime, anywhere allowing quick, convenient training especially at remote locations; learners can choose the mode that they prefer; e-Learning can provide monetary advantages over often costly classroom time; media content can be recycled cheaply with very low monetary and time costs; and training can be dispersed and used immediately.

There is some disagreement about the effects that VEs, and media training in general, has on trainee motivation. Some research indicates that media may not motivate users, and that "enhancing" training by incorporating technology can actually overload working memory (Clark, 2001). However, VEs may prove beneficial in especially mundane tasks, where three-dimensional interaction may engage a trainee. Salzman, Dede, and Bowen-Loftin (1999) found that IVE technology can be more motivating than traditional, two-dimensional training by utilizing multisensory cues to direct trainee attention. Training media can also be applied dynamically using VEs for tasks like air traffic control, where controllers experience periods of overload/stress followed by periods of monotony.

Wickens et al. (2004) provide many further advantages for implementing technology in training. Traditional training systems using computer-based training, lecture-style presentation of information, or low fidelity simulations may only utilize one or two sensory modes, while IVEs can incorporate various mediums that can exploit the multiple resource theory by incorporating parallel processing into training. Training technologies can also provide (1) immediate, timely feedback, (2) active problem solving support, (3) adaptive training by utilizing scaffolding, and (4) individually tailored training (e.g., intelligent tutoring systems).

Finally, VEs are particularly attractive to researchers because virtual experimental manipulations are relatively easy and experiments are more easily controlled, all but eliminating confounding variables (Hoyt, Blascovich, & Swinth, 2003, p. 184; Péruch, Belingard, & Thinus-Blanc, 2000). For example, social psychologists are able to test psychological constructs that are nearly impossible to test in the real world, providing a medium for presenting participants with many various experimental illusions (Loomis, Blascovich, & Beal, 1999). Similarly, VEs provide the "ability to implement contexts and relationships not possible to achieve in a traditional learning setting" (Bailenson, Yee, Blascovich, Lundblad, & Jin, 2007). However, this advantage comes at a cost, as VEs rarely perfectly represent the physical world that they represent. As VE technology improves, researchers are able to present participants with various modalities of precise information, alleviating any shortcomings in immersion. Another solution to the fidelity problem is to improve psychological fidelity using creative experimental methodologies or isolating sensory channels of interest. For example, VEs allow the continuous and precise tracking of participant location and behavior that is available for data collection.

Given the current technological training capabilities, it is important to reemphasize the importance of training pedagogy first and foremost before the implementation of media and technology. A sound training system construction, design, test, and evaluation is most important. However, there are some tasks that are best suited for the implementation of media, making the implementation of virtual training systems task dependent. Following are some examples of tasks that are best suited for the application of VEs.

— In an inherently difficult task or long training program, embodied agents that teach and learn can be used adaptively to optimize a training system. This technique allows a learner to progress through an adaptive training system at his/her own pace. This can ultimately save time in the long run, and prevent negative transfer from students who learn at a slower pace.

— In learning tasks where students learn together, some trainees are motivated by the performance of others. New research is showing that virtual co-learners can be used to enhance training (Bailenson, Yee, Blascovich, Lundblad, & Jin, 2007). These agents' behaviors can be manipulated to improve a learner's performance.

— Some tasks are very dynamic and spatial in nature. For example, flight simulations, architecture/engineering, and navigation tasks require an immense amount of spatial learning. Media and the use of visualizations can enhance training for these tasks.

— Tasks that require real-time data analysis, and advanced/sophisticated data analysis techniques may be better served by media training due to their ease and capability of complex data analysis.

— Tasks that are dangerous and/or expensive are better served by learning from media.

— Tasks that require some level of social interaction may be better served by media, especially if they are remotely located away from other learners. An example of this is language or cultural education for military members.

Despite these advantages and possible applications for VEs in training, Péruch et al. (2000) point out that there are also several drawbacks to the use of VEs including a "lack of realistic environmental modeling, slow image generation and rendering, narrow field of view, optical distortions, poor spatial resolution, etc." (p. 255). As most of these drawbacks are technology-dependent, they are quickly being overcome by advancements in VE technology.

Caird (1996) also stresses that VEs are only one of many ways to train people, and it is important to point out potential disadvantages of VEs. First, VEs can be relatively expensive in the short-term. A new VE built from the ground-up can consume exorbitant costs. Building on existing technologies (e.g., the use of games for training), collaborative efforts, and life-cycle planning are alleviating some of these costs. Immersive VEs are also oftentimes unfamiliar to the general public, requiring basic VE training before system training can commence. Continued VE proliferation is exposing the public to innovative training technologies making VEs more commonplace in the near future. In the past, VE systems were notoriously unreliable, but as the field develops these shortcomings are being overcome, especially in terms of system interoperability and networking. Given the before mentioned importance on fidelity, VE systems' display and interaction limitations are still far from the physical world. As technologies and implementation methodologies improve, physical and psychological fidelity likewise are improving, bridging the gap between the real world and VEs.

This section provides many advantages and applications for training in VEs. Given the benefits of training in VEs, the research project in this thesis will primarily train participants in a low-fidelity flight simulator. However, there is little empirical basis for VE implementation in aviation. Given the aviation relevance of this thesis, VEs as a means for introducing flight training is reviewed in the following section.

2. Virtual Environments for Flight Training

The use of VEs for flight training is not a new concept. The first flight simulators were introduced in the 1920s and were relatively rudimentary with no visual displays. The Link trainer is most often synonymous with the birth of flight simulators, and it featured a motion platform and working instruments to train students on instrument procedures. Following the introduction of analog computers, the aircraft flight models, flight instruments, and simulators themselves were all driven by computers. Limitations in simulator visual displays were quickly overcome with digital-processing, and simulators incorporated increasing levels of visual fidelity. The advent of sophisticated simulator motion platforms were first introduced to train student pilots to utilize their vestibular senses in instrument conditions. However, shortly after the vestibular system shortcomings were realized, motion platforms were utilized to introduce conflicting motion and visual cues to train non-reliance on the vestibular system. A current movement is flight training is concerned with the use of PC-based flight simulators for flight training. Currently all U.S. military flight training incorporates simulator training into its syllabus to some extent (Koonce & Bramble, 1998).

The FAA currently defines a flight simulator as a device that—

(i) is a full-size aircraft replica;

(ii) includes the hardware and software necessary to represent the aircraft in ground operations and flight operations;

(iii) uses a force cueing system that provides cues at least equivalent to those cues provided by a 3 degree freedom of motion system;

(iv) uses a visual system that provides at least a 45 degree horizontal field of view and a 30 degree vertical field of view simultaneously for each pilot; and

(v) has been evaluated, qualified, and approved by the Administrator. (ASA, 2008)

This definition places a large importance on simulator fidelity, and is mostly unsubstantiated by empirical research. Both of these issues will be discussed in Section B below.

Flight simulators are unquestionably beneficial for training flight skill and enhancing flight safety (McCauley, 2006). Flight simulators are also an attractive alternative to aircraft flight time due to their significantly lower purchase price, lower operation cost, and increased safety (Salas, Bowers, & Rhodenizer, 1998). Smode, Hall, and Meyer (1966) reviewed early flight simulator research to determine the efficacy and proper implementation of simulators. They concluded that the effectiveness of simulator training was "as much a function of the way in which simulator is used for training as it is a function of degree and fidelity of simulation" (p. 146). In a more recent literature review of 247 flight simulator studies, Hays, Jacobs, Prince, and Salas (1992) found that 90% of their reviewed studies favored flight training combined with simulator training to flight training alone. Further reviews of research pertaining to the transfer of skills and knowledge from flight simulators to aircraft is presented in the following section.

B. TRANSFER OF TRAINING

One of the most crucial yet challenging aspects of training development is the evaluation of training efficacy. Training evaluation is the process of determining the extent to which knowledge, skills, and abilities developed in a training task are utilized in their intended environment, or *transfer of training*. Transfer of training has been empirically studied for over 100 years by researchers in various fields including engineering psychology, human factors, learning theory, and management, to name a few. As Salas, Bowers, and Rhodenizer (1998) point out, due to the interdisciplinary nature of this line of research, an open dialog is required between these various disciplines. The following review examines research from each of these disciplines to, first, further define transfer of training. Secondly, common transfer or training measurement techniques are presented. Lastly, factors that affect training transfer are presented. The background presented in the last section is essential for determining the skills that may be transferred from a VE flight simulator to a stressful transfer task.

Many researchers from various disciplines have attempted to define transfer of training. Following are definitions from a variety of different disciplines:

> Does it train as well as, or perhaps, better than some other way we could use it? (Orlansky, String, & Chatelier, 1979)
>
> Degree to which learning one task is facilitated or hindered by prior or interpolated learning of another. (Roscoe & Williges, 1980)
>
> The degree to which trainees effectively apply knowledge, skills, and attitudes gained in a training context to the job (Baldwin & Ford, 1988) and subsequently maintain them over a period of time (Cheng & Ho, 1999)
>
> When learning in one context or with one set of materials impacts on performance in another context or with other related materials. (Perkins & Salomon, 1992)
>
> The effect that practice on one task will have on the performance of another task. (Abel & Merryman, 1987)
>
> Whether or not training carried out in a training environment transfers to the equivalent real world situation (Rose, Attree, Brooks, Parslow, Penn, & Ambihaipahan, 1998)
>
> The extent to which learning a new skill, or a skill in a new environment, capitalizes on what has been learned before. (Wickens & Hollands, 2000)
>
> How skills and knowledge acquired in one domain transfer to another. (Wickens & Hollands, 2000)

Although some definitions generalize to imply an advantageous transfer of skills, transfer of training can result in a decrease in performance if negative transfer occurs. As Osgood (1949) points out, when two tasks share the same stimuli or responses, transfer can occur (Roscoe, 1971); however, when two tasks share the same stimulus elements, but require different response mapping, the similarity paradox occurs (Osgood, 1949), and transfer will be negative. The similarity paradox is especially prevalent if there is incompatibility between the new and old responses (Annett, 1998; Wickens & Hollands, 2000). Negative transfer is a very real concern in the case of military training. For example, a motion simulator for flight training that gives a very real, yet slightly inaccurate representation of motion may result in an increase in simulator performance (i.e., both training and transfer tasks exist within the simulator, therefore they are identical and invariably share stimuli and responses). However, when the trainee is transitioned to a real aircraft that behaves in a slightly different manner than the simulator, the stimuli and responses now differ resulting in negative transfer. Training fidelity is an apparent remedy for negative transfer, but as shown in the below literature review (see "VE Factors Effecting Transfer of Training"), fidelity may not always be the solution and other safeguards may need to be employed.

1. Transfer of Training Measurement

"The ultimate purpose of asking the training effectiveness question is to improve training" (Blaiwes, Puig, & Regan, 2001), and as Roscoe and Williges (1980) and Wickens and Hollands (2000) point out, training effectiveness can be quantitatively measured using three different methods: percent transfer, transfer effectiveness ratio, and training cost ratio. These three methods are instrumental for determining stress training efficacy in this thesis. Therefore, each of these methods will be described here.

Percent Transfer

The first measurement technique, percent transfer, simply tells how much information is learned in a training system when compared to a group that receives no training (see Figure 2.1). The computation is:

$$percent\ of\ transfer = \frac{Y_o - Y_x}{Y_o}\ x\ 100 = \frac{savings}{control\ time}\ x\ 100 \qquad \text{(Equation 1)}$$

where:

Y_o = time, trails, or errors required by a control group to reach a performance criterion after zero training units on a prior or interpolated task (i.e., "control time," Wickens & Hollands, 2000; see "Control Group" in Figure 2.1);

Y_x = corresponding value for an experiment transfer group having received X training units (i.e., "transfer time," Wickens & Hollands, 2000; see "Training Group" in Figure 2.1).

(Roscoe & Williges, 1980)

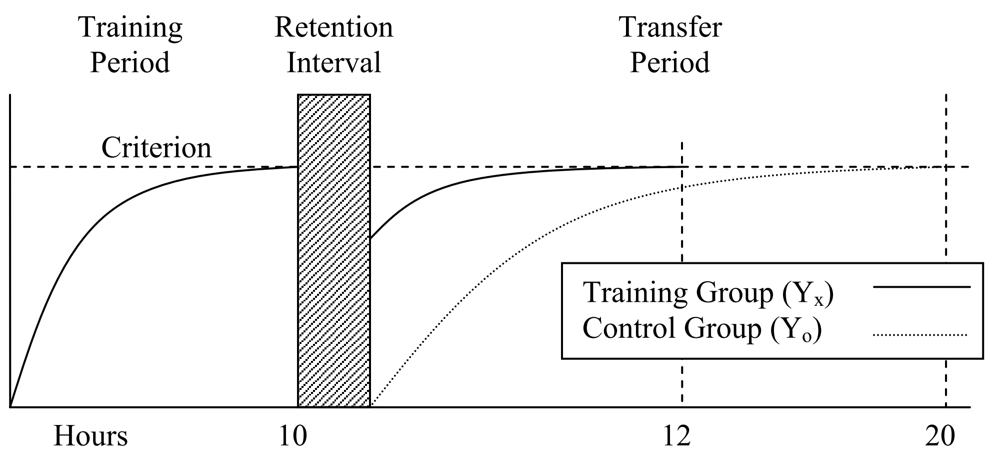


Figure 2.1 Transfer of training measurement

For example, consider the following flight training context: a group of novice students receives a predetermined and controlled amount of simulation training (say 10 hours), and then following a retention interval their time to reach a performance criterion (like hours to solo in an aircraft) is measured (e.g., Y_x = 12 hours). Another group of novice students receives no simulation training time, and their time to solo also is measured (e.g., Y_o = 20 hours). The percentage of transfer, or *savings*, for this example is:

$$\textit{percent transfer} = \frac{20 - 12}{20} \; x \; 100 = 40\%$$

In other words, 40% of the training that the simulation group received was transferred and utilized in the real world, flying task; or those students that participated in simulation training were able to solo in 40% the time of those with no training. It is important to point out that percent transfer can also express negative transfer if a training system results in a decrease in transfer performance. For example, if the flight simulator in the above example resulted in a mean solo time of 28 hours (Y_x = 28 hours), then the percent transfer would be -40%.

However, as Roscoe and Williges (1980) note, percentage of transfer calculations provide no insight into the amount of practice on the training task (see "Training Period" in Figure 2.1) and, therefore, provide no conclusions about transfer effectiveness and training efficiency. Looking at the example above, if the same results were observed after students were required to invest an exorbitant amount of training time (e.g., 40 hours of simulation time), a positive percentage of transfer of 40% would still be observed.

Transfer Effectiveness Ratio

The transfer effectiveness ratio (TER) takes into consideration the amount of time spent in training, and is expressed in the following equation:

$$TER = \frac{Y_o - Y_x}{X} = \frac{amount\ of\ savings}{transfer\ group\ time\ in\ training\ program} \quad \text{(Equation 2)}$$

where:

Y_o = see Equation 1 above;

Y_x = see Equation 1 above;

X = time, trials, or errors by an experimental transfer group during prior or interpolated practice on another task (i.e., "time in training," Wickens & Hollands, 2000).

(Roscoe & Williges, 1980)

Given the flight training example above, if the transfer group spends 10 hours in the simulator training,

$$TER = \frac{20 - 12}{10} = 0.8$$

In other words, for each hour spent in training, .8 hours are saved in transfer. Comparing this to the hypothetical situation where trainees spend 40 hours in the simulator,

$$TER = \frac{20 - 12}{40} = 0.2$$

A clear reduction in transfer effectiveness is observed. Also notice that if the time in training is equal to or less than the amount of savings (e.g., $X \leq 8$ in the above example), then $TER \geq 1$. If the TER is less than 1, the training system is less efficient than no training at all. Although Wickens and Hollands (2000) note, a TER of less than 1 may still be advantageous if (1) the training system is safer, and (2) the training system is cheaper.

Training Cost Ratio

"The effectiveness of training cannot be addressed meaningfully without also considering its cost" (Orlansky, String, & Chatelier, 1979). The training cost ratio (TCR) is used to determine the cost effectiveness of a training system, and is calculated as

$$TCR = \frac{training\ cost\ in\ target\ environment\ (per\ unit\ time)}{training\ cost\ in\ the\ training\ program\ (per\ unit\ time)} \quad \text{(Equation 3)}$$

(Wickens & Hollands, 2000)

The cost effectiveness of a training system is then determined by multiplying TER by TCR. A program is cost effective if TER x TCR > 1 (Wickens & Hollands, 2000). However, even if a training program is considered cost ineffective there may still be advantages of implementing such a program, as is shown previously regarding the advantages of virtual training systems.

Up until now, transfer of training has been discussed in general terms, using "training systems" to describe training in the real world (e.g., classroom training), simulation training (e.g., flight simulators), computer based training, etc. However, as mentioned previously, a recent trend in training system development is the utilization of VEs to provide training. Even given recent applications of this technology to training, very little is understood concerning the efficacy of virtual training systems measured in terms of transfer of training. The following section will address this current research deficit.

2. Barriers to Transfer of Training

There is an empirical vacuum in the literature on transfer of training. One motivation for conducting the empirical research in this thesis is to contribute to an empirical foundation for transfer of training. Despite the widespread use of VEs for training purposes, Caird (1996) points out that this empirical vacuum is "exacerbated by studies that have used poor experimental designs, small sample sizes, or suffer from multiple confounding variables" (p. 127). In addition, "there have been relatively few attempts to investigate empirically the virtual to real transfer process in terms of what sort of training shows transfer, in what conditions, to what extent, and how robust the transferred training proves to be" (Rose, Attree, Brooks, Parslow, Penn, & Ambihaipahan, 1998, p. 69). Some researchers have provided potential reasons for the lack of transfer or training research (Blaiwes, Puig, & Regan, 2001; Caird, 1996; Darken, 2009), and a summary of these reasons follows.

—Research has shown that training fidelity can enhance transfer of training. Increasing training fidelity is typically limited by technology and system cost, and surplus system costs are typically spent on operational systems as opposed to their training systems. Therefore, the fidelity required to enhance transfer of training and, thus, conduct a transfer of training study is typically not available.

—Although training fidelity can enhance transfer of training, no research has conclusively shown that the medium of a training program impacts learning outcomes. The lack of empirical evidence linking medium with training efficacy has decreased the motivation for performing future transfer of training studies. Until novel alternatives to fidelity are researched for enhancing transfer of training, future transfer of training studies may be infeasible.

—Although physical fidelity is directly observable, evaluated, and analyzed, psychological fidelity is not. A system can look and feel exactly like the system it is meant to train, but if the user does not perceive it to be the same, then there can be a loss in simulation effectiveness. The difficulty in observing psychological fidelity places speculation on transfer of training results. Psychological fidelity also emphasizes the individual confounds in transfer research—two individuals may judge the psychological fidelity of a training system very differently.

—Liability of end users is a concern in transfer of training studies. To perform a true transfer of training study, the training effectiveness must be measured in the real world. To do so, a participant must be trained in the new and often unverified training platform, and then allowed to operate the real system. If the alternate training system is inadequate, or negative training occurs, the end user could be at risk.

—Transfer of training studies have been performed inappropriately in the past, decreasing the general validity of these types of studies to system owners. "Studies that have used poor experimental designs, small sample sizes, or suffering from multiple confounding variables" (Caird, 1996) have contributed to this impression. For example, Tate, Sibert, and King (1997) conducted a transfer of training study to determine the efficacy of VE training for shipboard firefighting. They trained Navy firefighters in an IVE training program and compared their performance (time to locate fire, number of wrong turns finding the fire, time to fight the fire, and anecdotal evaluations) in an actual shipboard firefighting task to trainees who received traditional Navy firefighting training (mission review, study of ship maps, etc.) Data trends indicated that VE training resulted in lower elapsed time to locate the fire and put it out, less wrong turns, increased confidence, and better ability to concentrate. However, statistical analyses were inconclusive, because only 12 participants were recruited, which were then assigned to one of the two training groups.

—To adequately test a system it is essential that representative operators are used for testing, and "access to an appropriate community of trainees is often problematic" (Darken, 2009). This is due to either the unavailability or exclusivity of trainees. The sample size quandary also often results in the exclusion of a control group. As Roscoe and Williges (1980) point out, at least two groups are required for proper transfer studies: "an experimental group previously trained on another, usually similar, task is compared with the learning performances of a control group having no special previous training" (p. 182).

—A proper transfer of training study should at some point involve subject matter experts either for the study's design, implementation, or analysis. Similar to trainees, the availability of specialized experts often is limited. Subject matter experts also introduce subjective measures to a transfer study. Although subjective measures do provide insight into trainee preferences, they seldom provide any indication of training efficacy. Case in point, there is virtually no empirical proof for the benefit of motion-based flight simulators, although many researchers have been able to indicate pilots' preference for such trainers (McCauley, 2006).

—There are a large number of training systems to be evaluated, and this magnitude prevents the use of transfer of training studies for all of them (especially considering the previously mentioned technology, trainee and SME availability constraints). Therefore, proper transfer of training studies are often reserved for high-profile projects, or programs with larger budgets.

—A transfer of training study requires measurement of performance in a real-world setting. Field research like this often provides difficulty when collecting objective data. Alternatively, subjective (and arguably less descript) data is often collected. Worse still, in many cases no performance measures are taken in the field at all, making the efficacy of a training system entirely speculative.

—Transfer of training studies not only cost a lot financially, but they consume large amounts of time and resources. Similarly, within a VE training system, a world model must be developed before a transfer of training study is conducted. These models can be very costly and time consuming to construct before the training system is even tested and approved.

—Because of the cost of transfer of training studies, many organizations can not afford to produce/design their own unique training system. Therefore, many organizations purchase existing generic training systems which have not been validated for their specific needs.

—Technological limitations hinder proper transfer of training research. Kozak, Hancock, Arthur, and Chrysler (1993) conducted a study to determine the amount of VE transfer on a spatial pick and place task. They concluded that "learning did not transfer to the real-world task" (p. 782), and what was learned in the VE was "specific only to the context of virtual reality" (p. 782). However, each participant was given 20 minutes of individual instruction and practice with a VE system (unrelated to the research task), indicating a general unfamiliarity with VR and/or a counterintuitive VE design. It is unknown how much this apparent counterintuitive training environment affected the amount of transfer. The VR system also incorporated no tactile feedback, which is very beneficial for picking and placing of objects in a VE, and there was a noticeable lag within the VE training system. These shortcomings are certainly a result of the VR technology available in 1993, and accordingly the results should be interpreted with some reservation.

—Transfer of training studies are very susceptible to confounding variables and individual differences. Given trainee and SME availability issues, studies typically have small sample sizes greatly exacerbating exasperating these influences. Similarly, technological capabilities can affect transfer of training studies, as different systems introduce differential amounts of system and task error.

—"Transfer studies are rarely part of the system design or integration cycle," (Darken, 2009) decreasing the likelihood that program managers will spend superfluous funds on them.

—As Darken (2009) points out, the most critical reason why transfer of training studies are not as common is the current cultural view of training. "The training community still views training as a 'thing', not an outcome," and decision makers typically get influenced by technological capabilities before consideration is put on the human design of training systems.

—Studies typically are performed with college students performing a specialized, and typically non-stressful task, preventing generalization of results (Bliss, Tidwell, & Guest, 1997). As Baldwin and Ford (1988) point out, representative samples may be vital to transfer of training research.

—Some relevant transfer of training studies are 'lost' in other fields where they are considered "learning" or "education" (Kort, Reilly, & Picard, 2001). Inconsistent definitions of both transfer of training and VEs also exist in other fields of research. For example, Wagner and Campbell (1994), in a management context, use "virtual environments" to refer to an "outdoor-based experiential training program" in which participants practice leadership and management skills in an outdoor environment (e.g., "ropes course"). Furthermore, they refer to transfer of training from a VE as the skills learned in outdoor training transferring to a work situation. Finally, they provide guidelines for the implementation of such "virtual reality activities" without mention of any technology associated with typical VR apparatuses. These definitions are inconsistent with commonly used definitions presented above.

—Results of transfer of training studies are often task dependent, although Driskell et al. (2001) provide evidence that learning can transfer to novel tasks.

The research presented in this thesis attempts to learn from the difficulties encountered in past research on training transfer. This project's experiment designs were constructed keeping these barriers to transfer of training in mind.

3. Alternatives to Transfer of Training Research

In order to alleviate some of the above barriers to transfer research, alternatives to traditional transfer of training studies have been proposed and empirically evaluated. Two of these alternative methods, pseudo-transfer and backwards transfer, are discussed here.

Pseudo-Transfer

Some VE transfer of training studies measure transfer performance in a simulator rather than the real world. Pseudo-transfer studies are mostly done for systems in which real-world performance measures are either impractical or too costly. Aviation is one discipline in which pseudo-transfer research is prevalent due to the danger of practicing some abnormal or emergency scenarios in a real aircraft, and the cost associated with flight time. For example, Lintern and colleagues performed a series of pseudo-transfer studies to test the effects that flight simulators have on both air-to-ground attack and landing performance (Lintern, Sheppard, Parker, Yates, & Nolan, 1989; Lintern, Sivier, & Roscoe, 1990; Lintern, Taylor, Koonce, Kaiser, & Morrison, 1997; Lintern, Thomley-Yates, Nelson, & Roscoe, 1987). In these studies, participants were first trained in simulators varying in physical fidelity, scene detail, field of view, and augmented feedback. Following their simulator training, trainee's performance was tested in the simulator.

Another approach to pseudo-transfer research is exemplified by Loftin and Kenney (1994) to assess the “first large-scale implementation of VE technology for training personnel for a ‘real’ [space] mission” (p. 2). They performed an actual space mission following VE training of over 105 controllers and operators. However, they were unable to collect any empirical measures in the transfer task (repairing the Hubble space telescope), and measured the efficacy of their training using subjective questionnaires.

In pseudo-transfer studies, positive transfer does indicate that information was learned in the simulator. However, these results should be interpreted with some reservation, as the transfer performance gives no definitive indication of transfer to the real-world task. For this reason, there are many oppositions to pseudo-transfer studies. Salas, Bowers, and Rhodenizer (1998) point out "the determination that the training is effective should come from the trainee's performance rather than the performance of the simulation" (p. 203). Further, the trainee's performance should be observed in the transfer task with representative participants. Only when transfer research is conducted in this manner, can the efficacy of a simulator be definitively defined.

Pseudo-transfer research may be appropriate for exploring empirical relationships prior to conducting more elaborate and costly true transfer research. For this reason, this thesis utilizes a pseudo-transfer study (see Chapter III) to test relationships between stress and transfer of training prior to testing these effects in flight.

Backwards Transfer

Another alternative to traditional transfer research is backwards transfer, where participants are first trained in the real world and then tested in a training system (Adams

& McAbee, 1961). The backwards transfer approach is much more feasible for a system in which users are already proficient, and a new training system is implemented (e.g., flight simulators). The backwards transfer paradigm relies on Osgood's (1949) similar elements theory, indicating that regardless of the direction of transfer, if two tasks share similar elements then training will occur and the training system will be effective for those tasks. Therefore, if an expert operator performs adequately in a simulator (i.e., positive backwards transfer), then the simulator may be sufficient, and transfer from the simulator to the real world also is expected. However, if the similar elements paradox occurs and the training system is not congruent with the real-world task, negative backwards transfer will occur evidenced by a decrease in simulator performance when compared to real-world performance. The decrease in simulator performance is thought to be an indication of an inappropriate or ineffective simulator.

Few empirical studies have validated the backwards transfer paradigm. Kaempf, Cross, and Blackwell (1989) used a backwards transfer study to evaluate an AH-1 flight simulator for training emergency touchdown maneuvers (ETMs). Instructor pilots that were highly proficient in ETMs performed the maneuver in the AH-1 simulator. Results found negative backwards transfer, with an unsatisfactory SME rating of 82%, and 53% of simulated EMTs resulting in a crash. These findings in addition to unsatisfactory subjective participant evaluations resulted in the authors determining the simulator unsuitable for training ETMs.

Stewart (1994) conducted a similar study where 10 experienced AH-64 pilots performed simulator missions with no preflight orientation or warm-up. For the most part, the pilots were able to meet performance criteria for a variety of flying tasks in the simulator. Backwards transfer, in conjunction with mostly positive subjective ratings, provides insight about the efficacy of this particular flight simulator.

As Adams and McAbee (1961) and Stewart (1994) point out, backwards transfer studies are not limited to highly experienced, expert operators. In fact, by manipulating the level of operator experience, correlations can be determined between experience and trainer performance, although there is little research in this area. In addition, in terms of flight simulators, these researchers suggest that total flight time rather than time in type be used as a measure of pilot expertise. Total time measures determine likelihood of general piloting skills transferring to the simulator rather than skills in the particular airframe of interest.

4. VE Factors Affecting Transfer of Training

There are many factors that may affect transfer of training. Baldwin and Ford (1988) present a model of the transfer process that concisely outlines the relationships between training inputs, training outputs, and transfer of training (see Figure 2.2). Although their model is a management training approach to transfer, it provides a useful framework for the definition of transfer of training in general. Baldwin and Ford also present three categories of training inputs in their transfer process: training design, trainee characteristics, and training environment. Given that a training system's efficacy is reliant on its inputs, these three levels of analysis were used for the construction of the empirical research conducted in this thesis. Following is a review of training transfer research using these three inputs as levels of analysis.

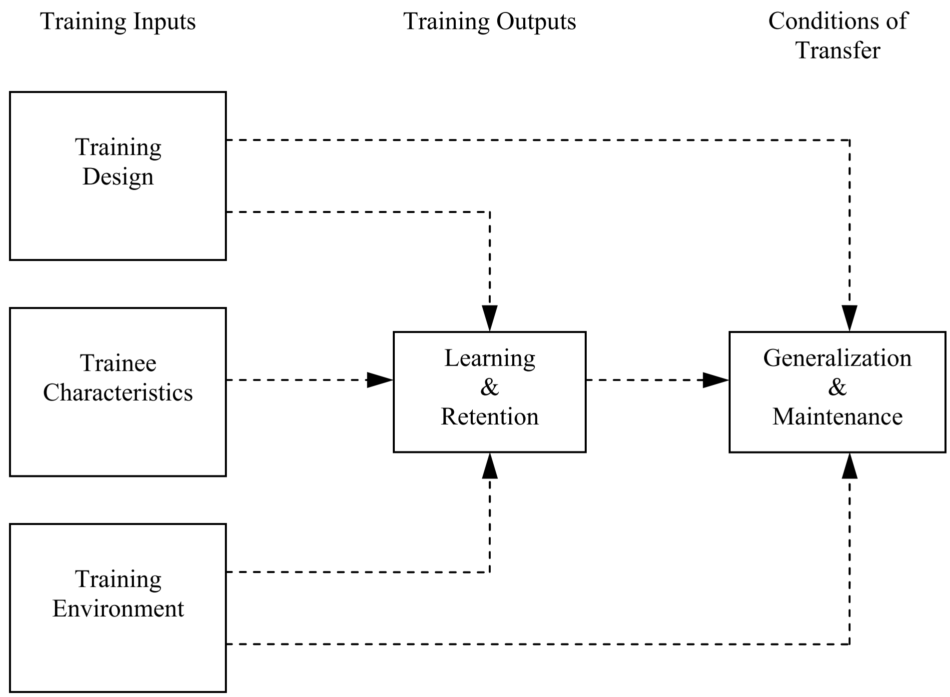


Figure 2.2 A model of the transfer process (adapted from Baldwin & Ford, 1988)

Training Design

Training design is the most thoroughly researched input for training transfer. Similar to Osgood (1949), Thorndike and Woodworth (1901) first demonstrated that training systems and the real world must share identical elements in order for positive transfer to occur. They asserted that, even if the training and transfer tasks differed somewhat, positive transfer would occur if the two tasks shared some identical elements. Transfer of training was further found to be dependent on the number of sensory and motor elements shared between the training and transfer task (Rose, Attree, Brooks, Parslow, Penn, & Ambihaipahan, 2000; Thorndike & Woodworth, 1901).

When designing training systems, it is imperative to first evaluate the cognitive, physical, and emotional elements of the real world task. Darken (2006) recommends accomplishing this by first conducting a cognitive task analysis of the task to be trained followed by a human ability inventory. By comparing these two products and pairing human abilities to training tasks, deficiencies in a training system's design can be determined. Using Darken's approach, the training for this research project was constructed by, first, considering the elements of the transfer task. The training was then constructed to incorporate these tasks. Finally, session dry runs were conducted to ensure that participants could accomplish the tasks given the training they received.

Another important question is how closely should a training system represent the real world transfer task? When considering fidelity for training design, most learning theorists agree that the learning pedagogy, training development and implementation, and accomplishment of learning objectives are of utmost importance for a training system's efficacy (Caird, 1996; Clark, 2001; Cobb & Fraser, 2005; Salas, Bowers, & Rhodenizer, 1998), although little empirical evidence supports these claims. The National Research Council (Druckman & Swets, ed., 1988) states that "effective instruction is the result of such factors as the quality of instruction, practice of study time, motivation of the learner, and the matching of the training regiment to the job demand"—not necessarily on the incorporation of fidelity. Furthermore, design should be directed by pedagogy, not technology—training systems should be constructed with a clear view of how they support learning and ensure that learning requirements are foremost (Mayer, 2005; Moshell & Hughes, 2002). Other researchers have gone so far as to say "the closer the similarity—the fidelity between the simulator and the aircraft—the more effective the instruction" (Dion, Smith, & Dismukes, 1996, p. 38). Following is research supporting these two different arguments.

In two quasi-transfer studies conducted by Lintern and colleagues (Lintern, Roscoe, Koonce, & Segal, 1990; Lintern, Sivier, & Roscoe, 1990) students were trained landing skills in low-fidelity flight simulators. They found that students soloed in less time than control conditions (i.e., no training) after as little as two hours of simulator training. They also found that, for low-fidelity simulator transfer, pictorial displays were more effective than symbolic displays. Lastly, transfer was further enhanced by supplementing low-fidelity simulators with augmented reality. Both studies conclude that fidelity has no effect on skill transfer. In addition, by adhering to specific principles of augmented feedback, reducing simulator fidelity may actually aid in training.

Taylor and colleagues (Taylor, Lintern, Koonce, Jefferson, Kaiser, & Morrison, 1991; Taylor, Lintern, Koonce, Kunde, Tschopp, & Talleur, 1993) not only found that low-fidelity simulators were suitable for training, but low-fidelity simulators actually out-performed higher-fidelity simulators. Taylor et al. (1991) taught landing skills to flight students in either a low-scene or a moderate-scene detail simulator and transfer performance was assessed in an aircraft. Students trained in the low-scene condition outperformed the moderate group in the number of aircraft landing training sessions required and instrument pattern performance. Taylor et al. (1993) also taught participants takeoff, closed patterns, and ILS procedures in simulators varying in scene detail and field-of-view. After measuring aircraft transfer performance, they concluded that a low-detail scene and a wide field-of-view was superior to the other combinations.

Waller, Hunt, and Knapp (1998) performed a study to test the effects that the VE interface (physical fidelity) has on the transfer of navigation skills. During the training phase of their study, their participants explored a 14-foot x 18-foot maze using their respective training method—a real world group explored the real test maze, a map group studied a map of the maze, a desktop group explored a model of the maze on a 21-inch color monitor, and an IVE group explored the same model using an HMD. During six different test trials, participants navigated the real maze blindfolded, attempting to follow the same predetermined route that they experienced in training. Transfer of training was measured as time to navigate the maze (touching predetermined waypoints along the way) and errors while navigating (hitting walls and following an incorrect path). Finally, in an integration task using the same waypoint locations, two walls of the maze were moved and blindfolded participants were asked to navigate the waypoints in a slightly different order than before; a true/false test was given to test their acquisition of new maze knowledge. Waller et al. found that, in general, participants that trained in the real world had superior performance to all other training conditions for early test trials. Further, participants in all VR conditions performed worse during these initial trials than those in both the real world and map conditions. However, the IVE-trained participants' performance drastically improved during later test trials (see *Trainee Characteristics* below).

Péruch, Belingard, and Thinus-Blanc (2000) performed two transfer experiments in a natural environment using a college campus as their test environment. In the first study, rather than manipulating the physical fidelity (Waller, Hunt, & Knapp, 1998), they manipulated the environment fidelity, using a rich, medium, or poor VE model of the college campus. They allowed three groups of 30 participants to explore the three different VEs and subsequently measured their ability to indicate the direction, travel distance, and direct distance to target objects in the real environment. They determined that the rich and medium groups outperformed the poor-fidelity group in direction and travel distance performance, indicating a possible reliance on an environment grid (the poor model did not incorporate campus roads) for spatial knowledge transfer. In their second experiment, they compared real world training to VE training using a high fidelity model of the same campus as the previous study. Two groups of sixteen participants were allowed to explore their respective training environment, and were tested using the same dependent measures as the previous study. Although they found the real environment training more conducive to transfer of training, they did observe transfer from the VE group.

Wilson, Foreman, and Tlauka (1997) attempted to test the efficacy of a low-fidelity computer-based simulation for spatial knowledge transfer to an office building that was trained by either real exploration or on a computer-based simulation. Their VE was not highly immersive, but it was interactive and modeled to-scale. A control group that received no training also was used. The real exploration group was allowed to freely explore the building while experimenters pointed out six target objects (e.g., specific doors, fire alarms, etc.) The computer group explored the building using the VE and the same virtual targets were identified. Each exploration group (including the control group) was then further divided into two test groups: either real building testing or virtual testing. Testing included (1) a pointing task that required Euclidean distance estimations to the obstructed target objects, (2) a route estimation task that required a distance estimation of a provided route through the building, and (3) a freehand drawing of the building. Depending on the perspective building floor that was tested, the real training, in general, resulted in better transfer of spatial knowledge than the VE training, and the VE training resulted in better transfer than the control group.

In summary, many studies have shown effective transfer of training from low-fidelity simulators. Despite these findings, many designers are over-reliant on high-fidelity simulation and customers also demand such systems. This situation has led to the increasing misuse of simulation to enhance learning of complex skills. Given the empirical evidence supporting transfer from low-fidelity simulators, the research apparatus in this project incorporates a low fidelity flight simulator. Low-fidelity simulators also introduce a certain level of experiment control into a research design. By limiting the simulator's displays and controls to only what is needed for the given tasks, research participants are less distracted by irrelevant information. For this reason, the simulator in this research project was simplified to include only pertinent information for the task (see Chapters III and IV).

Trainee Characteristics

Baldwin and Ford's (1988) transfer model also identifies trainee characteristics as an important training input (see Figure 2.2). The first trainee characteristic that is apparent when designing a training system is the trainee skill level, differentiated by Dreyfus' (2002) stages of learning: novice, advanced beginner, competence, proficient, and expertise learners. Each of these learner types can respond to a training system differently. For example, a high fidelity trainer for a complex system (e.g., a modern flight simulator) may be inappropriate for a novice learner as their information processing capabilities are quickly surpassed with a surplus of new information and stimuli. Likewise, an expert learner (e.g., a highly experienced aircraft pilot) may quickly become complacent in a classroom setting discussing basic flight information.

Darken and Banker (1998) tested the effects that trainee ability (beginner, intermediate, or advanced) has on VE navigation training transfer. Fifteen participants were recruited to test transfer of training from a high fidelity VE of a 1200x700 m^2 area in an outdoor, natural environment, sport orienteering course. All of the participants were trained using a map in addition to the VE, the real environment, or using a map only. After exploring the environment with their respective training method, participants were tested in a sport orienteering task that requires navigation through a natural environment using a map and compass to find natural landmark waypoints in a prescribed sequence. Task performance was measured as deviations from the route, distance off the path when a deviation occurs, and the number of times the map and/or compass are referenced. Their results indicate that the ability level of the participants was more important to navigation performance than the training method with the intermediate users benefiting from the VE training method more than the advanced or beginner users.

Hays et al. (1992) meta-analysis also investigated skill level and reported that flight simulator training that allowed trainees to progress at their own pace was found to be advantageous. A customized approach like this supports the assertion that individual differences are certainly a factor that affects transfer of training from VEs. Given the effects that skill level has on transfer performance, skill level was controlled in the current thesis research. In addition, objectively defining skill levels can be problematic (e.g., what constitutes an "expert" pilot?). Therefore, novice participants, which are easily defined as having no experience, are exclusively used in this research project.

Another factor that may vary between trainees is the amount of exposure to a training system. In their review of flight simulator studies, Carretta and Dunlap (1998) found that real world performance improved as the number of simulator sorties increased, although this gain leveled off after approximately 25 simulator missions. Lintern and colleagues (Lintern, Sheppard, Parker, Yates, & Nolan, 1989) found similar results, concluding that little additional benefit was found from simulator training after 24 training trials. In Waller et al., (1998) navigation study (see above) a sixth condition, IVE-long, was added where participants trained in the VE for longer than the other training groups. By the sixth test trial, the long VE training group's performance equaled the real world group, and was better than the other training groups (map, desktop, and VE). These results can not be attributed to learning during the test trials, as the control group was far worse than all other groups for all test trials. Waller et al. (1998) conclude that "when participants are allowed only short exposures, VE training may be no more effective than training with a map, and immersive VE training may be no more effective than a desktop VE." However, with more exposure, VE training may equal or even surpass map and real world training.

These two apparent conflicting views on simulator exposure present an interesting property of training in VEs. Due to relative unfamiliarity with VEs (when compared to the real, physical world) users may require sufficient training trials to become familiar with the VE (see Figure 2.3, A). However, once VE familiarity is achieved, as Waller et al. indicated (1998), VE training can quickly equal or surpass other training methods including real world training (B). After all of the possible skills and requisites are obtained from the simulator, relatively low fidelity may deter from further retention and real-world training may be required for subsequent skills acquisition (see C; Carretta & Dunlap, 1998; Lintern, Sheppard, Parker, Yates, & Nolan, 1989).

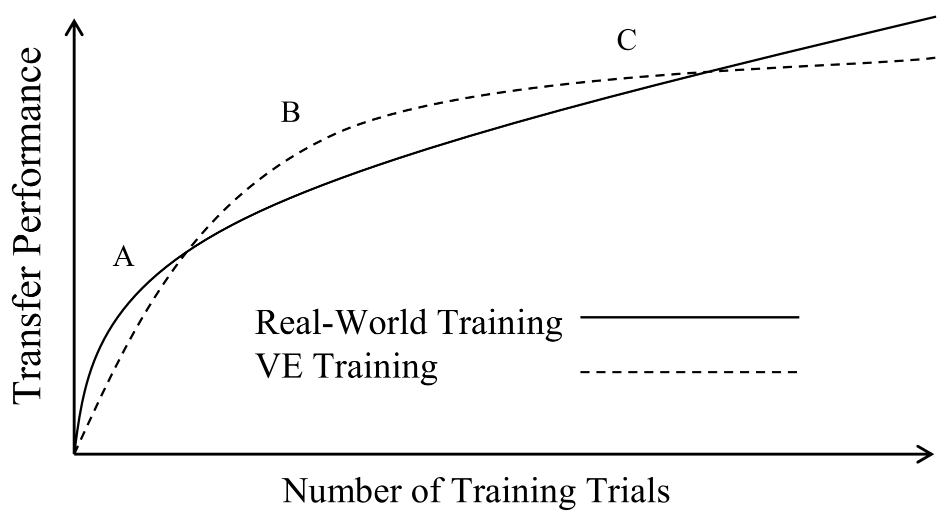


Figure 2.3 VE training exposure model

The theoretical model presented in Figure 2.3 has important implications for the research in this thesis. The VE flight training provided in this thesis can be represented by the dashed line. The participants in this research performed VE training until reaching asymptotic, or criterion, performance (see Figure 2.3, B). However, even though a criterion level is reached, this model indicates that a learning curve may still be achieved during the real-world transfer task (after C).

A final factor concerning trainee characteristics is gender differences. Very little research has investigated gender differences pertaining to transfer from VEs, but Waller et al. (1998) found that women performed consistently worse than men on transfer of navigation skills from VEs. More research is needed to determine gender difference for VE transfer. This research project evaluates gender as a covariate in statistical modeling (see Chapters III, IV, and V).

Other trainee characteristics such as ability, personality, motivation, and other individual differences still require attention regarding transfer of training research. As transfer of training is better understood with sound empirical research, these characteristics will eventually need to be addressed to better understand and enhance training efficacy.

Training Environment

Training environment is an important training input (Baldwin and Ford, 1988; see Figure 2.2). Related to environmental factors, many researchers are finding that transfer of training is very task dependent. For example, fidelity may be more crucial for highly dynamic, visual-spatial tasks (e.g., an aircraft pilot) than for more mundane, cognitive tasks (e.g., a flight engineer). Although there may be underlying training constructs that equally effect many different domains, the task for which a training system is developed and its associated, unique training requirements should not be overlooked.

Given the aviation context of this thesis, it is important to explore research relevant to the transfer of flying skills from flight simulators. Four comprehensive reviews spanning 59 years have evaluated transfer of training research pertaining to aviation. First, Orlansky, String, and Chatelier (1979) surveyed 34 studies from 1939 to 1977 that addressed transfer of flying skills from flight simulators. Their review found a median TER of 0.48 and determining that "pilots trained in simulators needed less time in aircraft to perform acceptably than pilots trained only in aircraft" (p. 101). Hays, Jacobs, Prince, and Salas (1992) performed a meta-analysis of 26 flight simulator transfer studies from 1957-1986. They found that "90% of the experimental comparisons favored simulator and aircraft training over aircraft training alone" (p. 73). They also found little support for the implementation of motion systems for flight training, a finding that other researchers have observed (Bürki-Cohen, Soja, & Longridge, 1998; Hays, Jacobs, Prince, & Salas, 1992; McCauley, 2006; Salas, Bowers, & Rhodenizer, 1998; Vaden & Hall, 2005). Finally, Carretta and Dunlap (1998) reviewed 13 flight simulator transfer studies and found that simulators facilitated the training of landing skills, bombing accuracy, and instrument and flight control.

Bell and Waag (1998) performed a literature review of 17 studies evaluating the effectiveness of flight simulators for training combat skills, and organized their review according to three general combat tasks. First, they reviewed literature that suggests that flight simulators are effective at transferring air-to-ground, conventional weapons delivery skills. Bell and Waag found insufficient data pertaining to interdiction and close air support missions, which prevents transfer determinations to be made for these tasks. However, more recent demonstrations have shown that "there is significant training potential for simulation training for both the ground attack and air-to-air environments" (p. 229). Transfer of training research regarding air combat maneuvering were contradictory, potentially attributed to inconsistent and possibly insensitive measures of performance in the transfer tasks.

Other transfer of training research has demonstrated effective transfer of spatial skills (Clawson, Miller, Knott, & Sebrechts, 1998; Péruch, Belingard, & Thinus-Blanc, 2000; Wilson, Foreman, & Tlauka, 1997; Witmer, Bailey, & Knerr, 1996), motor skills (Kozak, Hancock, Arthur, & Chrysler, 1993; Rose, Attree, Brooks, Parslow, Penn, & Ambihaipahan, 1998; Rose, Attree, Brooks, Parslow, Penn, & Ambihaipahan, 2000), and laparoscopic surgery skills (Seymour, Gallagher, Roman, O'Brien, Bansal, Andersen, & Satava, 2002). In some instances, VE training can even equal real-world training transfer performance (Clawson, Miller, Knott, & Sebrechts, 1998; Darken & Banker, 1998; Waller, Hunt, & Knapp, 1998) making the previously mentioned advantages of VEs even more appealing.

Aside from meta-analyses that explore transfer in different domains, very little research empirically and directly compares the effects of specific environmental characteristics on transfer of training. Given the apparent effects that a training environment has on information processing, it is not inconceivable that there are many environmental factors that may conversely either facilitate or inhibit transfer (Baldwin & Ford, 1988). One of these factors, stress and its associated levels of arousal, has been addressed in regard to training and is shown to be a critical factor in the transfer of knowledge and skills (Blascovich, 2007; Driskell, Johnston, & Salas, 2001; Friedland & Keinan, 1992; Kort, Reilly, & Picard, 2001; Overton, 1964; Prather, 1971; Rizzo, Morie, Williams, Pair, & Buckwalter, 2005; Saunders, Driskell, Hall, & Salas, 1996). Stress, and the human body's response to stress, is central to this research project. The following section will address stress and its effects on performance.

C. STRESS

There is great disparity surrounding the concept of stress due to its "exceedingly broad and poorly defined" nature (Berntson & Cacioppo, 2004, p. 59). Stress can be defined as when "the perceived demands of a situation tax or exceed the perceived resources of the system (individual, group, community) to meet those demands, especially when the system's well-being is at stake" (Lazarus & Folkman, 1984, p. 8). Early stress researcher, Hans Selye, contemplated if stress is "effort, fatigue, pain, fear, the need for concentration, the humiliation of censure, the loss of blood, or even an unexpected great success which requires complete reformation of one's entire life?" (1974, p. 12). Therefore, stress is often defined in terms of the environmental stressors that evoke stress and the human responses to stressors.

Selye (1974) proposes that there are three different responses to environmental stressors. First, from a biological standpoint an organism can exhibit a catatoxic response to stressors in which a fight response attempts to actively eliminate a stressor. This response can be either impossible or very taxing on an organism. In a combat context, fighting peripheral stressors (e.g., heat, radio noise, ambushing enemy troops, etc.) often distracts from the operational mission. A second response to stressors is the flee response in which an organism attempts to remove itself from a stressor. In a combat context, this response is typically infeasible short of retreat or aircraft rejection (which presents other stressors!) Finally, and often the most plausible response to stressors as defined by Selye, is a syntoxic response in which an organism learns to co-exist with stress. It is this coping mechanism that allows an organism to exhibit goal-seeking behaviors in the presence of stressors—this behavior is most favored in a combat setting and is the premise for this research endeavor.

Stress can have a profound effect on performance, but this effect is not always detrimental. This interaction is demonstrated by the Yerkes-Dodson Law (Broadhurst, 1957; Yerkes & Dodson, 1908; see Figure 2.4). An optimum amount of cortical arousal corresponds to the highest performance for a given task. Further, as more stress is introduced, resulting in heightened arousal and an appraisal of more demands than resources, performance will deteriorate. The same is true for low levels of stress (and associated arousal) in which boredom and/or complacency can result and performance suffers again.

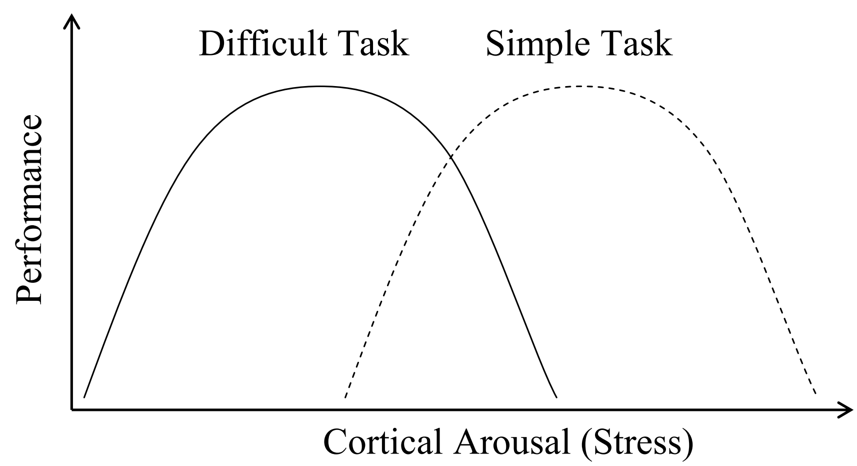


Figure 2.4 Yerkes-Dodson Law

It is important to emphasize the task dependency of the Yerkes-Dodson Law; simple cognitive tasks, which may insinuate relatively low attention and cognitive requirements, may be subject to higher levels of stress before optimum cortical arousal is achieved (dashed line in Figure 2.4). The inverse is true for difficult cognitive tasks that quickly saturate cognitive and attention capabilities. The Yerkes-Dodson Law also is subject to individual differences—novices may be quickly inundated with stress; experts may not only successfully moderate the effects of stress but they may *require* higher levels of stress to prevent complacency and optimize performance. Individual differences can refer to innate personality characteristics and coping strategies employed by different people. (Matthews, Davies, Westerman, & Stammers, 2000)

Some research criticizes the simplicity of the Yerkes-Dodson Law, indicating that the quality of performance is unknown; if performance diminishes during times of over-arousal, what types of errors will prevail? Welford (1973) suggests the signal detection theory a better vehicle for describing the relationship between stress and performance. He suggests that the response criterion equally divides noise and signal curves during times of optimum cortical arousal, thus equally distributing the probabilities of hits, false-positives, false-negatives, and correct rejections (see Figure 2.5, A). Likewise, during times of under-arousal and complacency, the response criterion shifts left, decreasing the probability of false-negatives, but also increasing the probability of false-positives (B). During these times, a person is underactive and errors of omission are more probable. Conversely, under high levels of cortical arousal and strain, the probability of false-negatives increases while the probability of false positives decreases (C). During these times, over activity increases the likelihood of errors of commission.

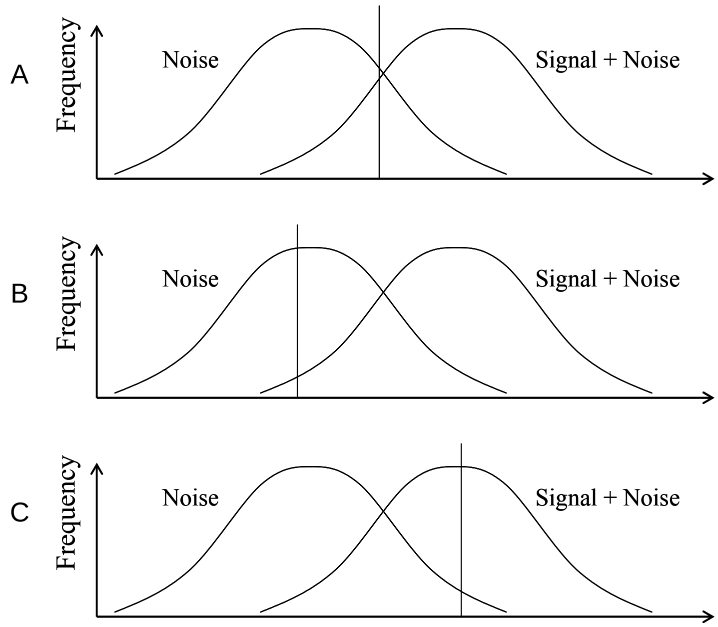


Figure 2.5 Signal Detection Theory

As the preceding discussion may indicate, "stress is neither a characteristic of the environment alone nor a characteristic of the person alone. Instead, stress is defined as a particular type of relationship between the person and the environment, in which the individual perceives the adaptive demands as taxing or exceeding available coping resources" (Meichenbaum, 1993, p. 382). In Beck's (1993) cognitive model of stress, Meichenbaum's person-environment relationship is further expanded: "The construction of a situation (cognitive set) is an active, continuing process that includes successive appraisals of the external situation and the risks, costs, and gains of a particular response. When the individual's vital interests appear to be at stake, the cognitive process provides a highly selective conceptualization" (p. 333).

There also is an apparent relationship between stress and motivation (Welford, 1973). Blascovich (2007) uses this relationship to differentiate between the type of physiological response to a particular stressor as either challenge or threat. He defines a challenge response as one in which an individual's evaluated resources outweigh situational demands; likewise, a threat response occurs when the evaluated situational demands outweigh an individual's resources. This model relies on goal relevancy (i.e., an individual is only motivated to act towards desired goals), situational demands (i.e., what the situation requires of the individual), and coping resources (i.e., the individual evaluation of required resources, or preparedness to act). For example, if a pilot's goal is to accomplish a combat mission safely, and during that mission an emergency situation arises for which the pilot perceives to be ill-prepared, a threat response to the emergency situation will result. However, given the same scenario, if the pilot appraises her available resources (i.e., previous training) as adequate, then a challenge response may result. This research project focuses on providing resources to trainees so that a challenge response results. Both stress training and skill training are utilized to accomplish a challenge response.

The challenge and threat model of stress response incorporates three levels of analysis: biological, psychological, and social psychological. First, the biological model focuses on how the cardiovascular system is affected by the interplay between autonomic and endocrine influences. Next, the psychological level refers to the evaluative process resulting from cognitive and affective influences. Lastly, the social psychological level integrates intraindividual, interindividual, and environmental influences. These three levels of analysis interact: "challenge and threat represent person/situation-evoked motivational states with affective, cognitive, and physiological antecedents and consequences" (Blascovich, 2007, p. 5). Blascovich (2007) has demonstrated a reliable and sensitive measurement of challenge and threat responses to stress as an alternative to the often disputed measures of arousal. A challenge response accurately indicates when a person perceives to have adequate resources for a situation, and is a reliable indication of performance in that situation. The converse is true for threat responses.

Critical to this research project is the effect that stress and subsequent emotional responses have on learning. Kort, Reilly, and Picard (2001) propose a model of emotional learning that they assert may be used for the development of customized, dynamic learning based on the tracking of affective learner states. A preliminary emotional set is proposed that admittedly may be incomplete and varies greatly between individuals. However, this model still provides some insight into the relationship between emotions and learning.

Axis -1. 0 -0. 5 0 +0. 5 +1. 0

Anxiety-Confidence	Anxiety	Worry	Discomfort	Comfort	Hopeful	Confident
Boredom-Fascination	Ennui	Boredom	Indifference	Interest	Curiosity	Intrigue
Frustration-Euphoria	Frustration	Puzzlement	Confusion	Insight	Enlightenment	Epiphany
Dispirited-Encouraged	Dispirited	Disappointed	Dissatisfied	Satisfied	Thrilled	Enthusiastic
Terror-Enchantment	Terror	Dread	Apprehension	Calm	Anticipatory	Excited

Figure 2.6 Emotion sets possibly relevant to learning (Kort, Reilly, & Picard, 2001)

As can be seen in Figure 2.6, these emotional states may be relevant to stress. For example, comparing the Kort et al. emotional set to the Yerkes-Dodson Law, boredom may result in low cortical arousal and a decrease in performance (i.e., learning). However, one criticism of the Kort et al. emotional set is the exclusion of the over-arousing, upper half of the Yerkes-Dodson Law inverse-U curve (e.g., an extreme positive rating on the Boredom-Fascination axis results in Intrigue, not inundation), and it is this over-aroused state that stress training attempts to alleviate.

Kort et al. (2001) further attempt to interweave their emotion axes (Figure 2.6) with the cognitive dynamics of the learning process (see Figure 2.7). In their model, the horizontal axis represents the emotional sets, and the vertical axis represents amount of learning or un-learning. For the purposes of the current research topic, if the learning axis is thought of as information learned in training, a third axis can be imagined extending from the page that represents performance in a transfer task—as positive affect and constructive learning are optimized, performance also is optimized (e.g., quadrant I).

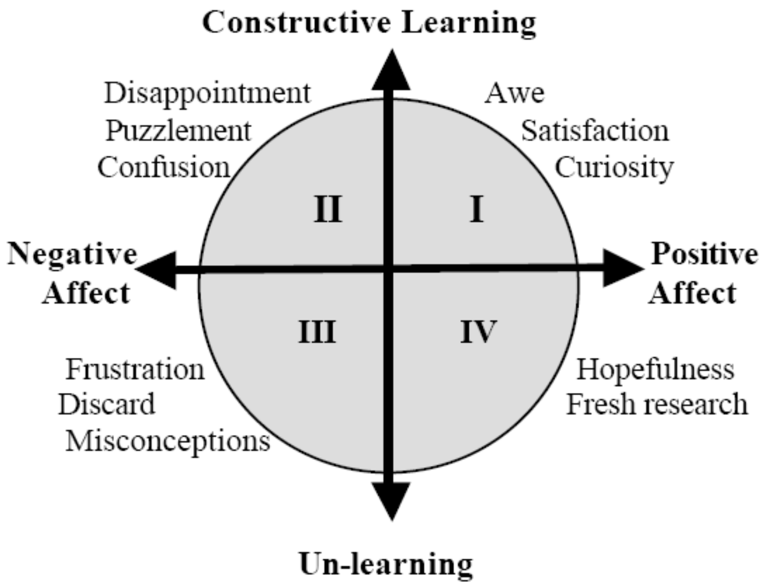


Figure 2.7 Emotional-learning model (Kort, Reilly, & Picard, 2001)

Although arousal often is used to describe the human response to stress, as Mullen, Bryant, and Driskell (1997) point out, "a major difficulty in the theoretical use of arousal has been the failure to develop a clear and unequivocal definition of this ubiquitous yet fuzzy construct" (p. 55). Therefore, arousal often is operationally defined in terms of its measurement (a thorough review of arousal measurement methods are reviewed in Chapter III, Section C).

For the purpose of this research review, arousal refers to the body's physiological response to a stressor, and stress will primarily be limited to short-term, or *acute*, stress (as opposed to long-term chronic stress).

D. STRESS TRAINING

Stress undoubtedly affects performance. In addition, many work environments are prone to high levels of performance-altering stressors. In an attempt to alleviate individual workload and stress, many attempts have been made to introduce automation, technologies, intuitive controls/displays, etc., although these approaches are often very costly, have theoretical limitations, and, in some cases, can actually increase stress.

Another approach to alleviate stress is the introduction of stress management into training. However, introducing stress training can often result in two counterproductive results: (1) high levels of anxiety that inhibit both training and post-training performance, and (2) stress that interferes with the acquisition of skills and knowledge for which the training is designed to promote (Friedland & Keinan, 1992). In an attempt to (1) effectively impart skills and knowledge on trainees, (2) expose and inoculate trainees to real-world levels of stress, and (3) alleviate the before mentioned shortcomings, Friedland and Keinan (1992) present three approaches to training for stress: graduated-intensity training, customized training, and phased training. These three approaches and associated research are presented here in order to develop a stress training protocol for the research in this thesis.

1. Graduated-intensity Training

One common problem with stress exposure training is that trainees are never provided a vehicle for becoming comfortable with the sudden onset of stress. A sudden, intense, and prolonged exposure to stress can, aside from being traumatic, be detrimental to performance. Graduated stress exposure also can decrease trainee confidence and give unreal expectations of a highly stressful and unmanageable real-world environment. A graduated-intensity approach to stress training attempts to alleviate these problems by slowly “inoculating” trainees with a progressive training protocol while theoretically not impeding task acquisition. While graduated-intensity training can provide trainees with a heightened sense of control and competency, the gradual introduction of stress can breed unrealistically low expectations about transfer task stress levels.

In a series of experiments, Friedland and Keinan (Friedland & Keinan, 1982; Keinan & Friedland, 1984) were able to empirically evaluate the effectiveness of graduated-intensity training and develop a set of implementation techniques. First, graduated-intensity stress training proved more beneficial than constant intensity training. However, this result was only true if the trainees knew the maximum stress level, or ceiling, of training. Training without a ceiling resulted in learned helplessness as trainees never knew when stress intensity (in this case electric shocks) would cease to increase.

They also showed that, as expected, stress training is no more effective, and sometimes inferior to control, no-stress conditions if the transfer task was not stressful. These results indicate that stress training can impose on skill acquisition. Finally, a graduated-intensity approach to training is most beneficial when trainees believe the stressors are contingent on their performance. This approach provides trainees with a locus of control that constant intensity trainees did not have. This last finding also reemphasizes a reoccurring trend in stress literature: stress is an individual's appraisal of a situation given the resources they have.

2. Customized Training

Due to the individual differences associated with stress, another potential approach to stress training is adapting the training to an individual's needs. By tailoring stress training to each individual, a training system can ensure that (1) appropriate levels of stress exposure are met that are congruent with a particular transfer task, and (2) specified criterion levels are met during task acquisition. One feasible customized approach to stress training is the incorporation of the challenge and threat model mentioned above (Blascovich, 2007). Given that model's reliance on individual differences, a trainee can undergo stress training while carefully monitoring the physiological responses to stress. If transfer task levels of stress are presented in a training context, then stress exposure training can be implemented until a challenge response is indicated. The challenge response would imply that the trainee perceives to have adequate resources to accomplish the transfer task under the projected levels of stress. However, this method relies on the trainee's subjective perception of resources, and may not always indicate proficiency for the task at hand.

Friedland and Keinan (1992) tested the effectiveness of customized training by first administering a confidence expectancy questionnaire to 297 male infantry soldiers. The questionnaire measures the generalized expectation about the risk of sustaining physical injury in dangerous situations, and can be likened to the challenge-threat response when confronted with danger. For example, an individual characterized by a high confidence expectancy will generally respond to a stressful situation with challenge as opposed to threat. Friedland and Keinan theorize that these individuals (i.e., high confidence expectance) may benefit more from stress exposure training when compared to individuals characterized by a low confidence expectancy. The training consisted of moving down a fortified trench in a crouch while shooting targets along their path with an assault rifle. Half of the participants performed the training in a stressful condition in which live rounds were believed to be fired over the trench and an abundance of medical provisions were on-hand. The non-stressful training group performed the same task in the absence of live rounds and medical provisions. Both groups performed a transfer task in the stressful setting.

Friedland and Keinan (1992) found that participants who scored high in confidence expectancy benefited the most from the high-stress training condition (i.e., they performed the best in the stressful transfer task after stressful training). This study provides further evidence that the stressful training task provoked a challenge response and allowed the high confidence expectancy soldiers to efficiently acquire task skills while simultaneously benefiting from stress exposure. This study also emphasizes the importance of addressing individual differences and the potential for customized stress training.

3. Phased Training

A final training approach for moderating stress while ensuring skill acquisition is to follow a phased approach to stress training. Friedland and Keinan (1992) defined and tested the interplay between three elementary phases of training: task acquisition (TA), stress exposure (SE), and practice under stress (PUS). Agreeing on the importance of task acquisition and the development of skills prior to any stress exposure, Friedland and Keinan tested five different training regiments: TA, PUS, TA/SE/PUS, TA/SE, and TA/PUS. Figure 2.8 illustrates the TA/PUS training regime.

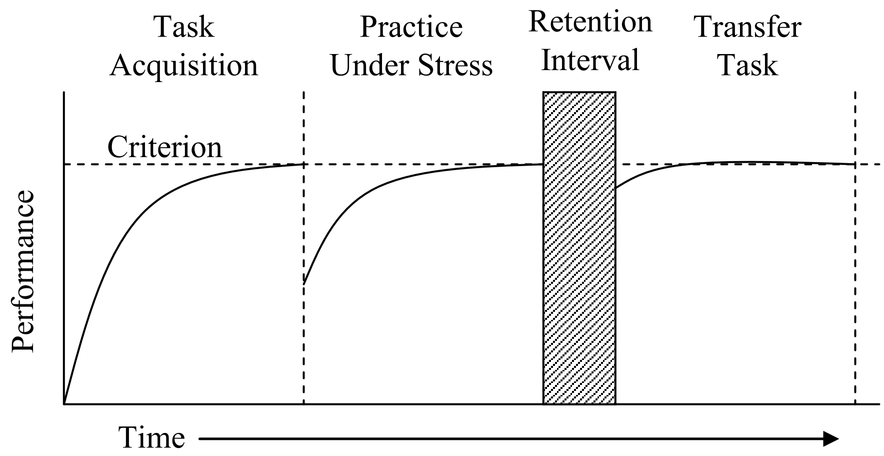


Figure 2.8 TA/PUS training regime

When performing a stressful target detection transfer task, they found that (1) the TA training protocol is the least effective, (2) the PUS protocol shows relatively poor criterion performance, and (3) the TA/PUS and TA/SE/PUS protocols are superior to the other three approaches (see Figure 2.9). These result emphasizes the importance of *both* skill acquisition and stress training when training for a stressful transfer task. These findings are promising for the future development of stress training. First, a phased approach to training is both effective and efficient (the total training time was controlled). Secondly, task acquisition is of foremost importance in any training program. Without the proper development of skills, transfer performance will suffer. Lastly, complete compartmentalization of stress exposure or skill acquisition (e.g., the TA/SE training protocol) is insufficient for a transfer task that incorporates both—rather, they both must be integrated into a total training plan. Training of the transfer task under stress is essential for transfer to a stressful task.

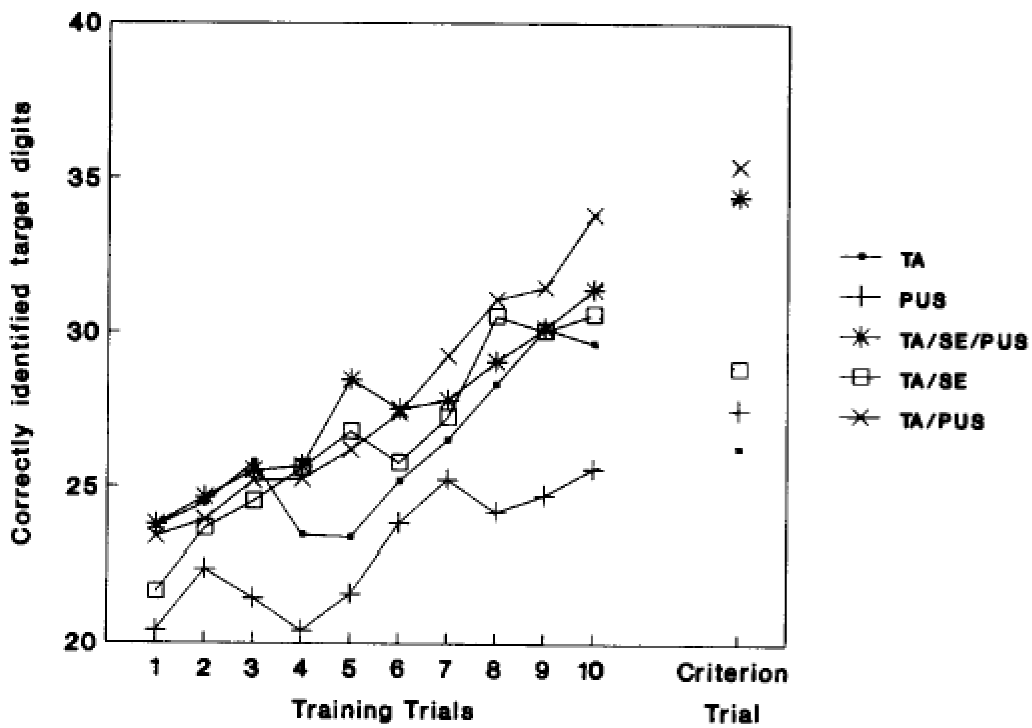


Figure 2.9 Criterion performance for target detection following five different phased training approaches (Friedland & Keinan, 1992)

One important observation from Figure 2.9 is the increase in performance between training trials and the criterion trial. As mentioned earlier regarding transfer of training measurement, two different methods may be employed during task training. As opposed to a criterion-based approach, Friedland and Keinan controlled for training time which may explain learning during the criterion trial (i.e., participants did not reach asymptote during training trials). This technique also may explain the discrepancy between their results and the theoretical model of their TA/PUS regime presented above (see Figure 2.8). Controlling for training time also may introduce more variance from individual differences unless a predetermined average time-to-criterion is established for a particular training task.

Following Friedland and Keinan's (1992) support for a phased stress training approach, many authors have expanded on the specific design and implementation of phased training. Most notably, Johnston and Cannon-Bowers (1996) provide guidance for Stress Exposure Training (SET) which evolved from three main objectives: building skills that promote effective performance under stress, building performance confidence, and enhancing familiarity with the stress environment. Similar to SET, stress inoculation training (SIT) is a phased training approach that has been applied in a mostly clinical context. As Meichenbaum (1993) points out, the primary notion of SIT is that "bolstering an individual's repertoire of coping responses to milder stressors can serve to defuse maladaptive responses or susceptibility to more severe forms of distress and persuasion" (p. 378). Both SET and SIT consist of three phases. The first stage entails conceptualization or education of work under stress and promotes the development of this requisite knowledge. Trainees are provided with knowledge of typical stress reactions and develop the know-how for dealing with stressors. As Meichenbaum points out, a critical outcome of this stage is a trainee understanding that they do not react to a stressor, rather their response is the result of their appraisal of the situation. For example, a pilot undergoing SET will first learn the physiological and psychological effects of working under stress to include increased temperature and sweating, difficulty in problem solving, and deficiencies in memory recall.

The second phase of SET is skilled practice with feedback which is similar to the second phase of SIT, *skill acquisition and rehearsal*. Skilled practice with feedback is similar to the task acquisition stage presented by Keinan, Friedland, and Sarig-Naor (1990) where trainees undergo task training in the absence of stress until a specified performance criterion is met. In the piloting example, if a pilot is performing SET for learning emergency procedures, a simulator (either static or dynamic) may be used to acquire the skills needed for aircraft recovery in the event of an emergency procedure. Criterion can be determined using one of two methods: (1) based on historical data, an average time may be determined for which most trainees will reach asymptotic performance, or (2) an adaptive approach may be employed where each trainee performs until asymptotic performance, or a performance criterion, is achieved. Given the inclusion of feedback, the outcomes of this phase as described by Johnston and Cannon-Bowers (1996) are (1) reduced negative attitudes towards self and stressors, (2) increased use of positive thoughts and behaviors, (3) reduced blood pressure, heart rate, and increased psychomotor steadiness, and (4) successful coping skill performance. Figure 2.8 illustrates phase 2 of the SET process as "Task Acquisition."

The third SET phase, skill practice with stressors, is likened to Keinan et al.'s (1990) performance under stress where the skills learned in phase two are practiced under transfer equivalent stressors. Returning to the above example, now that the emergency procedures are learned and rehearsed to proficiency, stress is introduced. A significant decrease in task performance should be observed at the beginning of this phase due to the introduction of stress (see Figure 2.8, "Performance Under Stress"). However, the training should continue (potentially requiring multiple exposures/trials) until performance returns to the phase two asymptote. Using this criterion level of performance should imply that the task can be performed under stress without any noticeable hindrance of stress, suggesting that cognitive capabilities are fully utilized.

In a meta-analysis of 37 studies testing the effectiveness of a three-phased approach to stress training (i.e., SET or SIT), Johnston and Cannon-Bowers (1996) determined that 67% of studies demonstrated that stress training significantly improved performance. In their meta analysis, Saunders, Driskell, Hall, and Salas (1996) determined that stress training was effective in reducing performance anxiety, reducing state anxiety, and enhancing performance under stress. They further analyzed their results by various stress moderators, and these analyses will be reviewed in the following section where appropriate.

Comparing the three training approaches above reveals that phased training is the most effective for imparting skills and introducing stress mitigation techniques. Following on work by Friedland and Keinan (Friedland & Keinan, 1982, 1992; Keinan & Friedland, 1984; Keinan, Friedland, & Sarig-Naor, 1990) and other phased-training research (Johnston & Cannon-Bowers, 1996; Meichenbaum, 1993; Saunders, Driskell, Hall, & Salas, 1996), a three-phased training protocol is developed for this research project and presented in Chapter III.

4. Factors Affecting Stress Exposure Training

There are numerous models addressing stress in training. The transactional theory defines the human reaction to stress as a product of the human-environment interactions. Hancock, Ward, Szalma, and Ganey (2002) propose a 3-dimensional descriptive framework for the transactional theory. They propose that moderators of stress, information processing, and sources of stress are all critical components of the appraisal of stress. When reviewing the before-mentioned VE factors affecting transfer of training (see Figure 2.2; Baldwin & Ford, 1988), Hancock et al. matrix is congruent with training design, trainee characteristics, and training environment, respectively, and it is apparent that these same three levels of analysis are useful for understanding the implementation of stress training.

Johnston and Cannon-Bowers (1996) suggest a systems-theory orientation for modeling stress exposure training that specifies inputs, throughputs, outputs, and feedback loops. Beck (1993) defines a cognitive model of stress from a clinical perspective that includes eleven principles affecting human stress responses. Finally, in addition to a systems-theory approach not unlike Johnston and Cannon-Bowers (1996), Saunders, Driskell, Hall, and Salas (1996) present seven effects of moderators when measuring the effects of SIT.

In an attempt to better understand the relationships between VR technologies, stress training, and transfer performance in stressful situations, Figure 2.10 provides a model of stress training. In the following discussion, this model is further defined using the following VE variables that affect stress training: training design, training environment, cognitive appraisal, trainee characteristics, stress type, and training outcomes. This analysis will determine what factors are important to consider when developing a stress training protocol for this research project.

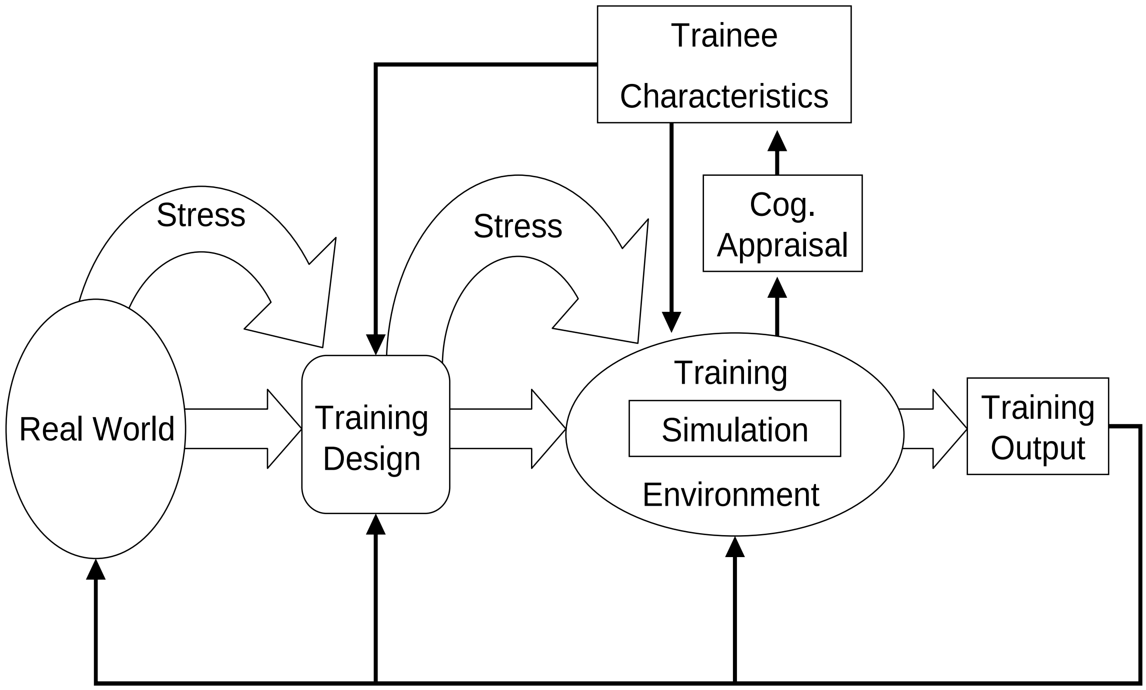


Figure 2.10 Model of stress training in VEs

Training Design

As demonstrated above concerning transfer of training, the training system design is paramount for effective stress training outcomes. A notable component of training pedagogy is Driskell et al.'s (2001) reference to structural and surface features. In an attempt to explain the generalization of stress training to various transfer stressors, they go beyond Thorndike and Woodworth's (1901) identical elements theory to emphasize the features relevant to both the training and transfer tasks. Structural features refer to "the underlying principles imparted in training," while surface features refer to "domain-specific characteristics such as the specific training examples used and the specific attributes of training context" (p. 108). For example, in an aviation context, their research suggests that a pilot will benefit from stress training even if the context (surface features) for which he received the stress training (i.e., a low-fidelity flight simulator) is different from the transfer task (i.e., an aircraft). In this case, the stress training itself is thought of as a structural feature, and in their study "it is likely that the stress training resulted in positive transfer to novel settings [and novel stressors] because the underlying principles in training were structurally consistent with the transfer environment" (p. 109).

In their meta-analysis of stress inoculation training, Saunders et al. (1996) provide further empirical evidence of the importance of structural features by comparing the setting in which trainees received stress training. By comparing stress training that was conducted in either a laboratory or field setting, they found both contexts effective for reducing performance anxiety and state anxiety.

Regarding stress training design, Saunders et al. (1996) also investigated the issue of how many stress training sessions and how long each session should be for effective transfer to occur. They found a negative correlation between the amount of stress training received and the transfer task anxiety, meaning as participants receive more stress training, their transfer anxiety decreased. They also found a reduction in task anxiety after as little as one stress training session.

Driskell, Copper, and Moran (1994) provide encouraging support for the use of VEs and simulators for stress training in their review of mental practice effects on performance. They found that, although not as effective as actual physical practice, mental rehearsal is an effective training approach. Saunders et al. (1996) further found that imagery practice was more effective than real-world practice for reducing transfer anxiety, although real-world practice of these stress reducing skills was more effective for enhancing transfer performance.

The above training design literature has important implications for the empirical work in this thesis. Given that training context is a surface feature, the research objectives in this thesis will be tested using a low fidelity flight simulator in a laboratory setting. In light of Saunders et al. (1996) results, the training in this thesis will also be introduced using a minimal number of training sessions. Finally, given the efficacy of mental rehearsal, SE will be used as part of the phased training protocol. Chapter III describes how the SE session capitalizes on mental rehearsal while exposed to stress.

As the current model depicts (see Figure 2.10), training design also is influenced by trainee characteristics. For example, given a target population of pilots, certain stress management skills may be assumed. Trainee characteristics are a crucial design consideration for training design and are discussed in the following sub-section.

Trainee Characteristics and Cognitive Appraisal

There is little doubt that different individuals respond to stressors in different ways, and "the autonomous and sociotopic personality types differ in the kind of stressors to which they are sensitive" (Beck, 1993, p. 345). Therefore, the reaction that an individual has to a stressor is largely contingent on the specific vulnerabilities of a certain personality, among other factors. Individual differences in response to stress are evident when reviewing transcripts and cockpit recordings from emergency and abnormal situations. For example, on December 6, 1999 a runway incursion occurred at Theodore Francis Green State Airport, near Providence, Rhode Island where the crew of United Flight 1448 became disoriented while taxing in instrument meteorological conditions (dense fog and one-quarter-mile visibility). Following is the transcript from the ground controller's communications:

> United 1448 stand by, please. DON’T' TALK! I have other things I need to do! … US Air 2998 isn't going yet, because I have a United that doesn't know where the hell he is… stop all traffic until advised, please. United has no idea where they are. (NTSB, 2000)

The unnerved and stern voice portrayed in this recording is further evidence that this controller responded to a stressful situation in an atypical fashion.

Parsa and Kapadia (1997) investigated the effects of stress on military participants by comparing stress levels of five different fighter pilot squadrons. The Beck Depression Inventory was administered to pilots and used as an indicator of stress during high tempo operations. Four of the five squadrons were assigned to an overseas air base with contingency operations and consisted of (1) a relatively successful permanently assigned unit, (2) an unsuccessful permanently assigned unit with a recent member death, (3) a temporary duty unit, and (4) a reserve unit, recently assigned. Finally, a fifth squadron consisted of stateside, non-combatant military pilots. The unsuccessful permanent squadron demonstrated higher stress levels than the successful squadron. In addition, combined active duty squadrons (1, 2, and 3 above) demonstrated higher stress levels compared to the reserve unit. Finally, there was no overall significant difference between combatant and stateside units. These results present evidence that both individuals and circumstances can greatly affect stress levels of military aviators.

In a review of research addressing stress coping in Vietnam veterans, Bourne (1970) attempted to find what solider attributes and characteristics allowed them to cope in the presence of stress. He found that "all of the subjects, using a wide range of defenses, were able to psychologically restructure their perception of reality in such a way that they avoided facing the danger and at the same time enhanced their own sense of immortality and invulnerability" (p. 78). Therefore, it is possible that stress training, rather than imparting skills, alters trainees' psyches—a certain amount of placebo effect makes trainees feel invulnerable to stress.

Further research addresses the relationship between experience and stress levels. Otsuka, Onozawa, and Miyamoto (2006) measured the ratios of stress hormones (catecholamine and cortisol) in both student and instructor pilots before and following training missions. They tested students and instructors in both preliminary and advanced flight training, and found that students in advanced flight training had lower levels of stress hormone than students in preliminary flight training. In addition, instructors had lower levels of stress hormone following preliminary flight training missions when compared to students, but this difference was not observed for advanced flight training. The difference effects that skill level has on stress indicates the adaptive nature of stress coping mechanisms. Although no formal stress training was provided in this study, the mere exposure to stressors (i.e., flight experience) reduced the body's sympathetic responses to stress. In addition, the relatively quick adaptation to stress is interesting—the advanced students had between 130-160 flight hours, while the advanced instructors had a mean 1,901 hours, yet they reported the same hormone levels following missions.

Beck's (1993) cognitive model of stress accounts for various cognitive constellations and suggests that these constellations are "chained" to specific stimuli. "This paring constitutes the specific sensitivity of a given individual and prepares the way for inappropriate or excessive reactions" (p. 342). Beck's approach explains why what is stressful for one individual may be benign for another, and provides support for both the customized stress training and the challenge-threat theory presented above.

Given that individual differences has a considerable affect on stress, the research presented in this thesis attempts to control the effects of individual differences using three different techniques. First, the sampling frame for selecting research participants was limited to decrease the amount of individual differences within the participant pool. Second, many efforts were made to control both the laboratory and flight research environments. Lastly, data analysis techniques measured dependent variables as a function of individual baseline or asymptotic measurements. These three techniques are discussed more in Chapters III, IV, V, and VI..

Training Environment

In response to the above research indicating the effectiveness of stress training, Saunders et al. (1996) question the generalization of these results to field settings outside of a controlled laboratory setting. Their meta-analysis revealed that stress exposure training was equally effective at reducing performance anxiety and state anxiety in both laboratory and field settings. Although their results indicate that SIT performed in a laboratory setting was effective for enhancing performance in a stressful transfer task, there were not enough hypothesis tests in field settings to draw a conclusive comparison. The research presented in this thesis addresses if stress training generalizes outside of a laboratory setting to a real-world stressor.

Meehan, Insko, Whitton, and Brooks (2002) demonstrated the effects that VE presence, as a result of the virtual training environment, can have on arousal. The authors hypothesized that "to the degree that a VE seems real, it will evoke physiological responses similar to those evoked by the corresponding real environment, and that greater presence will evoke a greater response" (p. 645). To prove this hypothesis, the researchers conducted a study in a VE with a Pit Room which incorporated a virtual unguarded hole in the middle of the floor that dropped 20 ft below to another virtual room. The within subject design had participants navigate the room while walking on a real, flat laboratory floor in one condition, and in the other condition by using a 1.5 inch platform on the laboratory floor to indicate where the virtual hole was (see Figure 2.11; Meehan et al., 2002).

Figure 2.11 A participant in the platform condition (Meehan et al., 2002)

Experimenter observations indicated a much more immersive experience when participants neared the edge of the virtual hole when using the platform. Participants were reported as feeling frightened, experiencing vertigo, and some participants asked to terminate the study immediately. Statistical results indicated that an increase in simulation fidelity (using the platform) increased both subjective measures of presence and physiological indications of arousal using heart rate variability (HRV). However, the authors admit that this study used a relatively stressful environment, and both the subjective presence ratings and physiological measurements of arousal may be attributed to the simulation's stressful nature. Further research is needed to measure the effect that presence has on arousal in more mundane tasks (Meehan et al., 2002).

Regarding this thesis research, the training context (a surface feature) is not expected to affect transfer of skills or stress coping (structure features). However, a certain amount of VE presence is desired to facilitate task acquisition. Realistic controls and displays in addition to an aviation headset are incorporated in the experiment design to enhance presence within a low fidelity flight simulator (see Chapters III, IV, and V).

Training Output

As Figure 2.10 demonstrates, and as most systems theories indicate, system output and feedback are essential to system performance. Training feedback also has been shown to enhance trainee performance (Goldstein, 1993). Prather (1971) distinguishes between stress training with and without feedback as trail-and-error and errorless, respectively. In order to test the importance of feedback obtained in trial-and-error stress training, he conducted a study in which pilots judged target distance in simulated combat strafing runs. Pilots who were trained using trial-and-error were provided electric shocks for errors in their distance judgments. Errorless pilots were presented with a light that cued them when they were the appropriate distance from the target (without electric shock feedback). Following initial training, both groups received stress training similar to the trial-and-error initial training phase. The two groups were then tested in the same range estimation task both with and without stress. Results indicate that the trial-and-error training resulted in superior transfer of training (no stress) and superior transfer performance under stress. This study emphasizes the importance of a closed feedback-loop in training for stress exposure to be effective.

Prather's (1971) results beg the question, why does stressful feedback enhance transfer performance both with and without stress? One explanation is that stressful feedback can improve performance confidence, or self-efficacy, which will provoke a challenge response and further improve performance. This improvement in performance further decreases stress feedback and, in turn, increases self-efficacy. However, there are two precautions that should be taken when providing trial-and-error stress training. First, if a trainee's performance does not improve, the increasing levels of stress feedback can result in a threat response and a subsequent decrease in performance. Similarly, as Friedland and Keinan (1992) point out, if a feedback ceiling is not identified *and* communicated to the trainee, learned helplessness can occur which also decreases training efficacy and transfer performance.

Given the effects that performance feedback can have on performance, it is important to control for performance feedback when testing other research objectives. Therefore, in this thesis research, performance feedback was strictly controlled. During both training and transfer sessions, participants were never provided clues to their performance. In an attempt to increase individual performance confidence and self-efficacy, research participants, in general, often ask how they are performing. This information must be withheld to control for performance feedback effects.

Stress Type

As research progresses towards understanding stress training, the specific method for introducing stress needs to be addressed. Researchers have attempted to introduce stress in training using various stressors. A summary of stressor types used in empirical research is reported in Table 1.

Table 1. Empirically tested training stressors

Type of Stress Exposure	**Reference**
Amphetamine	(Hockey, 1984)
Cold	(Friedland & Keinan, 1992; Kanfer & Seidner, 1973; Rosenbaum, 1980)
Electrical Shock	(Friedland & Keinan, 1992; Prather, 1971)
Heat	(Pepler, 1958)
Incentive	(Hockey, 1984)
Noise	(Broadbent, 1958; Driskell, Johnston, & Salas, 2001; Klonowicz, 1989)
Prolonged work	(Hockey, 1984)
Sleeplessness	(Hockey, 1984)
Time pressure	(Driskell, Johnston, & Salas, 2001)

Although the research presented in Table 1 introduces a variety of stressors into training, there is little empirical evidence that training under one stressor will generalize to a novel stressor. Hans Selye (1936) first explored the notion that humans respond in a generalized biological manner to various types of stress. He found that regardless of what type of stress a person (or rat in early studies) was exposed to, the same autonomic nervous system responses resulted. This generalization of stress responses, or *the general adaptation syndrome*, allows a trainee to benefit from stress training even in the presence of a novel stress.

Driskell, Johnston, and Salas (2001) attempted to test this theory by, first, training participants to perform a computer-based task under either time pressure or auditory distraction (noise stress). Following a familiarization phase (under stress), the participants were subjected to a regimented SET intervention that included further practice under the same stressor. During a final phase, participants performed the same task under a novel stressor (i.e., the stressor they were not trained on). Performance was measured as accuracy of performance on the computer-based task and self-reported stress. Driskell et al. found that SET did generalize to a novel stressor (i.e., performance under a novel stressor improved following SET on a different stressor), although they found the greatest generalization (i.e., lowest self-reported stress) when trainees received SET under the time pressure condition.

The fact that stress training may generalize to novel stressors is most intriguing when viewed in light of Thorndike and Woodworth's (1901) identical elements theory. As Driskell et al. (2001) suggest, their results indicate the relative importance of training structure features versus domain-specific surface features. Furthermore, these results indicate the promise for training stress exposure in a VE as long as the structure features of the VR training to include the expected stress levels are congruent with the real-world task(s).

The preceding literature review provides the premise for conducting three empirical experiments described in the following chapters. In general, the three experiments will test (1) the efficacy of stress training to mitigate the strain of stressors and improve performance during a stressful simulator criterion task, (2) the transfer of stress training to a flight environment, and (3) the transfer of skills from a VE to a flight environment.

These experiments were performed over the course of 12 months. The schedule for this research endeavor is provided in Appendix A. The following three chapters provide the detailed methods for each experiment.

THIS PAGE INTENTIONALLY LEFT BLANK

III. EXPERIMENT 1: STRESS TRAINING EFFICACY

There is nothing training cannot do. Nothing is above its reach.

— Mark Twain

A. OVERVIEW

As seen in Chapter II, transfer of training can be enhanced if the training protocol includes elements of the transfer task. One of the elements that may be included in training are the mental, emotional, and physiological states of the transfer task. These states can be considered structure features and have the greatest effect on transfer. One emotional state synonymous with aviation is stress, and this experiment addresses the efficacy of introducing stress into flight training. The purpose of stress training is to produce a syntoxic biological response in the trainee, which, in turn, may improve performance in a stressful transfer task.

Experiment 1 utilizes Friedland and Keinan's three-phased stress training approach to determine if stress training in a flight simulator transfers to a stressful simulator criterion session. This experiment is a validation of Friedland and Keinan (1992) with many important distinctions. First, this experiment explores the three-phased stress training approach in an aviation context using an operational task and operational performance measures. This experiment also considers longer stress exposures (approximately 10 minutes) which are more congruent with real-world acute stressors found in an aviation context. During this experiment, human physiological and subjective measures of stress are used to determine responses to stress (i.e., strain). Finally, this experiment uses a VE as a vehicle for testing this research question.

B. METHODOLOGY AND EXPERIMENT DESIGN

Equipment

The tasks in this experiment were conducted using a desktop PC-based flight simulator (see Figure 3.1).

Figure 3.1 Research apparatus

The participants' virtual environment consisted of a desktop computer, a visual display, a flight yoke control input, and an audio headset. In addition, the experimenter workstation had a second display of the flight environment including other real-time data (i.e., audio recordings, physiology, etc.) A schematic of this equipment is presented in Figure 3.2 with a description of each component following.

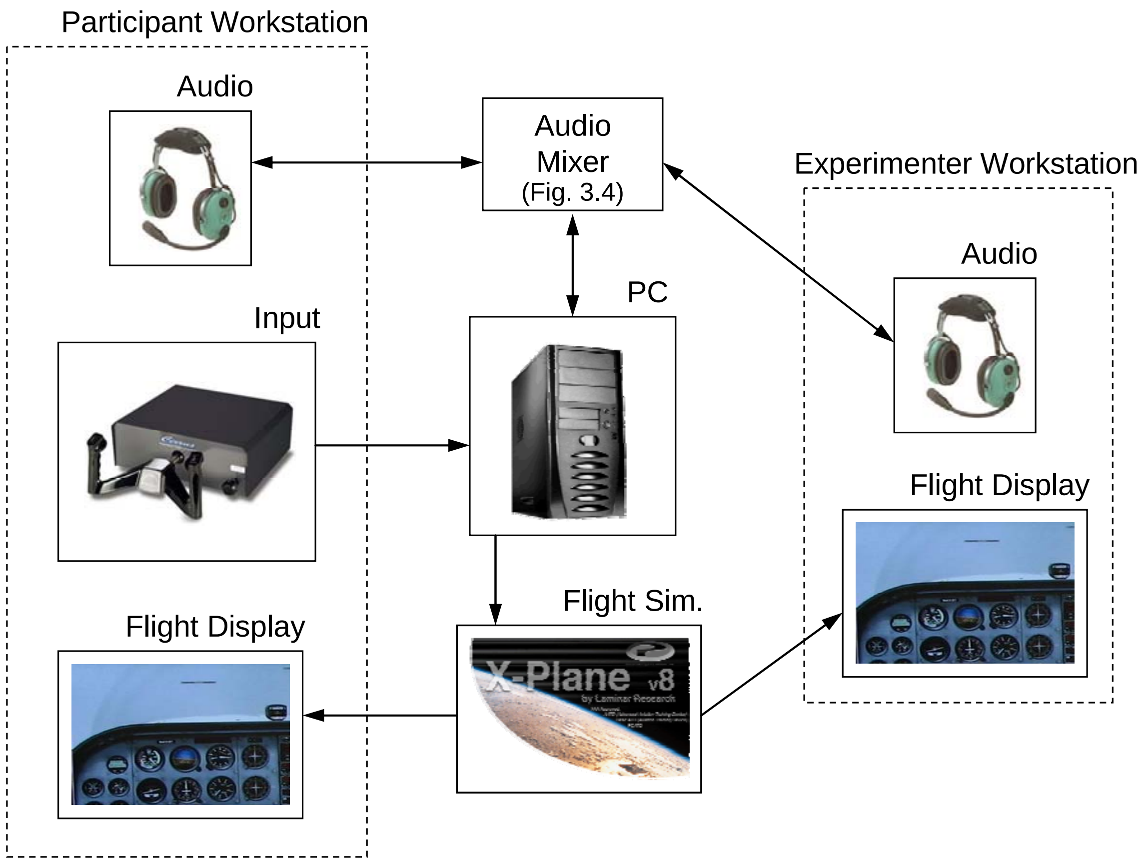


Figure 3.2 Flight simulator equipment

The Precision Flight Instruments Cirrus yoke was the only participant control input for the simulator. The rudder (yaw) controls were coupled to the yoke and the throttle was set at a constant setting by the experimenter. This task simplification was to limit the amount of training required before testing the applicable research questions.

The X-Plane® v8.0 flight simulator ran on a Monarch Computing Systems Antec PC. The X-Plane® simulator was chosen for its ability to customize the flight display, create unique scenarios, and easily customize the data output stream. All tests used a Cessna 172 flight model and instrument panel (see Figure 3.3). The simulator outside view was replaced with instrument meteorological conditions (IMC; i.e., visibility less than one nautical mile) so that only instruments were used for the flying tasks.

Figure 3.3 Simulator display

Simulator audio output consisted of the simulator audio feedback (i.e., engine and wind noise), air traffic control (ATC) instructions (provided by the experimenter), ATC "chatter" (described below), and experimenter/participant voice (see Figure 3.4). These sources of information were synthesized using a Behringer® MX802A 12-channel audio mixing board and presented to the participant at set, calibrated sound levels via a David Clark® H10-30 aviation headset. The passive noise canceling headset and mixing board set-up allowed the audio environment to be controlled, and allowed for consistent noise levels across participants. The experimenter also heard the synthesized information, and this information was sent digitally to a PC for recording using the Audacity® audio recorder.

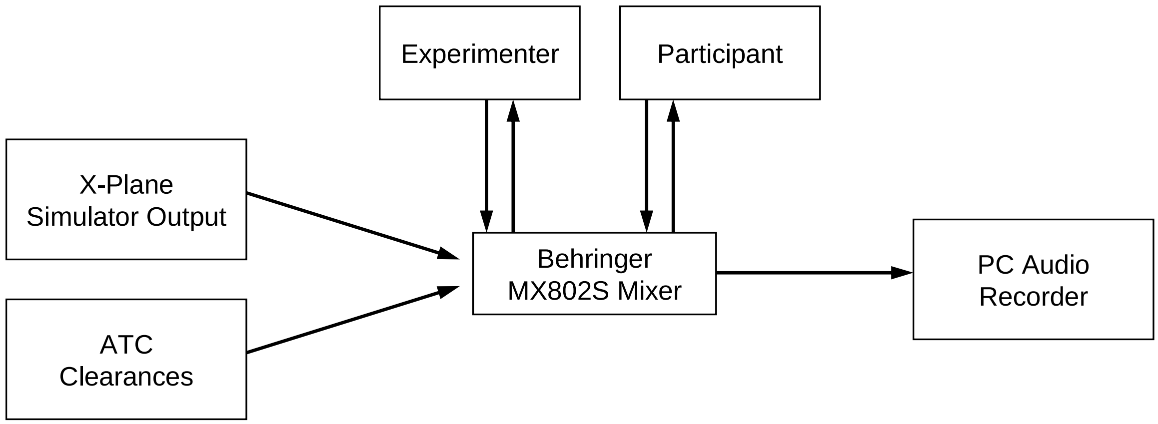


Figure 3.4 Audio information

The physiological data collection equipment consisted of the Thought Technologies Ltd. ProComp Infinity™ Biofeedback System using both electrocardiogram (EKG; see Figure 3.5) and Skin Conductance (SC; see Figure 3.6) sensors. The BioGraph Infinity 4.0® software was used to collect and analyze the physiological data. Skin Conductance and subsequent extrapolations from EKG data (see Chapter IV) have been shown as sensitive and reliable indicators of stress (Hoover & Muth, 2004; Perala & Sterling, 2007; Sloan, Shapiro, Bagiella, Boni, Paik, Bigger, Steinman, & Gorman, 1994).

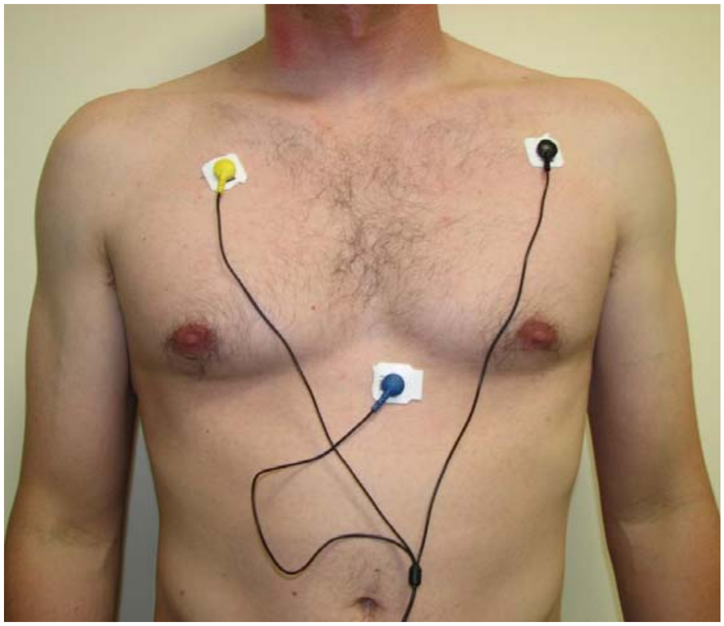


Figure 3.5 Thought Technology EKG sensors

Figure 3.6 Thought Technology SC sensor

These physiological measures were used to determine the arousal responses, or *strain*, to various stressors. The data management, conversions, and analysis of these measures is discussed below. The stressors for the pilot study and experiment are described below.

Participants

Twenty two participants were recruited for Experiment 1 from a pool of NPS students. Participants ranged in age from 25 to 39 years and all were in good health. The same survey as in the pilot study (see Appendix B) assessed recruits' qualification for participation. All participants had 20/20, or corrected to normal, vision and personal computer experience. None of the participants had flying or air traffic control experience. Similar to Friedland and Keinan (1992), participants were randomly assigned to either a stress-trained treatment group (N=11) or a control group (N=11).

Procedure

The procedures for Experiment 1 followed the schedule presented in Table 2. Participants first arrived at the MOVES Virtual Environment and Training Laboratory and heard a short welcome and description of the experiment. Next, participants completed the Informed Consent Form (see Appendix C1). If they declined participation at any time they could be excused from further participation. Participants were then given the Participant Survey, which acquired their personal information and their experience with flight simulator-related apparatus (see Appendix B). The physiological equipment was then connected and tested for accuracy.

Table 2. Training Protocol

Event	**Duration**
Administrative (Participant Survey, Informed Consent, etc.)	5 min.
Physiological Equipment Hook-up	10 min.
Verbal Instructions	10 min.
Baseline Video	5 min.
Practical Instructions	12 min.
Task Acquisition	12 min.
Stress Training (for treatment group)	5 min.
Task Acquisition 2 (control) or Performance Under Stress (treatment)	12 min.
Criterion Trial (Performance Under Stress)	12 min.
Exit Administrative	7 min.
Total	**90 min**

The participants then completed the Cohen Perceived Stress Scale (CPSS; Cohen, Kamarck, & Mermelstein, 1983; see Appendix D). The CPSS is a global measure of perceived stress and was used to further determine potential outliers due to adverse chronic life stress.

The first part of the experiment design involved verbal instructions regarding aircraft navigation. With the participant sitting at the simulator, the experimenter (a licensed pilot) provided instructions on the use of the aircraft instruments for navigation. A dialog of the verbal instructions can be found in Appendix E1. All of the instruction provided in this study was generated, in part, from Jeppesen and U.S. DoT training materials (ASA, 2008; Department of Transportation, 2001; Jeppesen Sanderson Inc., 1996, 2006a, 2006c). Following the verbal instructions, a short verbal quiz was administered to test the participant's knowledge of the provided information (see Appendix E1). If an answer was incorrect, more instruction was provided about that particular part of the task.

Following the verbal instructions, the laboratory lights were turned off and a short professional instructional film about aircraft navigation was shown to the participants (Jeppesen Sanderson Inc., 2006b). The film was presented on the same simulator screen with the simulator instruments in view for reference. The purpose of this film was two-fold: (1) the professional knowledge provided further instructed the participants on aircraft navigation while adding credibility to the training protocol, and (2) the video provided an environment conducive to relaxation in order to collect accurate baseline physiological data. As Perala and Sterling (2007) suggest, the anticipation associated with attaching physiological measures to a research participant and knowledge of imminent experiment trials can greatly skew baseline physiological data.

During the practical instructions, the participant applied the knowledge learned thus far to simulated flight in the simulator. The flight maneuvers included: level turns to a heading, climbs and descends to an altitude, climbing turns, and descending turns. Participants wore an aviation headset with microphone, and they were instructed that they must maintain criterion values of heading, bank, pitch, and altitude. A detailed account of the instructions can be found in Appendix E1.

The knowledge and practical experience given to the participants up to this point was not comprehensive (i.e., asymptote performance was not yet achieved). This ensured that learning could be observed and compared between groups in the following training and criterion sessions (see Table 2 and Figure 2.8). Following the practical instruction, the participants were asked if they had any questions.

Once the instruction phases of the pilot study were completed and the participant demonstrated practical knowledge of the test task, the training portion of the study began with the Task Acquisition phase (see Figure 3.7). During this phase, the participants communicated with a simulated air traffic controller. The controller gave the participants heading and altitude directions (i.e., *clearances*). Participants were then required to respond quickly and accurately to the commands given to them by the controller. The pilot study also was used to determine the length of time for Task Acquisition to reach mean performance criterion levels (i.e., asymptotic performance). See Appendix E1 for an account of the Task Acquisition session.

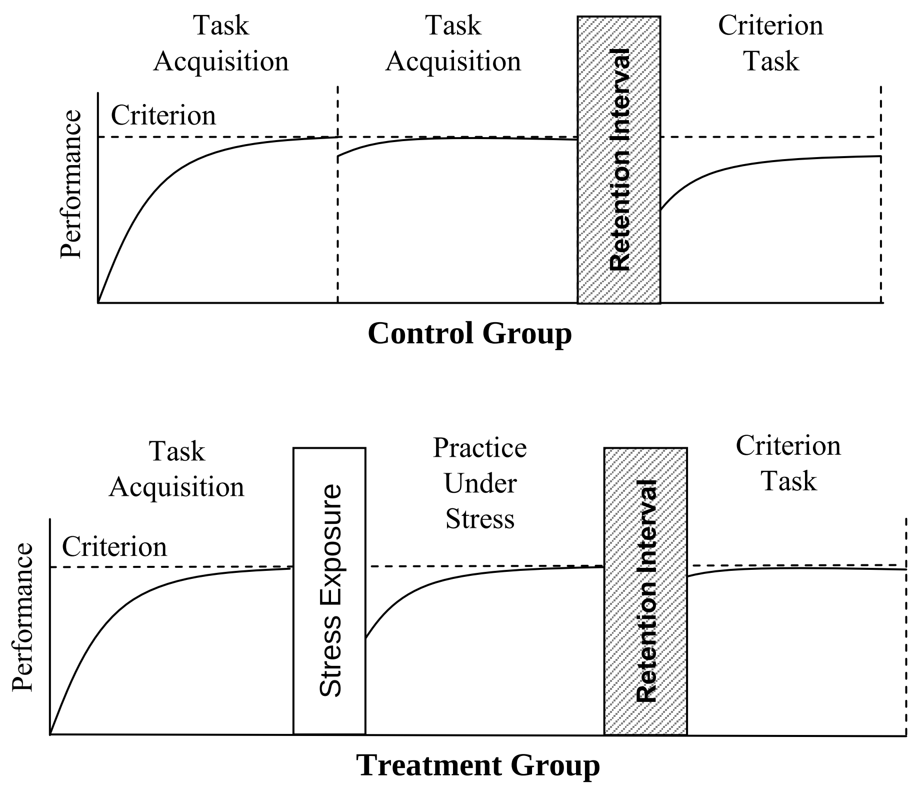


Figure 3.7 Experiment 1 design

Following the initial Task Acquisition session, the participants randomly assigned to the PUS group were given a short stress exposure protocol. During this training, the participants were first familiarized with the stressor to be used (i.e., ice water). The stressor was then applied to the participants followed by a series of relaxation techniques including controlled breathing, focused attention, and performance enhancement strategies (Meichenbaum, 1993). The participants were then given a short break.

Participants then performed either a second Task Acquisition session or the Performance Under Stress session, depending on which group the participants were assigned. This session's task was identical to the first Task Acquisition session, although the clearances (i.e., specific heading and altitude directions) were changed. Although different clearances were used for subsequent sessions, the total heading and altitude change was controlled for all sessions. The PUS group also performed this session under the cold pressor stress described above. During this session, all participants were asked every other maneuver to answer a Subjective Arousal Assessment similar to Perala and Sterling (2007, see Appendix F).

The final experiment session was the criterion trial. All participants performed the session under stress similar to the PUS session above. The only difference was the actual clearances provided.

After all sessions, the participants were thanked for their participation. They were given a short survey to ensure that they were not ill or otherwise adversely affected by the simulator or stressor (see Appendix G1).

Variables

The independent variable in this Experiment 1 was the training protocol that the participants received with two levels: stress training and no stress training. Dependent measures can be divided into two different categories: performance measures and physiological, or arousal, measures. The measures are described further in the analysis section below.

Hypotheses

As Friedland and Keinan (1992) showed, a performance enhancement was expected during stressful criterion tasks for participants who received a stress training protocol. In addition, an arousal decrement was expected for the stress-trained participants as they learned coping strategies for mitigating stress. These two hypotheses are described below.

$$H_0: \quad p_T = p_C, \quad a_T = a_C$$
$$H_1: \quad p_T > p_C, \quad a_T < a_C$$

where,

p_T: criterion performance of stress-trained treatment group

p_C: criterion performance of no-stress-training control group

a_T: criterion arousal of stress-trained treatment group

a_C: criterion arousal of no-stress-training control group

These hypotheses are reflected in Figure 3.7. This research validates stress training procedures discussed in Chapter II, and demonstrates their effectiveness and feasibility in an aviation task.

C. DATA ANALYSIS

Experiment 1 followed a between-subject design with each participant randomly assigned to either a treatment or control group. Blocking was performed on gender in an attempt to evenly distribute variance due to these factors between the two groups. Unless otherwise stated, an Analysis of Variance (ANOVA) was conducted on all performance, physiology, and subjective stress rating response measures with "condition" used as a predictor variable (IV). An alpha level of 0.05 was used to identify any significant effect of condition. The two groups were compared and found to have equal variance, therefore the ANOVA's equal variance assumption applies.

Power analyses also were conducted for all statistical tests to provide some evidence of whether the sample size was appropriate. The β values, or probability of a type-II error, are reported for each statistical test along with the test statistics and significance levels. Power values are also reported for specific insignificant results that were counter to expectations.

The following sections provide descriptions of the measures used, the collection and aggregation of the data for each of those measure, and the data analysis. In addition, an approach for determining appropriate physiological measures is presented. The methods described in this chapter were developed during the Experiment 1 pilot study.

1. Performance Measures

The aircraft trajectory data was downloaded from the simulator and analyzed according to the following procedures.

Aircraft Performance

Data was first separated into maneuvering and straight/level phases of flight. Two and a half seconds were removed from the beginning and end of each flight phase to account for transitions between phases.

Straight and level error is defined as a participant's ability to maintain the assigned trajectory based on provided clearances. For example, if a participant was asked to fly at a heading of 330 and 7200 feet, their ability to maintain that heading and altitude was determined once they arrived at that clearance. Any deviation from the assigned heading and altitude was determined to be an error.

Maneuvering error was determined in a similar manner except pitch and bank angles were used as performance measures during maneuvering portions of flight (i.e., transitions between clearances). Any deviation from 20 degrees of bank and 10 degrees of pitch during maneuvering portions of flight was determined to be error. These performance parameters were explained clearly in training and reiterated before each session (see Appendices E1 and E2). A participant may have completed the climbing/descending portion of a maneuver before the turning portion, or vice-versa. For this reason, both turning and climb/descent maneuvers were tracked separately and analyzed accordingly.

The Root Mean Square Error (RMSE) was used to calculate error during both straight/level and maneuvering segments of flight (Jagacinski & Flach, 2003; see Equation 4). The error was then analyzed for (1) each minute of task acquisition to indicate learning effects, (2) the beginning of the criterion session to determine stress training efficacy for the onset of stress, and (3) the total criterion session to indicate overall stress training efficacy.

$$X_{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(X_i - T_i\right)^2} \qquad \text{(Equation 4)}$$

Where:

X_{RMSE} = RMSE for performance variable X

X_i = the observed performance values for time $_i$

T_i = the target performance value for time $_i$

n = number observations

Figure 3.8 provides a sample depiction of straight/level error (i.e., heading and altitude) within a trial for a treatment group participant. The graph indicates (1) a decrement in performance (increased error) as difficulty increases during practice, (2) an improvement in performance during task acquisition as performance approaches asymptote (learning effect), (3) a decrement in performance during stress onset of the PUS session, (4) an improvement in performance during PUS (main effect), and (5) an improvement in the criterion trial when compared to the PUS trial (main effect).

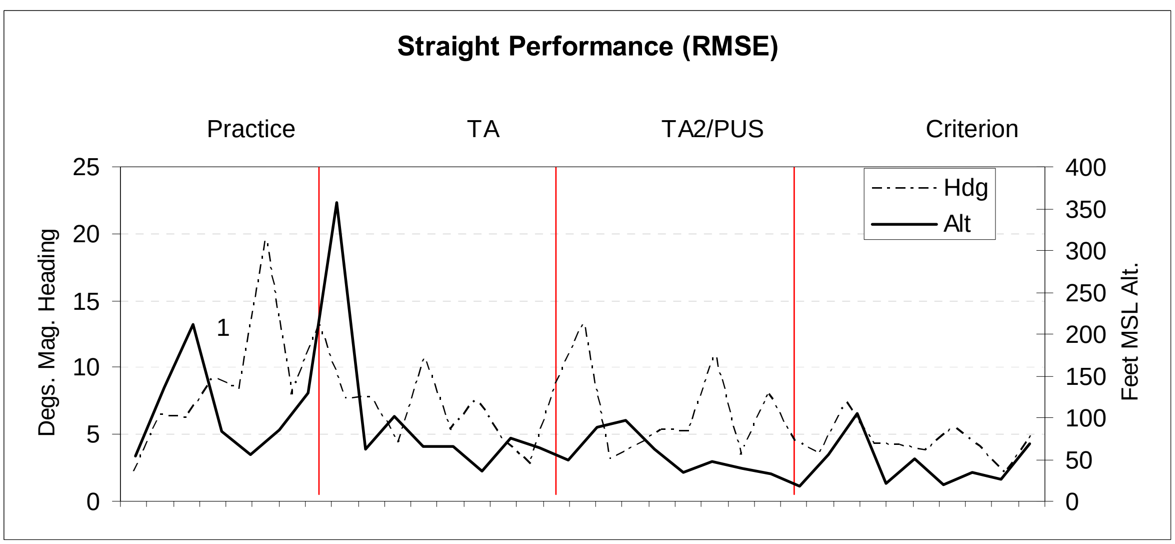


Figure 3.8 Sample performance error

Two approaches were used to determine main effects in the presence of individual differences. First, the performance measures were calculated as a function of each participant's asymptotic performance (e.g., the last three minutes of task acquisition) and total task acquisition performance. Second, step-wise regressions were performed on the performance measures to determine significant covariates in a statistical model. The following variables were added to test effects: condition, gender, age, gaming experience, computer experience, flying experience, air traffic control experience, task acquisition total performance, and task acquisition asymptote.

Transfer Effectiveness Ratio

The percent transfer and transfer effectiveness ratio were computed for all performance variables as described in Chapter II (see Equations 1 and 2). To illustrate how this was performed, data from Experiment 2 aircraft performance is provided (see Figure 3.9). To determine criterion performance, the control group mean performance was first visually inspected to determine at what time into a session the performance stabilized/optimized (1). This point was then traced forward or backward to determine at which point the treatment group's performance reached that same level (2). Finally, the time at which the treatment group reached this level of performance was determined (3).

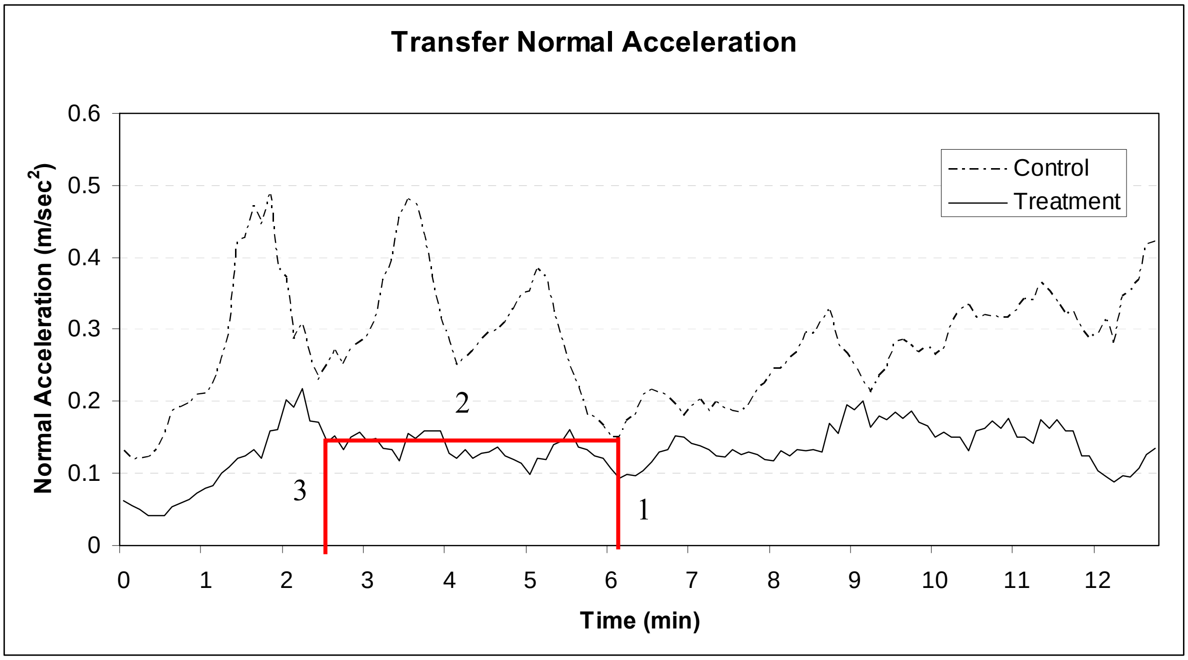


Figure 3.9 Experiment 2 transfer normal acceleration variance

In the Figure 3.9 example, the time for the control group to reach criterion performance was 6.1 minutes, and the treatment group took 2.5 minutes to reach the same level of proficiency. Substituting for Y_o and Y_x, the percent transfer can be calculated as:

$$percent\,transfer \quad = \quad \frac{6.1 - 2.5}{6.1} \; x \; 100 \quad = \quad 59\%$$

Considering that the mean time spent in the stress exposure was 4 minutes, this was the only additional training time that the treatment groups received in both Experiment 1 and 2. Therefore, the TER can be calculated as (see Equation 2):

$$TER \quad = \quad \frac{6.1 - 2.5}{4} \quad = \quad 0.9$$

These results can be interpreted to mean that 59% of the stress training the treatment group received transferred to the flight. In addition, for every hour spent in stress training, 0.9 hours are saved in the transfer task (flight).

2. Physiological Measures

The autonomic nervous system (ANS) is responsible for regulating physiological responses to external stressors and maintaining homeostasis within the body as a result of stress. The ANS works in conjunction with the endocrine system and its associated endocrine glands to secrete corticoids and adrenalin, among other chemicals, to maintain homeostasis. The ANS consists of two different branches, the parasympathetic nervous system (PNS) and the sympathetic nervous system (SNS), which collectively maintain balance within the body. In response to a stressor, increased SNS activation, often called a "fight-or-flight" response (Cannon, 1915), stimulates physiological reactions such as increased heart rate, strengthened cardiac contraction, increased breathing and lung capacity, sweating, increased temperature, pupil dilation, etc. The PNS branch of the ANS is responsible for homeostasis, or returning the body to "rest.” Specific PNS activity includes a reduction in heart rate, decreased contraction of arterial muscle, increased delay between atrial and ventricular contraction, decreased blood flow, etc. (Berntson & Cacioppo, 2004; Friedman & Thayer, 1998; Hoover & Muth, 2004; Kroemer, Kroemer, & Kroemer-Elbert, 1986; Selye, 1956, 1974)

The predictable and reliable responses of the SNS and PNS allow accurate inference about stress levels based on physiological responses. Although there are many physiological measures of stress (e.g., temperature, respiration, electroencephalograph, etc.) the most common measures are based on electrocardiogram (EKG) and electrodermal response data. These two measures and their uses in this study are described below.

Electrocardiogram

Many physiological measures of stress are derived from an electrocardiogram (EKG), or the measure of heart activity. Not unlike an electromyography (EMG) which measures the action potential of a large muscle group, an EKG measures the action potential of the largest muscle group in the human chest, the heart. There are many advantages to performing EKG-based measures of stress including their ease of use, noninvasiveness, and reproducibility (van Ravenswaaij-Arts, Kollee, Hopman, Stoelinga, & van Geijn, 1993).

An EKG trace is comprised of three primary components which are instrumental in identifying EKG-based physiology measures (see Figure 3.10). The first process in an EKG trace is the P wave, which is primarily associated with atrial activity. During this phase, blood is drawn into the heart through the atrium and then into the ventricle. This blood flow is associated with a decrease in pulmonic and aortic pressures, while the pressures within the atria and ventricles rise slightly (Fox, 1993; Kroemer, Kroemer, & Kroemer-Elbert, 1986; Ricci, 1967).

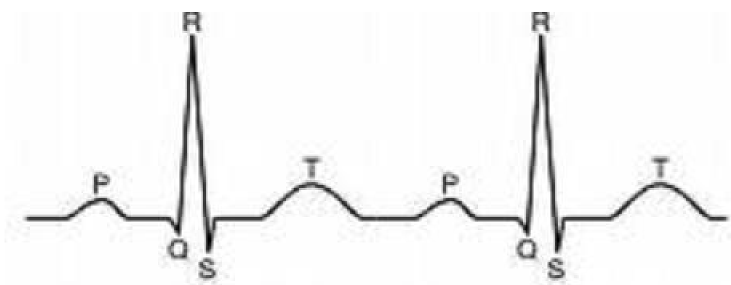


Figure 3.10 A typical EKG trace

The systolic phase follows and is a highly complex combination of vagal activity that is associated with the audible heart beat. The first *QRS process* of the systolic phase begins with an increase in intra-atrial pressure as blood is pumped from the atriums to the ventricles. Depolarization of the ventricles then occurs in conjunction with closing of the atrioventricle valves followed by an increase in ventricle pressure. Finally, the T wave represents maximum ventricular contraction and repolarization as the ventricles force blood into the aorta and pulmonary arteries (Fox, 1993; Kroemer, Kroemer, & Kroemer-Elbert, 1986; Ricci, 1967).

Following the T wave is the diastolic phase marked by decreases in aorta, pulmonary artery, and ventricle pressures as the heart returns to a resting state. The diastolic phase ends at the completion of the P wave.

Heart rate is the first general measure of vagal activity derived from EKG data and it often used as a measure of pilot workload (Hart & Hauser, 1987; Kalsbeek & Ettema, 1963; Metalis, Bieferno, & Corwin, 1989). Hasbrook and Rasmussen (1970) effectively used heart rate as a measure of stress in an aircraft instrument flying task. Their study asked experienced pilots (600-12,271 total flight hours; 60-3,057 total instrument hours) to perform ten Instrument Landing System (ILS) approaches under simulated instrument conditions. They found that pilot stress (heart rate) increased as (1) the aircraft descended closer to the ground, and, at the same time, (2) the ILS sensitivity increased. In addition, they also found lower stress levels for subsequent approaches.

Although heart rate can be measured as a rate (commonly beats per minute), the distance in time between heart beats, or inter-beat-interval (IBI), is most often used to determine the speed at which the heart is beating. Landmarks on the EKG trace must be used to mark heart beats and measure IBI. Most often, the R wave of the QRS process is used as a landmark for its ability to saliently mark a heart beat in the presence of noise. However, the P wave also is used for its computational simplicity following the relative inactivity of the diastolic phase (see Figure 3.11).

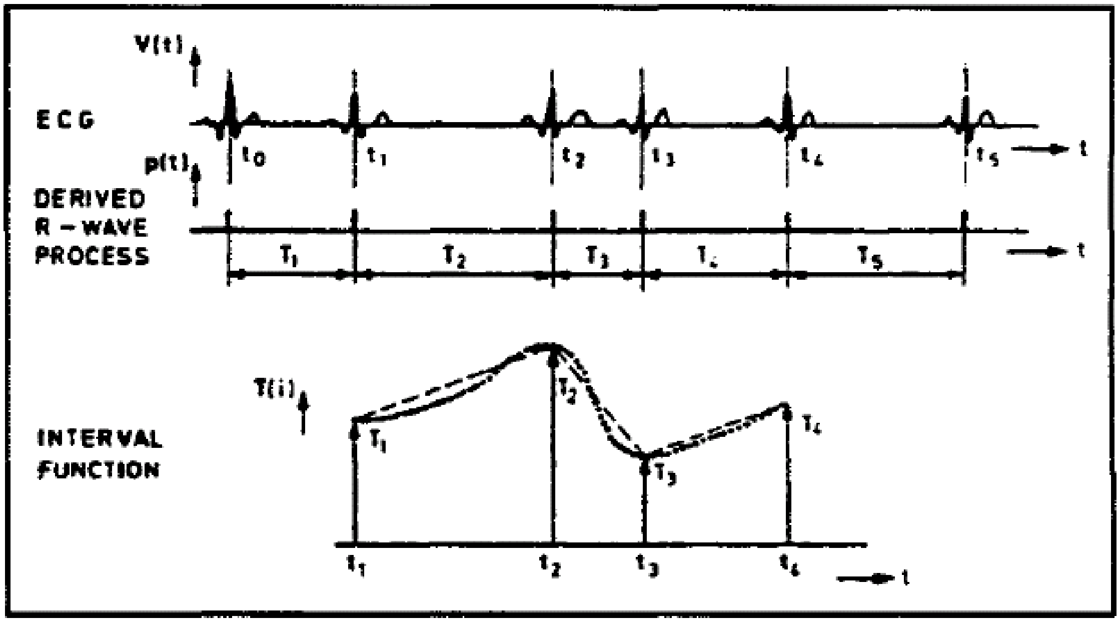


Figure 3.11 Deriving IBI from EKG
(Pfieper & Hammill, 1995)

Not only is a decrease in IBI expected during stressful events when compared to baseline or non-stressful periods, but a decrease in the variability of IBI also is expected, or heart rate variability (HRV). Heart rate variability is commonly described by "the standard deviation of intervals between successive R waves (SDRR) of the cardiac cycle" (Dishman, Nakamura, Garcia, Dunn, & Blair, 2000, p. 122), although other nomenclatures and IBI landmarks may be used (e.g., SDAN, rMSSD, pNN50, etc.)

Aside from the previously mentioned time domain, HRV may also be expressed within a frequency domain. A power spectral density (PSD) defines the frequency content of a time-based stochastic process by performing a Fourier transformation of the time domain data:

$$F(v) = \int f(t)e^{-2\pi ivt}dt \qquad \text{(Equation 5)}$$

where:

t = time

v = real number *i* in ordinary frequency (Hz).

"The main advantage of spectral analysis of signals is the possibility to study their frequency-specific oscillations" (van Ravenswaaij-Arts, Kollee, Hopman, Stoelinga, & van Geijn, 1993, p. 437). Power spectral density analysis provides the ability to distinguish between different frequency spectra and associated *types* of vagal activity, or autonomic heart rate modulation (Dishman, Nakamura, Garcia, Dunn, & Blair, 2000). The physiological basis for the high frequency spectrum (HF; 0.15-0.5Hz), or parasympathetically mediated respiratory sinus arrhythmia (RSA), "is well known and many studies have validated the use of various RSA measures as indexes of PNS activity" (Hoover & Muth, 2004, p. 198; Pfieper & Hammill, 1995). The low frequency spectrum (LF; 0.05-0.15 Hz) is derived from both parasympathetic and sympathetic activity, and is "believed to be vagally controlled" (Dishman, Nakamura, Garcia, Dunn, & Blair, 2000, p. 122). Finally, the very low frequency (VLF; 0.0033-0.05 Hz) is less understood but is believed to be indicative of sympathetic activity. Physiological responses to stress often include an increase in LF power, a decrease in HF power, and an increase in the LF/HF ratio (Berntson & Cacioppo, 2004; Pfieper & Hammill, 1995; see Figure 3.12).

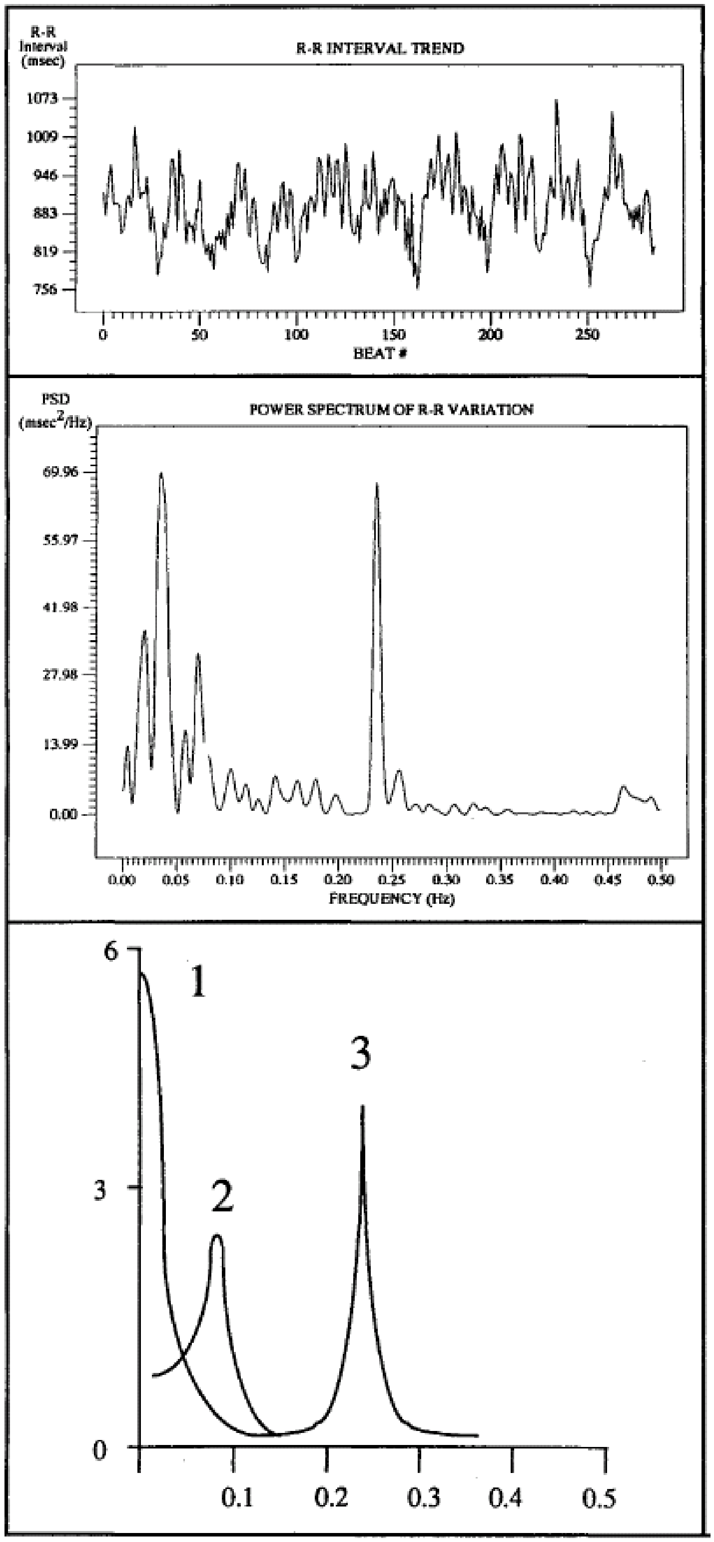


Figure 3.12 Deriving VLF (1), LF (2), and HF (3) PSD from IBI
(Pfieper & Hammill, 1995)

One common difficulty with collecting and interpreting EKG-based measures of physiology is unwanted noise, or artifact data. Artifact data can be produced by electrical interference, incorrect electrode placement, subject movement (e.g, sneeze, scratching, etc.), and tensing of abdominal muscles. During the aggregation of IBI data, artifact data can (1) introduce erroneous heartbeats, drastically decreasing the IBIs, (2) mask EKG signals, drastically increasing the IBIs, or (3) both. The removal of artifact data is accomplished by either automatic or manual means. All data for this research was filtered using a technique and computer script developed by the author. The filtering technique included three steps: (1) the data was sampled every 60 seconds, (2) the mean and standard deviation IBI of each sample were calculated, then (3) any IBI that was 3.5 standard deviations or more from the sample mean was removed.

Determination of appropriate measures

Given the prominence of many different EKG-based physiological measures, for the purposes of this research the author analyzed the efficacy of 38 different measures produced by BioGraph®. This analysis was performed during a prior pilot study and determined six measures that were (1) sensitive to mental state changes, (2) sensitive to physical stressors, and (3) reliable within/between participants within the tested simulation. The 38 measures tested are listed in Table 3 with a short description of each:

Table 3. BioGraph® EKG-based measures

Measure	Description (units)
Raw EKG	Raw EKG voltage produced by the heart (mV)
IBI	Distance between heart beats (ms)
HR from IBI	The rate of heart beats derived from IBI (bps)
HR (smoothed)	Damper filter and smoothing average of heart rate
IBI peak frequency	Moment-to-moment peak frequency of IBI using power spectrum
VLF % power LF % power HF % power	Moment-to-moment percent of total power spectrum represented by a user-defined or default constant within the spectrum
VLF total power LF total power HF total power	Moment-to-moment total amplitude of the power spectrum for a given frequency band
HR epoch mean	Average heart rate for a given period (epoch)
IBI epoch SD	IBI standard deviation for a given period (epoch)
VLF % power epoch mean LF % power epoch mean HF % power epoch mean	Mean percent of total power spectrum represented by a user-defined or default constant within the spectrum for a given period (epoch)
VLF total power epoch mean LF total power epoch mean HF total power epoch mean	Total amplitude of the power spectrum for a given frequency band for a given period (epoch)
LF/HF (epoch mean)	Average LF/HF ratio for a given period (epoch)
HR mean (beats/min)	Average heart rate in beats per minute
HR SD	Moment-to-moment heart rate standard deviation
EKG peak frequency mean	Moment-to-moment peak frequency of EKG using power spectrum
IBI SD (SDRR)	Standard deviation of IBI using R-wave landmarks
VLF % power mean LF % power mean HF % power mean	Moment-to-moment mean percent of total power spectrum represented by a user-defined or default constant within the spectrum
VLF total power mean LF total power mean HF total power mean	Moment-to-moment total amplitude of the power spectrum for a given frequency band
LF/HF (means)	Overall running average of LF/HF
IBI peak amplitude	Moment-to-moment peak IBI recorded
IBI peak amplitude maximum	Overall running maximum IBI recorded
IBI peak frequency trigger IBI peak amplitude trigger	Boolean function of maximum values recorded given user-defined or default trigger
IBI NN intervals	Time interval between consecutive normal heart beats
IBI pNN intervals	Number of instances two R-R intervals differ by user-defined or default constant
IBI pNN intervals (%)	Percent of instances which differ by above difference

Preliminary analyses compared all raw measures across trials. Figure 3.13 shows all 38 measures for a pilot study participant in the control condition (i.e., no stress training, TA and TA2 are identical, and no stress exposure until criterion). There are many difficulties when comparing raw measures in this manner. First, dissimilar units between the 38 measures make comparing the measures difficult. In addition, presenting the data as in Figure 3.13 places emphasis on larger magnitude response units as they become more salient. Finally, comparing raw measures of stress neglects the variance of each measure. A measure that has very little variance, but is very sensitive to state changes may not appear as a large spike in the figure below.

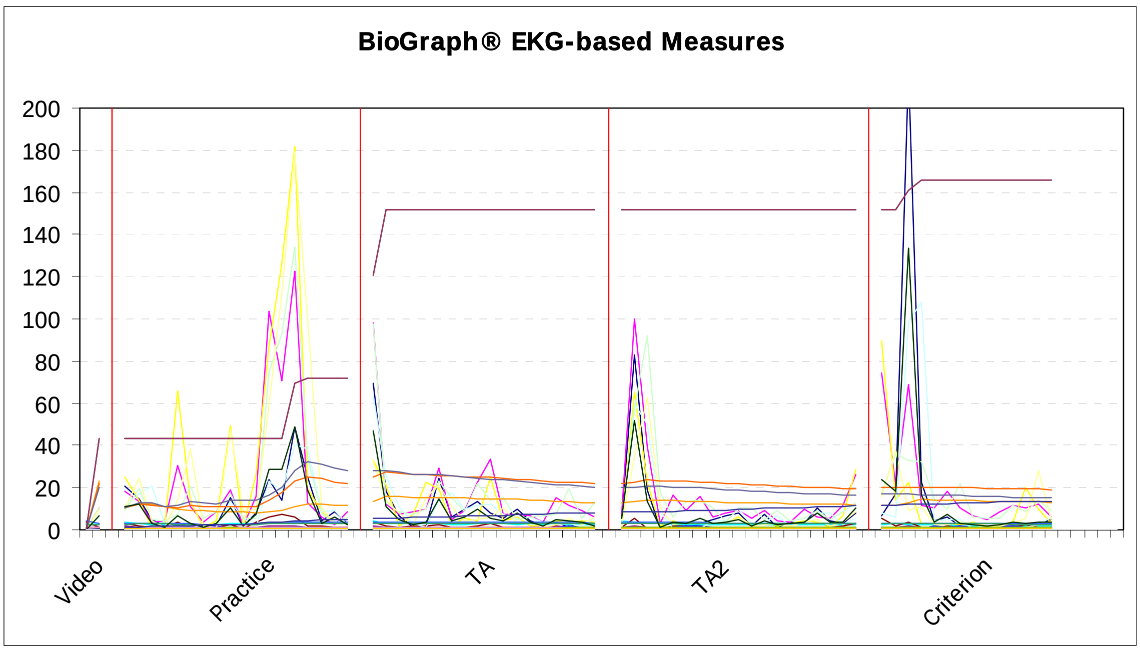


Figure 3.13 EKG-based measures for control condition (various units)

An alternative approach the author proposes to determine which of many response measures are sensitive and reliable indicators of stress is a variance/baseline-based analysis of multiple measures, or VBA. The first step in performing VBA is to compute the variance, or standard deviation, of each measure. Second, the measures are plotted as the number of standard deviations from the baseline for each segment of the trial. In this case, the baseline is the physiological response while participants were viewing the video.

Preliminary analysis determined this to be an effective measure of baseline. Each segment represents either a transition from straight/level flight to maneuvering flight or a transition between experiment sessions (e.g., Practice, TA, Criterion, etc.) The final step in the VBA examines known periods of mental and physical stress to determine sensitivity to state changes (e.g., transitions from straight/level to maneuvering, cold pressor, etc.) Figure 3.14 shows the same control participant as in Figure 3.13 after performing the VBA, and indicates the most reliable and sensitive measures resulting from the VBA analysis. Similar to the RMSE analysis above, this analysis and these variables indicated emotional state changes both between different sessions and within a session.

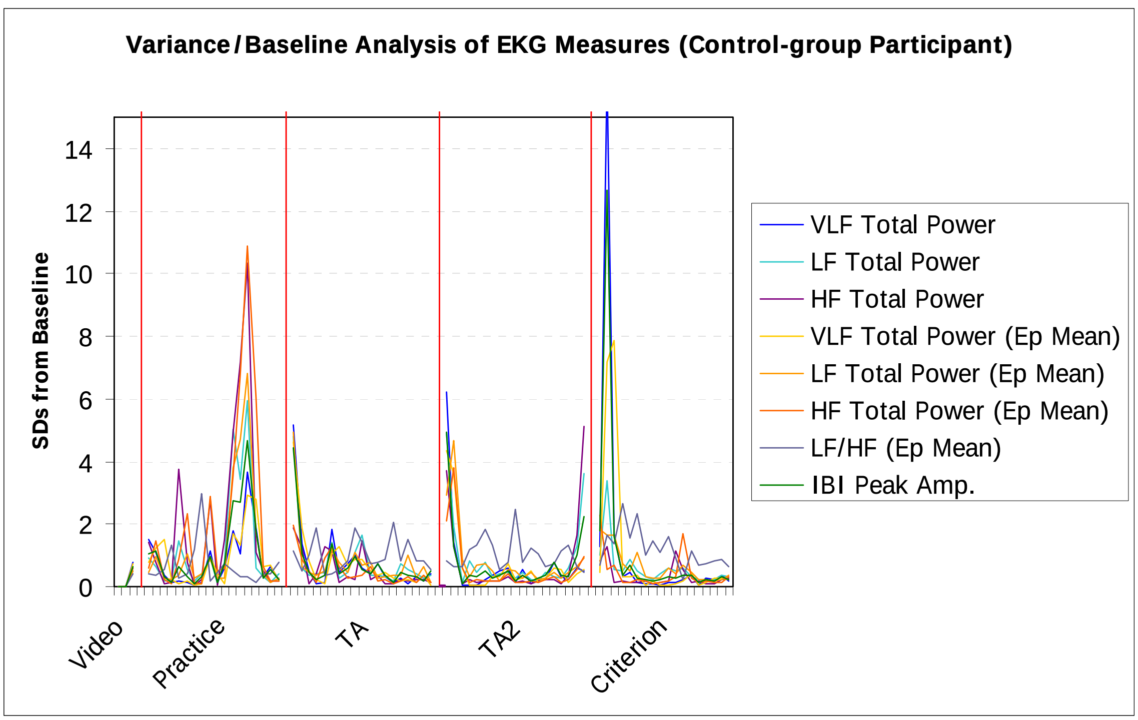


Figure 3.14 Variance/Baseline analysis

The VBA identified eight variables that were sensitive to experiment state changes using a criterion of 2 or more SDs from baseline: VLF, LF, and HF Total Power; VLF, LF, and HF Total Power Epoch Mean; and IBI Peak Amplitude. The only measure chosen that did not meet the 2 SD criteria was LF/HF Epoch Mean. However, this measure was chosen due to its high sensitivity to simulation state changes (i.e., straight/level vs maneuvering flight). These eight variables will be used as indications of stress during experiments.

Electrodermal Response

Another physiological correlate of human emotion is electrodermal response. The relationship between emotions and properties of the skin has been studied for over a century. One form of measuring electrodermal response is skin resistance, or "the electrical resistance of the skin to the flow of electromotive current and it is measured in ohms" (Grossman, 1967, p. 504). The reciprocal of skin resistance, skin conductance (SC; measured in micro-Siemens, μS) is "a method of capturing the autonomic nerve response as a parameter of the sweat gland function" (Perala & Sterling, 2007, p. 2), and, thus, allows inference regarding emotions. Galvanic skin response and psychogalvanic reflex are other terminology for SC (Grossman, 1967).

Unlike EKG data, SC is a relative measure, meaning only an increase or decrease in individual SC relative to a baseline can indicate a heightened level of emotional arousal. In addition, "although there is a relationship between sympathetic activity and emotional arousal, determining the specific emotion being elicited is difficult" (Perala & Sterling, 2007, p. 2)—anxiety, fear, anger, and sexual feelings all elicit the same SC response. Sample SC recordings can be found in Chapter V.

Subjective Queries

As mentioned in Chapter II (Section C.1), stress is an individual's appraisal of a situation given the resources that person perceives, and subjective measures of stress have been shown to indicate changes in stress levels (Johnston & Cannon-Bowers, 1996; Sloan, Shapiro, Bagiella, Boni, Paik, Bigger, Steinman, & Gorman, 1994). Subjective queries were analyzed in an attempt to portray this information. Similar to performance data analysis described above, subjective queries were aggregated for each session to determine main effects, and data was analyzed within sessions to determine learning effects. Finally, a linear regression determined the relationships between subjective and physiological measures.

D. RESULTS

This section provides the results obtained from Experiment 1 using the data analysis described above. Chapter VI further describes the applicability of these results to the research objectives.

Two outliers were removed from the following results. The first outlier's time to complete the task practice was 4.4 standard deviations greater than the mean completion time implying a general misunderstanding of the task, a disproportionate amount of practice, or both. The second outlier exhibited very erratic performance and physiological recordings as he continually asked what the manipulation in the study was, such as, "What are you changing back there?" This participant's performance was often 2 or more standard deviations from the mean. Following the exclusion of these two participants, the total sample size was 20 and there was an equal distribution of participants to the control and treatment conditions.

The 20 participants included 16 males and 4 females, and their age ranged from 25 to 39 years old with a mean age of 32.5. In addition, mean ratings for gaming experience, personal computer experience, flying experience, and air traffic control experience were 2.6, 4.6, 1.3, and 1.1, respectively (see Appendix B).

1. Performance Measures

Individual differences are accounted for using both a measure-based technique and a statistical model approach. The results of these two methods are presented below. In addition, training effectiveness ratios are computed in this section.

For all performance variables (altitude, heading, pitch and roll), no significant difference was found between the control and treatment groups for practice and task acquisition performance ($p>0.05$). In addition, when comparing the final minutes of the third (TA2/PUS) and fourth (Cr) sessions the treatment group had a decrease in error and the control group had an increase in error for every variable. This difference in performance indicates general anecdotal evidence of the stress training's efficacy, and is represented by theoretical performance curves in Figure 3.7.

Measures Analysis

The performance data was analyzed, first, comparing each variable separately. Next, the straight and level variables (altitude and heading) and the maneuvering variables (pitch and roll) were combined to draw inference concerning these two portions of flight. Finally, all data was combined to indicate main effects.

Separate Variable Analyses

The following measures were used to compare control and treatment groups for each of the four dependent variables (altitude, heading, pitch and roll):

Cr/TA: Mean criterion performance divided by mean task acquisition performance

Cr/TAend: Mean criterion performance divided by task acquisition asymptote

CrBeg/TAend: Early criterion performance (stress onset, the first three minutes) divided by task acquisition asymptote

CrBeg-TAend: The difference between task acquisition asymptote and early criterion performance

Any results found using the Cr/TA measure were expected to be compelling evidence of stress training efficacy, as the variance of the entire session (TA and Cr) is included in this calculation. For example, if stress training is only effective for improving performance during the onset of stress (CrBeg), then the converging performance of the control and treatment groups may null these differences in a total Cr-based measure. Likewise, other individual differences within the task acquisition session (e.g., initial performance, time to reach asymptote, slope of the learning curve, etc.) are all incorporated into the mean task acquisition measure.

Using Cr/TA as a measure of performance, all four variables indicated higher control group error compared to the treatment group. The difference between groups was significant when measuring altitude ($F(1,18)=4.90$, $p=0.04$, $\beta=0.45$), while heading ($\beta=0.65$), pitch ($\beta=0.76$), and roll ($\beta=0.88$) were all found insignificant (see Figure 3.15). These results provide compelling evidence that throughout the entire criterion session altitude performance is enhanced with stress training.

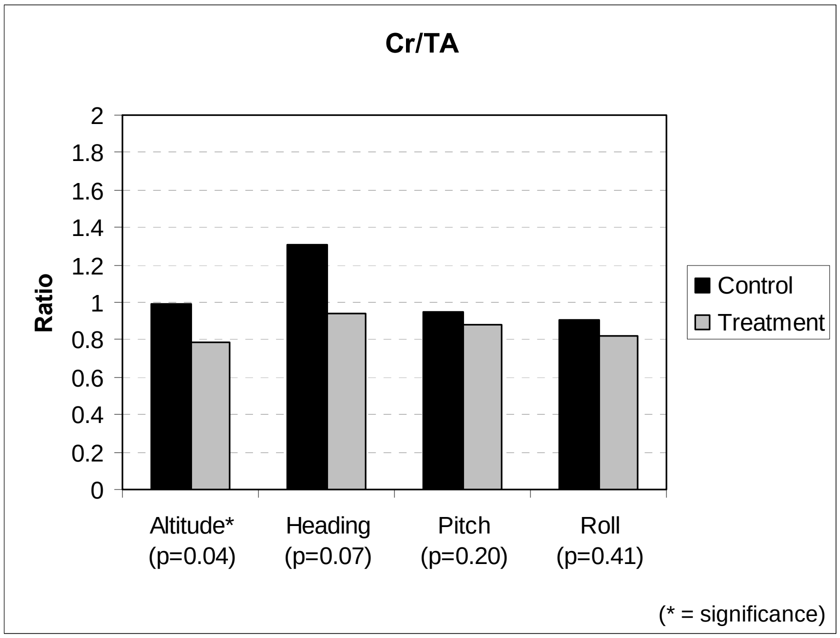


Figure 3.15 Cr/TA analysis

The Cr/TAend measure removes some of the variance from the task acquisition measure, but showed similar results for both altitude ($p=0.05$, $\beta=0.48$) and heading ($p=0.07$, $\beta=0.66$). However, the pitch ($p=0.09$, $\beta=0.62$) and roll ($p=0.08$, $\beta=0.59$) p-values, though still insignificant, decreased substantially. Again, the variable trends were all as expected (see Figure 3.16).

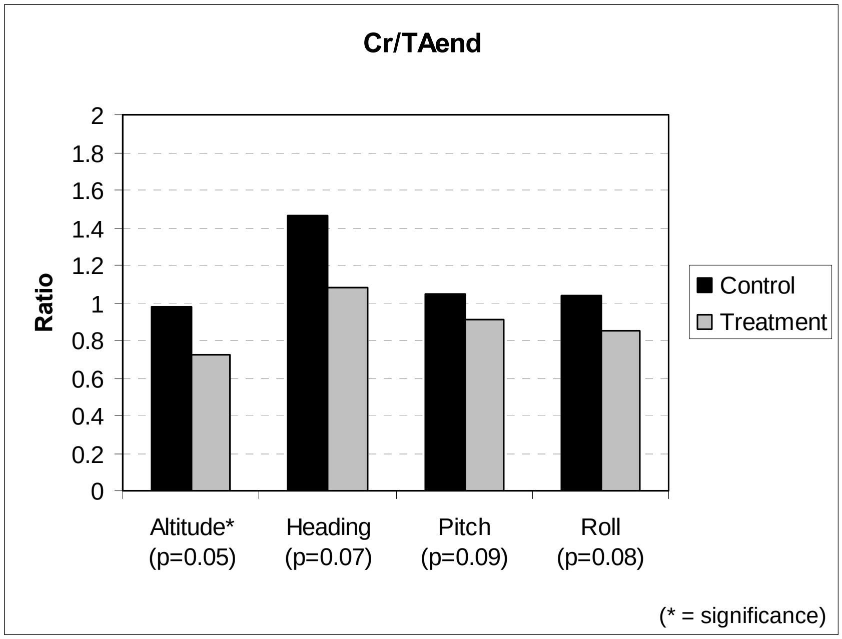


Figure 3.16 Cr/TAend analysis

The CrBeg/TAend tests the efficacy of stress training to improve performance for the onset of stress—this measure only indicates if stress training improved a participant's initial performance in the early minutes of the Cr session. The treatment group again indicated lower error compared to the control group for all variables with a significant difference when measuring pitch ($F(1,18)=7.04$, $p=0.016$, $\beta=0.29$). Altitude ($\beta=0.63$), heading ($\beta=0.50$), and roll ($\beta=0.72$) were all found insignificant (see Figure 3.17).

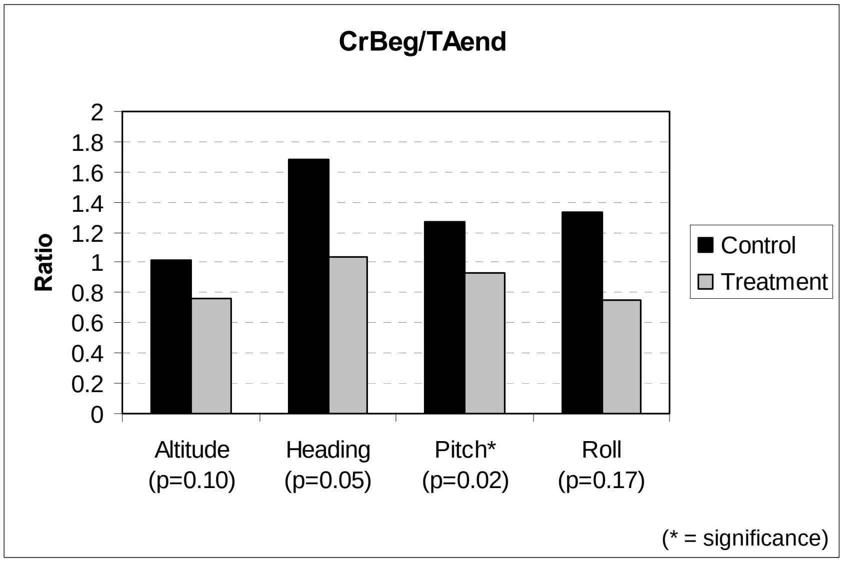


Figure 3.17 CrBeg/TAend analysis

Finally, using CrBeg-TAend as a measure, there was a significant difference between the control and treatment groups when measuring both heading ($F(1,18)=5.01$, $p=0.038$, $\beta=0.44$) and pitch ($F(1,18)=7.69$, $p=0.012$, $\beta=0.25$). Altitude ($\beta=0.68$) and roll ($\beta=0.75$) were found insignificant, although all variables showed differences in the expected direction. The two groups' altitude was not significantly different given the substantial trend data seen in Figure 3.18. Although the two groups appear to have much different mean data, the overall mean (-7.6 feet) and standard deviation (19.6 feet) of altitude for this measure indicates a very high variance in the group as a whole.

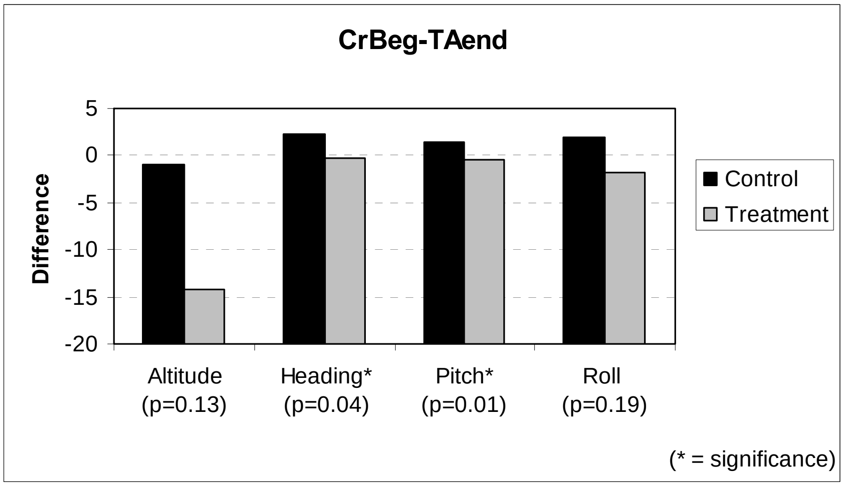


Figure 3.18 CrBeg-TAend analysis

Combined Variable Analyses

The straight/level (altitude and heading) and maneuvering (pitch and roll) variables were then combined using the below Euclidean transformation of the data:

$$c = \sqrt{x^2 + y^2} \qquad \text{(Equation 6)}$$

where:

c = combined data

x, y = straight/level (altitude, heading) or maneuvering (pitch, roll) variables.

There were no significant findings for the CrBeg/TAend and CrBeg-TAend measures ($p>0.05$). Stress training did significantly improve the performance of the treatment group for straight/level portions of flight ($F(1,18)=4.76$, $p=0.043$, $\beta=0.44$) and total performance ($F(1,18)=4.87$, $p=0.041$, $\beta=0.44$) when using the Cr/TA measure (see Figure 3.19). Stress training also improved the treatment group's performance for maneuvering portions of flight ($F(1,18)=4.43$, $p=0.049$, $\beta=0.48$) and total performance ($F(1,18)=4.87$, $p=0.041$, $\beta=0.44$) when using the Cr/TAend measure (see Figure 3.20).

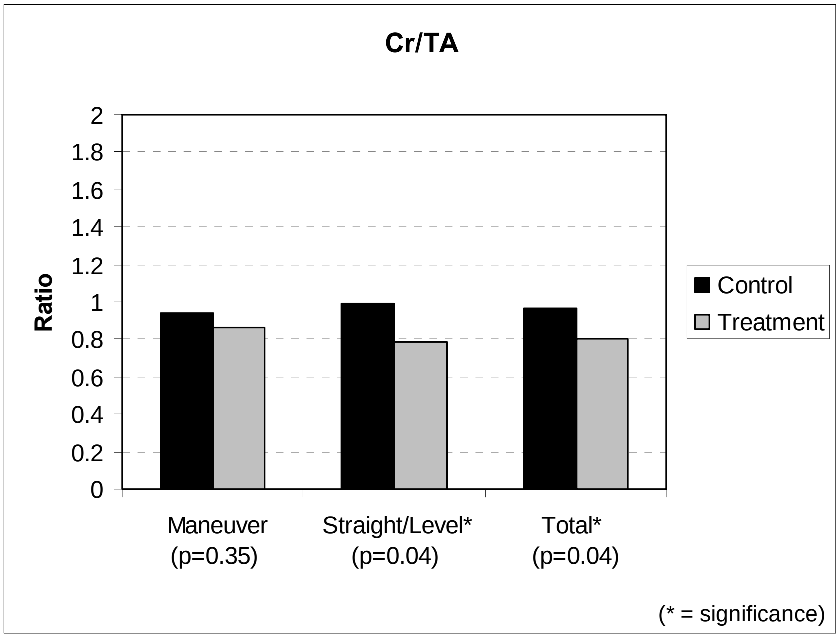


Figure 3.19 Cr/TA analysis for combined data

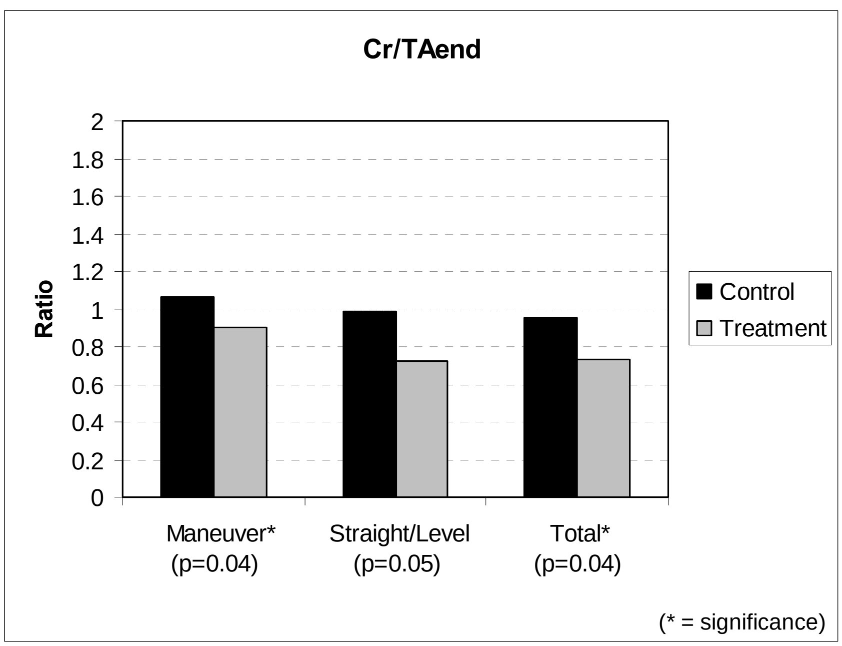


Figure 3.20 Cr/TAend analysis for combined data

Variance Analyses

In addition to the performance measurements used above, variability within each measurement was collected and analyzed. As Jagacinski and Flach (2003) point out, variability is a particularly useful measure of performance if the target is difficult to decipher or easily forgotten. Participants were required to remember target altitudes, headings, pitch, and bank in addition to ranges of acceptable performance. Although the specific performance tolerances (e.g., 20-degrees of bank, plus or minus 5 degrees; see Appendices E1 and E2) were not used for analysis, they were given to provide a performance goal. However, participants often asked to have these re-read, and it is possible that these tolerances made the target performance confusing. Given this potential confusion, performance variance was explored as a measure of performance.

A variance-based measure has been shown effective at measuring performance within a driving context (Mackie & Miller, 1978; Marcotte, Roberts, Rosenthal, Heaton, Bentley, Grant, & Group, 2003; Wylie, Shultz, Miller, Mitler, & Mackie, 1996). The method pioneered by Mackie and Miller (1978), called the "standard deviation of lane position" (SDLP), measures the amount a driver weaves within a lane.

For the purposes of this analysis, variance was used to determine how much a participant varied regardless of the assigned clearance. Variance was used, rather than standard deviation, to retain the normal distribution properties of the data for subsequent ANOVAs. Chapter VI uses Experiment 1 data to demonstrate why using a variability-based measure of performance is preferred in an aviation context compared to RMSE.

Data were sampled once every six seconds; variance was calculated for each sample, and then averaged across sessions. Six seconds was determined a sufficient sample size to allow (1) short enough samples that large heading and/or altitude changes did not adversely effect error scores, and (2) long enough that error was sensitive to the manipulation.

The control group indicated higher variance for all four dependent variables using the Cr/TAend measure. The difference in performance was statistically significant when measuring heading ($F(1,18)=11.08$, $p<0.01$, $\beta=0.22$) and pitch ($F(1,18)=5.60$, $p=0.029$, $\beta=0.39$; see Figure 3.21).

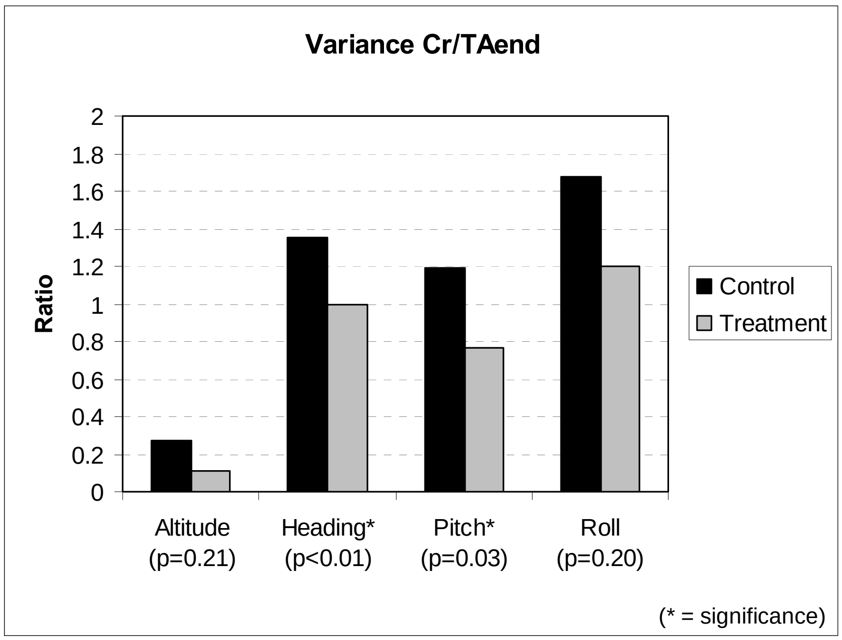


Figure 3.21 Cr/TAend variance analysis

The variance analysis also indicated significant differences between groups when measuring CrBeg/TAend. The control group showed higher variance error for both heading ($F(1,18)=5.17$, $p=0.03$, $\beta=0.42$) and pitch ($F(1,18)=4.67$, $p=0.04$, $\beta=0.47$; see Figure 3.22).

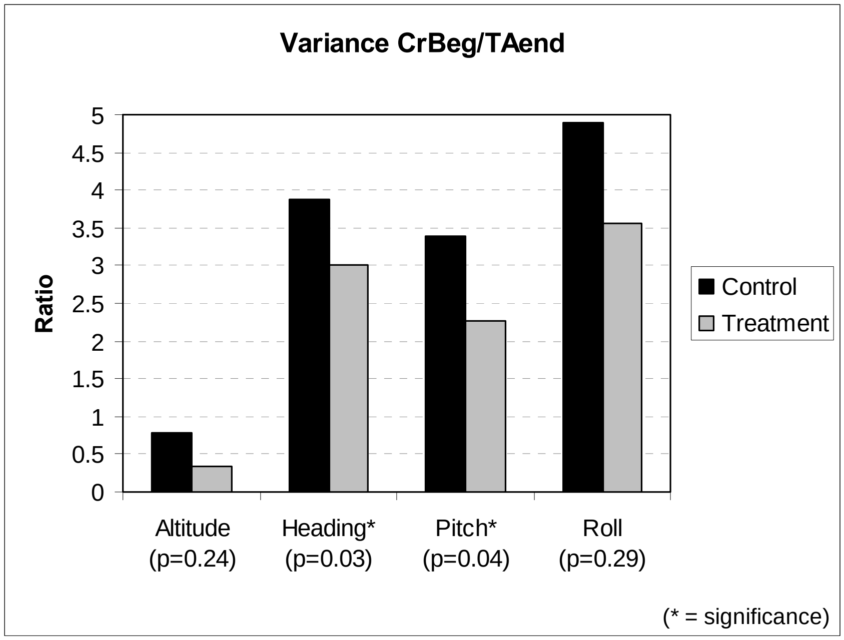


Figure 3.22 CrBeg/TAend variance analysis

The variance analysis also indicated a significant difference between groups during the criterion session (Cr and CrBeg) without normalizing the data using the TA asymptote. The treatment group showed a significantly lower variance error than the control group for both the beginning of the criterion session (CrBeg; $F(1,18)=5.33$, $p=0.03$, $\beta=0.41$) and the total criterion session (Cr; $F(1,18)=6.36$, $p=0.02$, $\beta=0.33$). This finding in the presence of individual differences suggests that the variance method may be more sensitive than traditional performance measures like RMSE in an aviation context.

Statistic Modeling Analysis

A step-wise regression was performed for each variable to determine other covariate factors other than condition. For the altitude performance variable, condition (C; $t(18)=7.04$, $p=0.016$) and mean task acquisition performance (TA; $t(18)=10.18$, $p<0.01$) explained a significant proportion of the variance in mean criterion error (Cr; $R^2=0.47$, $F(2,18)=7.85$, $p<0.01$):

$$Cr = 16.39 + 5.05C + 0.48TA + \varepsilon$$

Heading analysis determined that condition (C; $t(18)=1.90$, $p=0.074$) and asymptote performance (TAend; $t(18)=3.13$, $p<0.01$), explained a significant proportion of the variance in mean criterion performance (Cr; $R^2=0.43$, $F(2,18)=6.31$, $p<0.01$):

$$Cr = 2.06 + 0.71C + 0.65TAend + \varepsilon$$

Pitch analysis determined that condition (C; $t(18)=1.49$, $p=0.153$) and mean task acquisition performance (TA; $t(18)=5.01$, $p<0.01$) explained a significant proportion of the variance in mean criterion performance (Cr; $R^2=0.60$, $F(2,18)=13.32$, $p<0.01$):

$$Cr = 2.08 + 0.22C + 0.60TA + \varepsilon$$

Roll analysis determined that condition (C; $t(18)=1.43$, $p=0.17$) and asymptote performance (TAend; $t(18)=7.92$, $p<0.01$), explained a significant proportion of the variance in mean criterion performance (Cr; $R^2=0.79$, $F(2,18)=31.36$, $p<0.01$):

$$Cr = 0.086 + 0.54C + 0.93TAend + \varepsilon$$

The results of these models indicate that individual performance differences measured using mean task acquisition (TA) and asymptote performance (TAend) are the only contributing covariates to the dependent variable (criterion performance; Cr). Therefore, the dependent measures used in the above ANOVA analyses (Cr/TA, Cr/TAend, CrBeg/TAend, and CrBeg-TAend) are sufficient for determining main effects between groups.

Transfer Effectiveness Ratio

The variance measures were used to calculate the percent transfer and TER for each variable. Table 4 shows these results:

Table 4. Experiment 1 transfer effectiveness

Variable	**% Transfer**	**TER**
Altitude	-12.5%	-0.075
Heading	32.1%	0.23
Pitch	38.5%	0.25
Roll	81.5%	0.55

These results are consistent with Orlansky et al. (1979) meta analysis reported in Chapter II. When measuring the transfer of flying skills from flight simulators they reported a median TER of 0.48.

2. Physiological Measures

Electrocardiogram

The only significant EKG finding for the practice or task acquisition sessions was an unexpected higher control group VLF compared to the treatment group during task acquisition ($F(1,18)=7.28$, $p=0.015$, $\beta=0.27$). There were no other significant EKG findings during these practice and task acquisition sessions ($p>0.05$). There also was a surprising lack of significant difference between the groups' HRV during the third TA2/PUS session ($p>0.05$). To reiterate, during this time the control group was performing the task without the cold pressor exposure and the treatment group was performing the task under stress of the cold pressor.

When comparing overall mean session HRV, an ANOVA determined that during the criterion session, the control group had a higher mean LF PSD compared to the treatment group ($F(1,18)=7.74$, $p=0.012$, $\beta=0.25$). This significant difference in LF PSD indicates that the treatment groups' stress training significantly improved the participants' ability to mitigate the stress of the cold pressor in a stressful criterion task (see Figure 3.23).

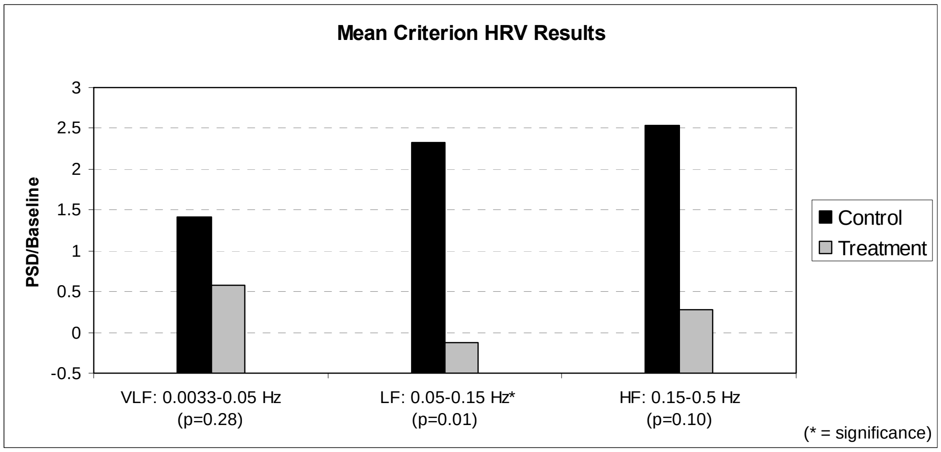


Figure 3.23 Mean criterion results for each PSD band

Given these results and prior research emphasizing the importance of LF PSD as a measure of stress response, the mean LF PSD between the two groups throughout the criterion session is presented in Figure 3.24. This figure demonstrates the treatment group's relatively quick mitigation of the criterion session stress with an abrupt drop in LF response. The treatment group's LF response then stabilizes around the same level this group reported in the baseline measurement. Conversely, the control group's LF PSD never stabilizes and indicates between one and four times this group's baseline LF response.

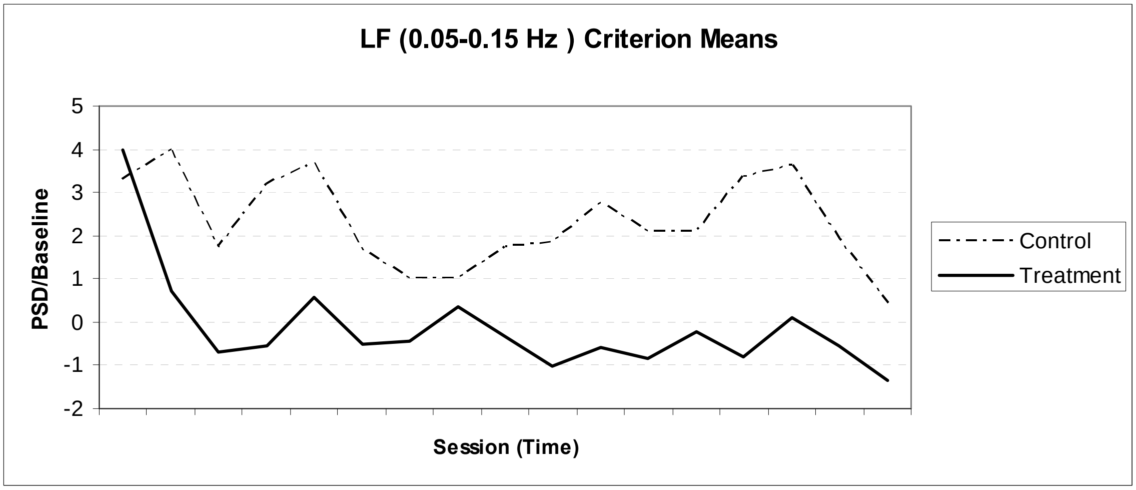


Figure 3.24 Mean LF criterion results during session

In an attempt to determine if stress training is effective for mitigating stress during both straight/level and maneuvering portions of flight these periods were analyzed separately. The treatment group had significantly lower LF PSD during both overall straight/level ($F(1,18)=6.00$, $p=0.025$, $\beta=0.36$) and maneuvering ($F(1,18)=9.57$, $p<0.01$, $\beta=0.16$) portions of the criterion session (see Figure 3.25). When each segment of flight was analyzed separately, the treatment group demonstrated lower LF PSD for six of the eight maneuvers, but only two of the eight straight and level portions ($p<0.05$). The different results for these two different portions of the flying task indicates that stress training may be more effective for decreasing strain only for portions of a task that are inherently stressful to begin with (i.e., maneuvers).

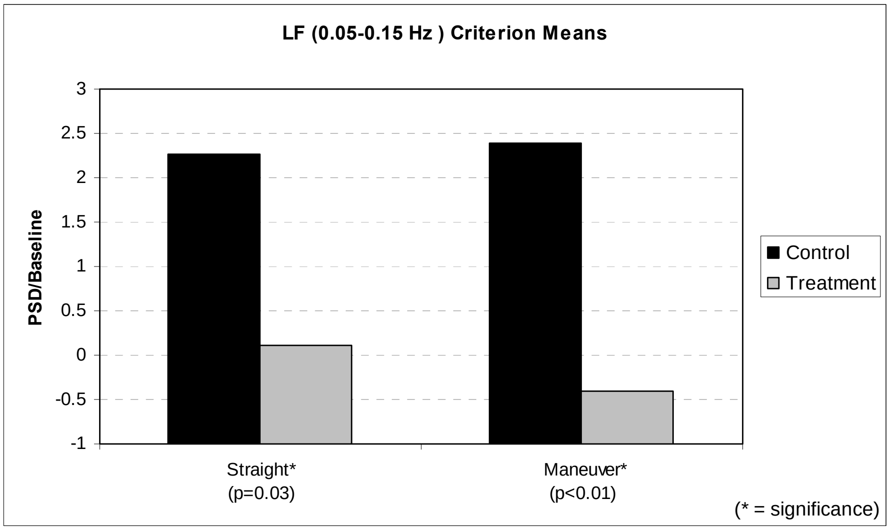


Figure 3.25 Mean LF criterion results for straight and maneuver portions

Electrodermal Response

After review of the SC data, one participant's data indicated erroneous conductance readings (~90 standard deviations from baseline). After removing this participant (N=19), trends indicated a decrease in electrodermal response for the treatment group during the TA2/PUS and criterion sessions (see Figure 3.26).

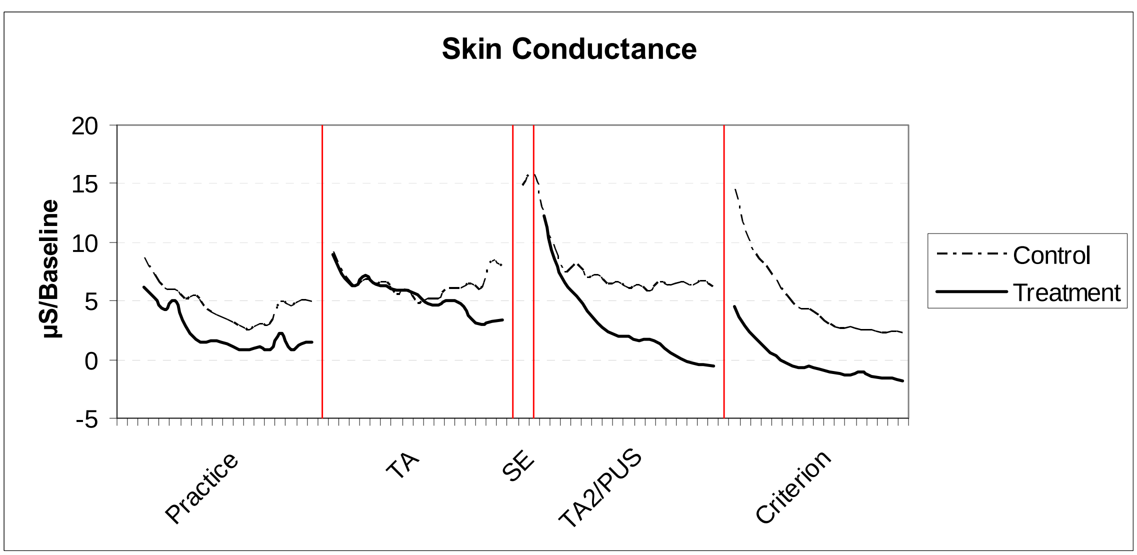


Figure 3.26 Mean SC results

Further ANOVAs indicated no significant differences for the practice, task acquisition, and TA2/PUS sessions ($p>0.05$), and a significant difference between groups for the final criterion session ($F(1,17)=4.61$, $p=0.047$, $\beta=0.48$). Further analysis of the criterion session indicated a significant difference for only the first portions of the session (see Table 5).

Table 5. Criterion SC ANOVA results

Event	**ANOVA Result**
Straight and level 1	$F(1,17)=6.62$, $p=0.02$, $\beta=0.32$
Maneuver 1	$F(1,17)=4.77$, $p=0.043$, $\beta=0.46$
Straight and level 2	$F(1,17)=4.64$, $p=0.046$, $\beta=0.47$
Maneuver 2	$F(1,17)=5.35$, $p=0.033$, $\beta=0.41$
Straight and level 3	$F(1,17)=5.49$, $p=0.032$, $\beta=0.40$
Maneuver 3	$F(1,17)=4.47$, $p=0.049$, $\beta=0.48$

Subjective Queries

Initial analysis compared subjective query responses as a function of each individual participant's baseline. Each response (five per session) ratio was analyzed separately, and mean responses for each session also were analyzed. A baseline subjective query was not collected for one of the participants, therefore this participant was not included in the subjective query analysis ($N=19$).

There was no significant difference between groups for the baseline query ($p>0.05$). During task acquisition, there was an unexpected significant difference between groups for the fourth query ($F(1,17)=7.94$, $p=0.011$, $\beta=0.24$) and the overall session mean ($F(1,20)=4.76$, $p=0.044$, $\beta=0.46$; see Figure 3.27).

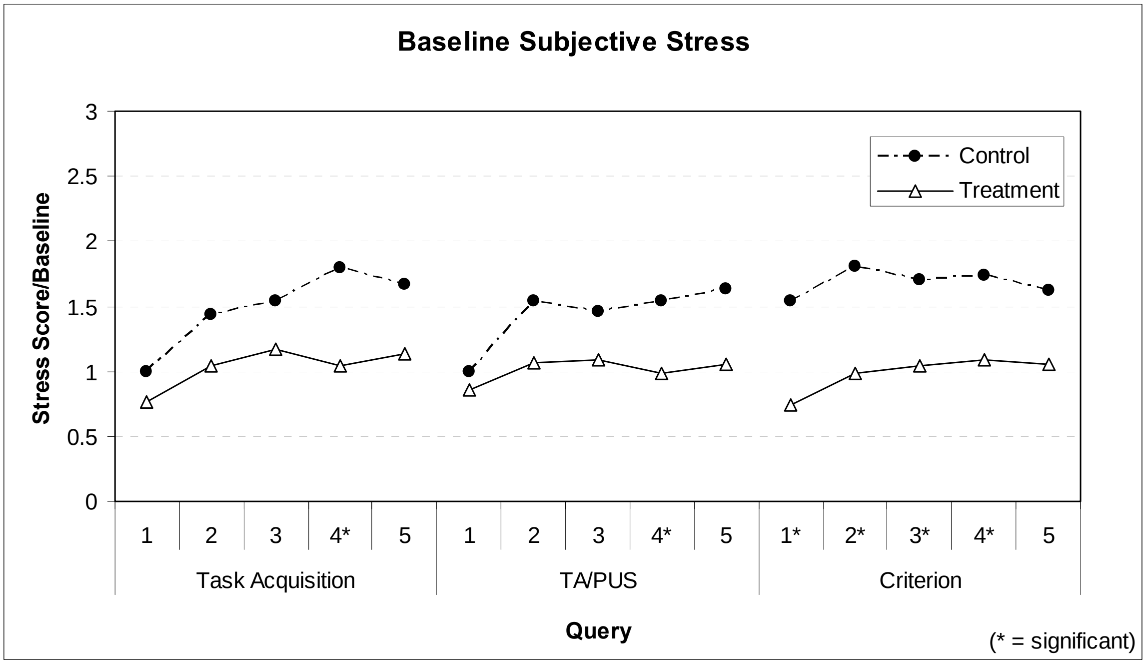


Figure 3.27 Trial subjective query results

There is no explanation for any significant differences between groups during the task acquisition session given that the two sessions, and the methods leading up to the task acquisition sessions, were identical. Similar to the performance analysis, a step-wise regression was performed to determine other covariate factors (including baseline) that may significantly influence the mean task acquisition raw stress ratings (no ratio measures). The regression determined that condition (C; $t(16)=1.64$, $p=0.12$) and baseline (B; $t(16)=5.18$, $p<0.01$) were the only covariates that explained a significant proportion of the variance in task acquisition ratings (TA; $R^2=0.63$, $F(2,16)=13.61$, $p<0.01$):

$$TA = 1.39 + 0.42C + 0.78B + \varepsilon$$

During the second test session (where control is a second task acquisition session and treatment is performance under stress), the control group unexpectedly indicated significant higher stress levels as a function of their baseline during the fourth query compared to the treatment group ($F(1,20)=5.95$, $p=0.026$, $\beta=0.37$). A step-wise regression did not determine condition to significantly contribute to the participants' responses during this session ($p>0.05$).

As expected, analysis of the final criterion session revealed significantly higher anxiety for the control group during the first ($F(1,20)=13.93$, $p<0.01$, $\beta=0.06$), second ($F(1,20)=13.13$, $p<0.01$, $\beta=0.07$), third ($F(1,20)=6.49$, $p=0.021$, $\beta=0.33$), and fourth ($F(1,20)=4.60$, $p=0.046$, $\beta=0.47$) queries in addition to the overall session average ($F(1,20)=8.00$, $p=0.011$, $\beta=0.24$; see Figure 3.27).

Again, a step-wise regression was performed to determine other covariate factors (including baseline) that may significantly influence the mean criterion raw stress ratings (no ratio measures). The regression determined that condition (C; $t(18)=2.43$, $p=0.028$) and the second test session mean (M; $t(18)=7.85$, $p<0.01$) explained a significant proportion of the variance in mean criterion ratings (Cr; $R^2=0.80$, $F(2,18)=44.32$, $p<0.01$):

$$Cr = 0.28 + 0.51C + 1.00M + \varepsilon$$

Correlations Analysis

A multivariate analysis was performed between the subjective queries and both EKG and SC. The analysis determined relatively low correlations between subjective scores and physiological measures (see Table 6). Lack of high correlation values is not surprising given the different insight subjective and physiological measures provide—subjective queries indicate a participant's appraisal of a situation given their available resources (i.e., threat), where physiological measures provide actual human responses to a stressful situation.

Table 6. Correlations between physiological measures and subjective queries

	EKG			
	VLF	LF	HF	SC
TA	.41	.13	.55	.07
TA2/PUS	.17	.22	.33	.02
Cr	.22	.12	.34	.28

These findings may appear surprising in light of research that found physiological measures sensitive to variations in workload (Berntson & Cacioppo, 2004; Pfieper & Hammill, 1995). However, it is important to point out that stress, as a subjective measure, is fundamentally different from workload—stress is a psychological state where workload is a more concrete construct comparing available resources to task requirements. Therefore, empirical workload findings should not always be generalized to stress. In addition, the subjective stress tool used in this research is not an empirically tested and proven tool. The tool used in this research, and subjective appraisal of stress in general, requires further investigation.

E. EXPERIMENT 1 CONCLUSIONS

The results of Experiment 1 indicate that, in an aviation context, when performing a stressful virtual environment criterion trial, stress training (1) improved performance, (2) decreased physiological responses to stress, and (3) decreased subjective appraisals of stress. In addition, stress training was more effective for more difficult portions of the flying task (i.e., maneuvers) and the onset of stress (i.e., the beginning of the criterion session). Experiment 1 validates the findings by Friedland and Keinan (1992) in an aviation context and using operational measures.

These results also support theories by Thorndike & Woodworth (1901) and Osgood (1949) reaffirming that a training program must share elements with the transfer task. Experiment 1 also empirically verifies that retention and retrieval is dependent on an individual's emotional, physiological, and mental states during both training and recall (Overton, 1964).

However, as pointed out in Chapter II, there are many oppositions to pseudo-transfer studies, and the results of Experiment 1 should be interpreted with some reservation. These results provide no definitive indication that stress training in a flight simulator will transfer to a stressful flying task. Therefore, a second experiment was devised to test the transfer and generalization of Experiment 1 results to the real world.

THIS PAGE INTENTIONALLY LEFT BLANK

IV. EXPERIMENT 2: STRESS TRAINING TRANSFER

Wars may be fought with weapons, but they are won by men.

— General George S. Patton, Jr.

A. OVERVIEW

Numerous studies have found that fight simulator skills will transfer to flight (Taylor, Lintern, Koonce, Jefferson, Kaiser, & Morrison, 1991; Taylor, Lintern, Koonce, Kunde, Tschopp, & Talleur, 1993; Smode, Hall, & Meyer, 1966; Carretta and Dunlap, 1998; Lintern, Sheppard, Parker, Yates, & Nolan, 1989; Hays, Jacobs, Prince, & Salas, 1992; Orlansky, String, and Chatelier, 1979; Bell & Waag, 1998). However, little is known regarding the transfer of emotional states from flight simulators to aircraft. Given inherent stress levels associated with real flight, Overton's (1964) state-dependent learning suggests that flight training also should be stressful.

The general adaptation syndrome (Selye, 1936) suggests that stress training may benefit pilots in a stressful flight environment even if the stress exposure in training is not congruent with flight. Generalization from incongruent stressors also relies on Driskel et al.'s (2001) distinction between structure and surface features. If the generalized biological response to stress is thought of as a structure feature, and the type of stress is thought of as a surface feature, then positive transfer should be observed between incongruent stressors.

As reviewed in Chapter II, Driskel et al. (2001) empirically demonstrated the generalization of stress training to novel stressors. Experiment 2 of this thesis expands on these findings in an aviation context, using a VE flight simulator, and measuring stress responses using subjective and physiological measures.

B. METHODOLOGY AND EXPERIMENT DESIGN

Experiment 2 used a between-subject experiment design with training (stress training vs. no stress training) as the manipulation (see Figure 3.7). However, in

Experiment 2, the criterion session was replaced with a transfer task in a flight environment. Other differences between the two experiments are noted below.

Equipment

The transfer task took place in a 2006 Piper Archer owned and operated by the Travis Air Force Base Aero Club (see Figures 4.1 and 4.2). This aircraft provided the capability to download telemetry data following each flight from the primary flight display (PFD).

Figure 4.1 Transfer aircraft

Figure 4.2 Transfer session

The same virtual environment and experiment apparatus was used as in Experiment 1. However, the simulator display was a representation of the transfer aircraft PFD—an Avidyne 10.4-inch, full color, flat panel, liquid-crystal display (LCD; see Figure 4.3). The simulated display filled the entire simulator screen with no outside view.

Figure 4.3 Avidyne PFD

Participants

Thirty-two participants were recruited from the pool of active duty military members at both the Naval Postgraduate School and the Defense Language Institute, Monterey, California. Research volunteers were randomly assigned to either a control (no stress training) or a stress-trained treatment group. The same surveys and requirements as in Experiment 1 were used. However, Experiment 1 analysis determined a potential influence of the informed consent—the control group may have expected the ensuing stress exposure, and this may have influenced physiology results. Therefore, for Experiment 2, the treatment and control group received slightly different informed consent forms. The control group, which did not have any stress exposure in training, received an informed consent absent of any reference to the cold pressor or "stress" (see Appendix C2). The treatment group received a similar informed consent to the Experiment 1 form (see Appendix C3). In addition, participants were asked to rate their susceptibility to fear of heights, fear of flying, and motion sickness (see Appendix G2). No participants were found highly susceptible to any of these conditions (greater than 5 on a 10-point scale). These responses also were used as statistical covariates (see Results).

Procedure

The Experiment 2 procedures followed the same schedule as described in Table 2 and Figure 3.7 above, and similar instructions were read (see Appendix E2). However, the criterion session was replaced by an approximately 60-minute flight including a 15-minute transfer task. All procedures prior to the flight were identical to Experiment 1, except (1) they were conducted at the Monterey Navy Flying Club (MNFC) located at the Monterey Airport, (2) a different baseline training video was used that was relevant to the Avidyne PFD (Dishman, Nakamura, Garcia, Dunn, & Blair, 2000), and (3) a "Transfer Task Prep" section was added after all training to prepare participants for the flight operations (see Appendix E2). In addition, no altitude changes were used for the flying task because dry-runs determined this to be too difficult for a novice in conjunction with turns in a real flight environment. Therefore, all procedures in Experiment 2 tested level turns only. Following the Task Acquisition 2 (control)/Practice Under Stress (treatment) session, the participants were escorted to the transfer task aircraft. Figure 4.4 and 4.5 show the participant and experimenter simulator consoles, respectively.

Figure 4.4 Participant console

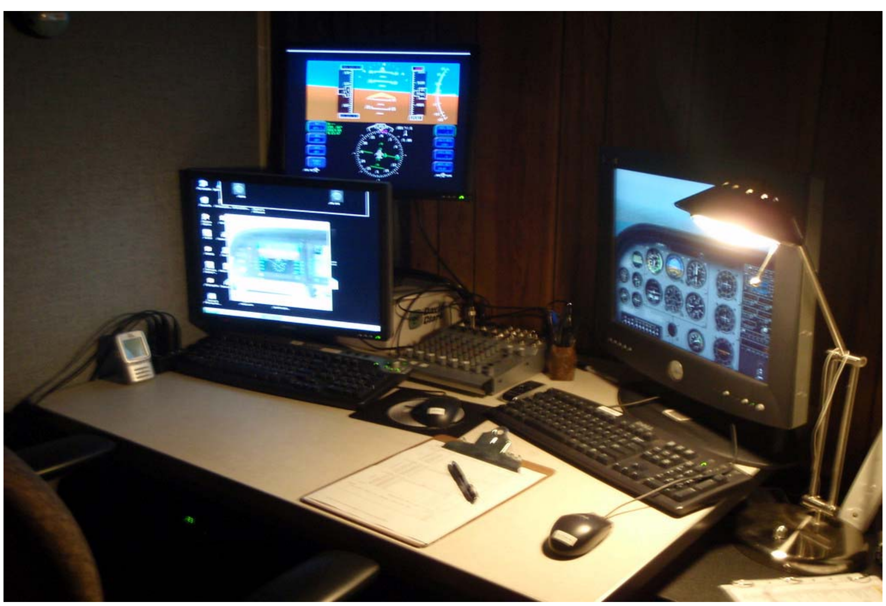


Figure 4.5 Experimenter console

Prior to the participant's arrival at the aircraft, a certified pilot performed a pre-flight inspection, calculate weight and balance, and reviewed current and forecasted weather. The participant was then assisted into the front-left seat, the experimenter sat in the rear-right seat, and a certified flight instructor (CFI) sat in the front-right seat of the aircraft (see Figure 4.6). The participant was given an FAA-compliant safety briefing during aircraft taxi.

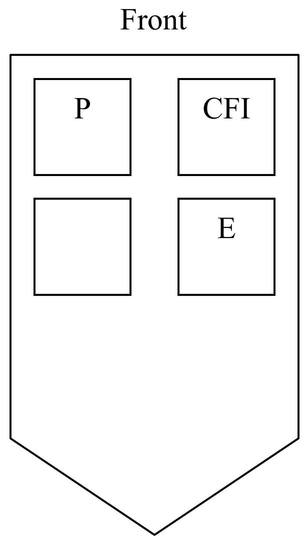


Figure 4.6 Aircraft seating arrangement
(P = participant, CFI= certified flight instructor, E = Experimenter)

The two CFIs were FAA-licensed flight instructors with instrument instructor ratings (CFII). The two CFIs combined had more than 19,000 flight hours and over 50 years of flying experience. In addition, they were both instrument instructors—they are experts at evaluating a student's flying experience during instrument maneuvers. Therefore, the CFIs were subject matter experts.

Their job was to safely fly the aircraft to and from the training area. On the short flight to the training area, the participant wore a pair of glasses that blinded the view of the controls and instruments. The blinders prevented the participant from gaining any further knowledge about the controls and displays in route to the training area. However, the outside view was still visible to prevent motion sickness and undue discomfort. Upon arrival at the training area, the participant dawned a second set of blinders that shielded the outside view and all instruments other than the PFD. These blinders ensured that only the skills learned in training (instrument flying) were used during the transfer task.

The CFI then aligned the aircraft with the starting heading and altitude. Session instructions were read by the experimenter and aircraft control was transferred to the participant. The experimenter then provided clearances following the same procedures from the VE training. In the event that an unusual attitude was obtained (greater than 60 degrees of bank, greater than 20 degrees of pitch, or less than 20 degrees of pitch), the CFI recovered the aircraft. The recovery procedure was practiced in the simulator and during taxi; this procedure is familiar to all CFIs and is a required part of their certification. Other than these duties, the CFI was not allowed to interact with or instruct the student in any way until after the transfer task was completed.

The experimenter ensured a sterile cockpit environment—prior to and during the transfer task, communications within the aircraft were limited. Only necessary communication for flight safety and for the experiment were allowed. In addition, the aircraft was equipped with split communications capability. This feature allowed the CFI to communicate with air traffic controllers in isolation from the participant and experimenter.

The experiment officially concluded after all clearances were accomplished. At this point, the CFI resumed control and returned the aircraft to the airport. Upon arrival back at the Monterey Navy Flying Club, the experimenter debriefed the participant (see Appendix G2). See Figure 4.2 for a trial in progress.

Variables

The same variables and analysis were used as in Experiment 1 with the following exceptions:

—The aircraft location and performance measures were collected using primary flight display telemetry data. Data were extracted and then analyzed similar to Experiment 1.

—During each of the in-flight subjective stress queries, the CFIs were required to rate the participants' performance on a 10-point scale (see Appendix H).

—The physiology data was recorded on a Compact Flash using the ProComp Infinity encoder (see Figure 4.7). After each trial, data was uploaded to a PC for analysis. A button switch was used to mark events during trials.

Figure 4.7 ProComp Infinity encoder with compact flash drive

—Audio recordings were used by connecting an Olympus digital voice recorder (VN-3100PC; see Figure 4.8) to the empty intercom jack (rear left seating position) in the aircraft. This data also was uploaded to a PC following each trial.

Figure 4.8 Olympus audio recorder

Data synchronization took place prior to each take-off. During this time, (1) the PFD altitude and heading bugs were set to 10,000 feet and the participant ID, respectively; (2) the audio recording was started; and (3) a unique key press was performed on the physiology button switch. This synchronization allowed all three data formats to later be compared at any specific time within a trial.

Hypotheses

The Experiment 2 hypotheses are as follows:

$$H_0: \quad p_T = p_C, \quad a_T = a_C$$
$$H_1: \quad p_T > p_C, \quad a_T < a_C$$

where,

p_T: transfer performance of stress-trained treatment group

p_C: transfer performance of no-stress-training control group

a_T: transfer arousal of stress-trained treatment group

a_C: transfer arousal of no-stress-training control group

C. DATA ANALYSIS

The same data analyses were performed as in Experiment 1. Due to a more diverse sample of participants from the Monterey military population, blocking for Experiment 2 was performed on both gender and military rank. In addition, participants were equally distributed between the two CFIs to control for the amount of variance between the two CFI's ratings. As in Experiment 1, ANOVAs with an alpha level of 0.05 were conducted on all performance, physiology, and subjective stress rating response measures with "condition" used as a predictor variable (IV).

D. RESULTS

Two outliers were removed from the Experiment 2 results. The first outlier became airsick during the flight, and the experiment was terminated approximately half way into the transfer task. Another outlier was removed due to a nine day delay between the simulator training and the transfer task. This delay was due to inclement weather following the simulator portion of the trial. These two outliers resulted in a sample size of 32 with an equal distribution of participants to both groups (16).

The 32 participants included 28 males and 4 females, and their ages ranged from 18 to 37 years old with a mean age of 27.1. In addition, mean ratings for gaming experience, personal computer experience, fear of heights, fear of flying, and motion sickness were 2.10, 4.4, 2.0, 1.5, and 1.6, respectively (see Appendix B and G2).

A preliminary analysis was performed to determine if the flight environment was indeed perceived as stressful by participants. This analysis was important for determining if stress training was viable in this context—if the flying task was not stressful then it should not benefit from stress training. For this analysis, subjective stress ratings were compared between the Experiment 1 control group criterion session and the Experiment 2 control group transfer session (see Figure 3.7). These two groups were similar other than the actual stress they were exposed to during these sessions: cold pressor versus flight.

First, an ANOVA of the first two sessions' (TA and TA/PUS) queries was performed to ensure that this was a fair comparison between groups. No significant differences were found ($p>0.05$) suggesting that the stress responses during these sessions were similar between the two experiments (see Figure 4.9 Task Acquisition and TA/PUS). Subjective stress scores during the final transfer session determined that the two groups perceived to be equally strained when exposed to their respective stressor ($p>0.05$). In other words, as reported by these participants, flying an aircraft is as stressful as putting your foot in a bucket of ice water. This analysis determined that the instrument flying task is stressful for a novice and may benefit from stress training.

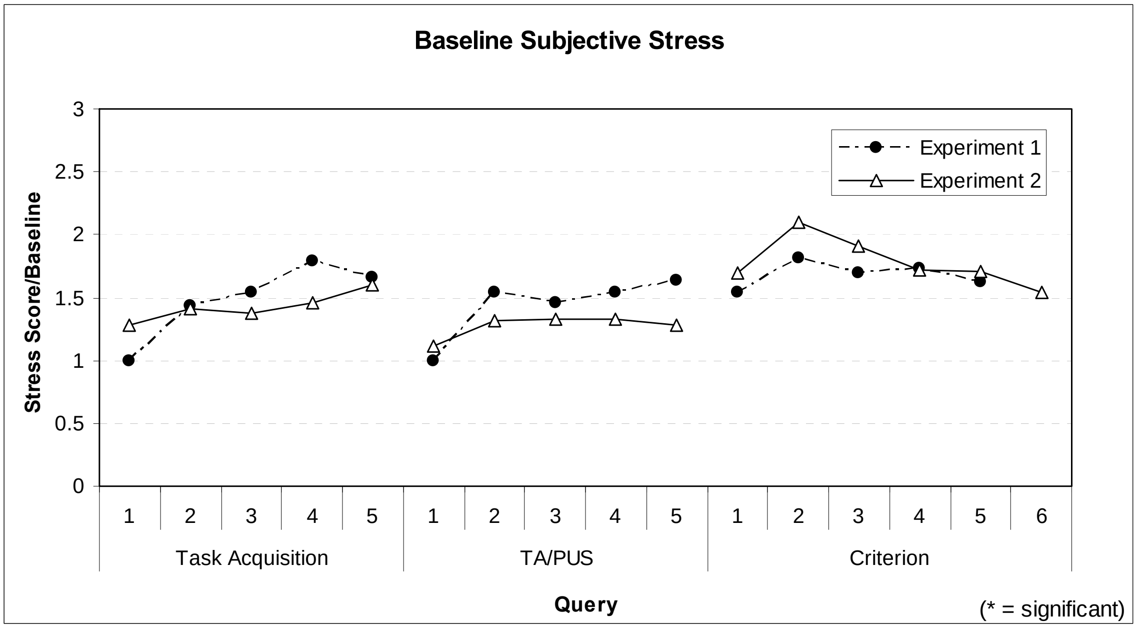


Figure 4.9 Experiments 1 and 2 subjective stress comparison

1. Performance Measures

Aircraft Telemetry Data

Given the success of the Experiment 1 variance analysis, the Experiment 2 analysis exclusively used variance as a measure of performance. (See Chapter VI for a discussion regarding the effectiveness of this measure in an aviation context.) Given the large effect sizes and performance significance ($p<0.05$) observed in the following variance analysis, measures that account for individual differences were not needed—the two groups' total mean transfer performance were compared absent of simulator performance.

There were problems with comparing the two groups' performance using both the altitude and heading measures as used in Experiment 1. In Experiment 1, the task clearances were carefully selected so that no participants traveled through a heading of due North. Computing performance measures of a flying task as the heading transitions from 359-degrees to 1-degree, or visa-versa, can present erroneous error scores. The operational constraints associated with Experiment 2's flying task prevented this control.

In addition, the experiment aircraft's telemetry data did not provide indicated altitude. Rather, it provided true altitude which varies from the indicated altitude due to temperature and barometric pressure variations. As outside temperature and pressure are constantly changing, true altitude becomes an unreliable measure of flying performance—the reading on the display (i.e., indicated altitude) can have zero variance while the true altitude is constantly changing. For these two reasons, heading and altitude were not used in this analysis.

The aircraft's telemetry data did provide some variables that allow for very accurate inferences regarding pilot performance, foremost being the acceleration and rate variables. The telemetry data included directional acceleration in the lateral (side-to-side), longitudinal (forward-to-back), and normal (up and down) aircraft axes in addition to rotational rates due to roll (roll rate) and pitch (pitch rate). These measures are direct, objective measures of how *smoothly* the aircraft is flying. Also included in this analysis are bank and pitch angles.

Using variance as a measure of performance (see above) there was not a significant difference between the control and treatment groups for the total transfer task mean using the pitch and roll variables ($p>0.05$). The trends are in the expected direction (see Figure 4.10).

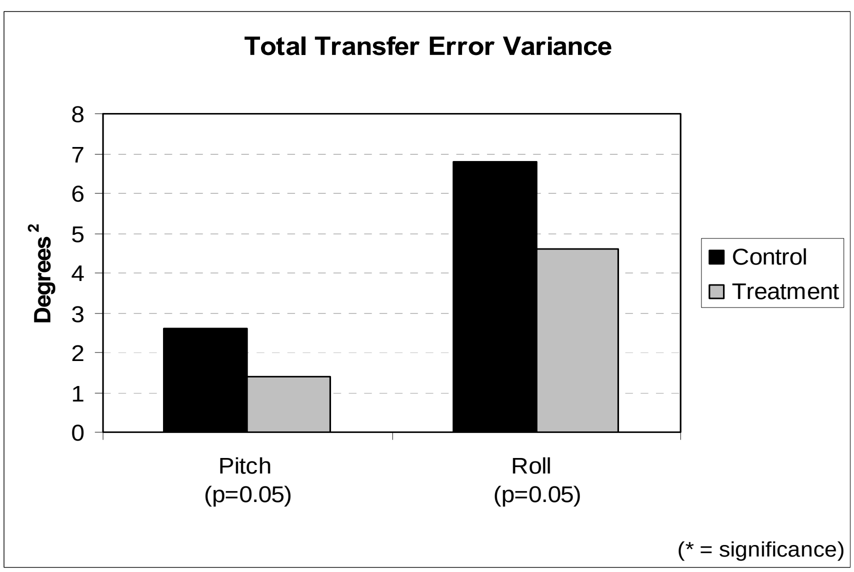


Figure 4.10 Total pitch and roll variance analysis

The variance analysis did indicate significant differences between groups when measuring the total transfer task mean acceleration variables. The control group showed higher variance error for lateral ($F(1,28)=5.52$, $p=0.026$, $\beta=0.38$), longitudinal ($F(1,28)=6.34$, $p=0.008$, $\beta=0.32$), and normal ($F(1,28)=6.18$, $p=0.019$, $\beta=0.33$) acceleration (see Figure 4.11).

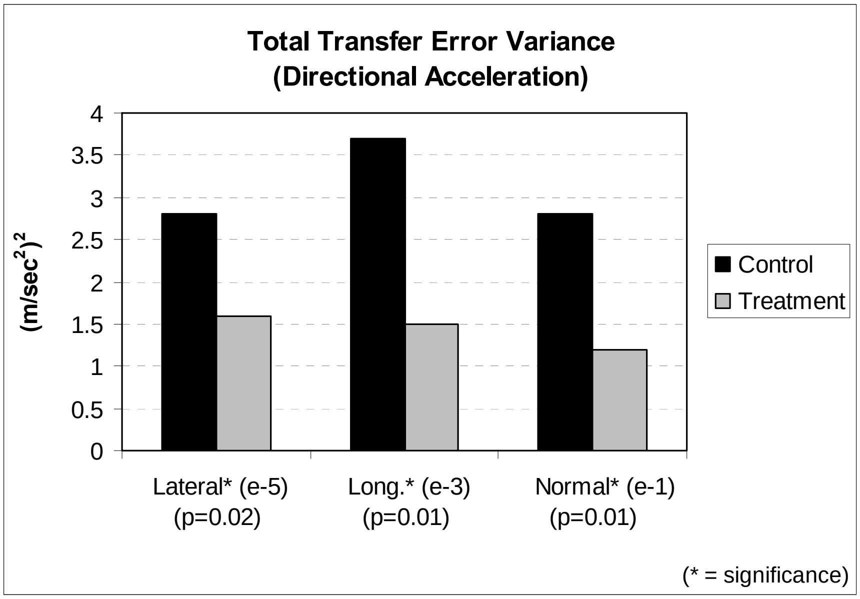


Figure 4.11 Total acceleration variance analysis

The control group also showed higher variance error for roll rate ($F(1,28)=5.42$, $p=0.027$, $\beta=0.39$) and pitch rate ($F(1,28)=5.68$, $p=0.024$, $\beta=0.37$) when computing total transfer task mean variance (see Figure 4.12).

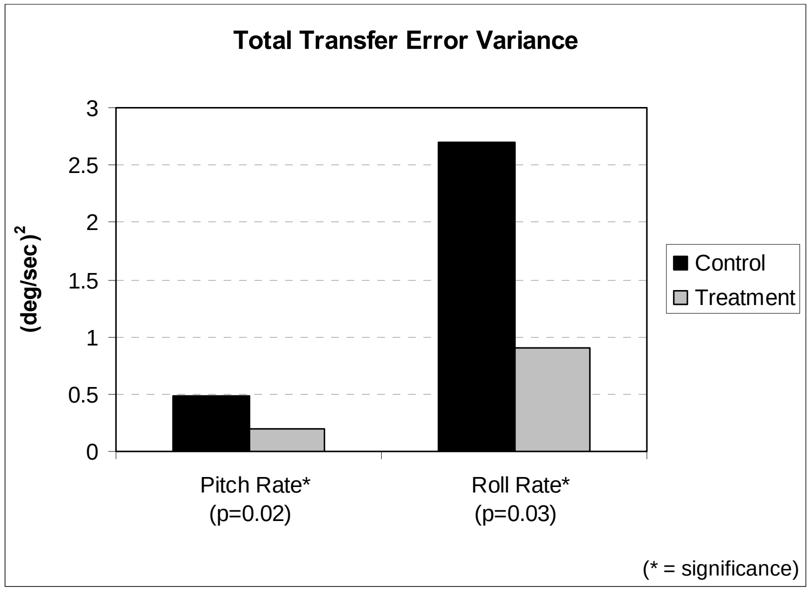


Figure 4.12 Total rate-or-rotation variance analysis

A separate variance analysis was performed to determine the efficacy of stress training to enhance performance during the onset of stress in a transfer task. To this end, the average variance was computed for the first three minutes of the transfer task. This analysis indicated a significant difference between groups using the roll variable ($F(1,28)=4.98$, $p=0.033$, $\beta=0.43$; see Figure 4.13).

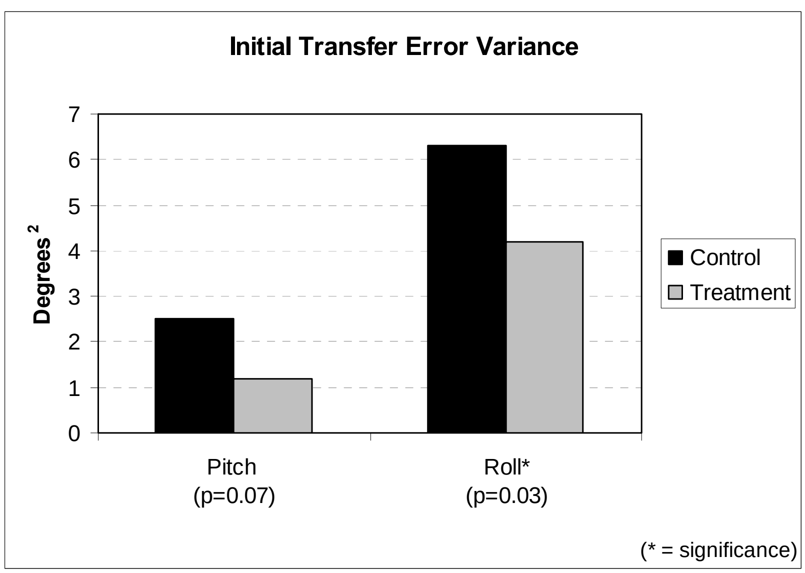


Figure 4.13 Initial pitch and roll variance Analysis

When measuring initial transfer task variance, the two groups also were significantly different using the lateral ($F(1,28)=5.46$, $p=0.027$, $\beta=0.38$), longitudinal ($F(1,28)=6.12$, $p=0.019$, $\beta=0.34$), and normal ($F(1,28)=5.25$, $p=0.029$, $\beta=0.39$) acceleration variables (see Figure 4.14).

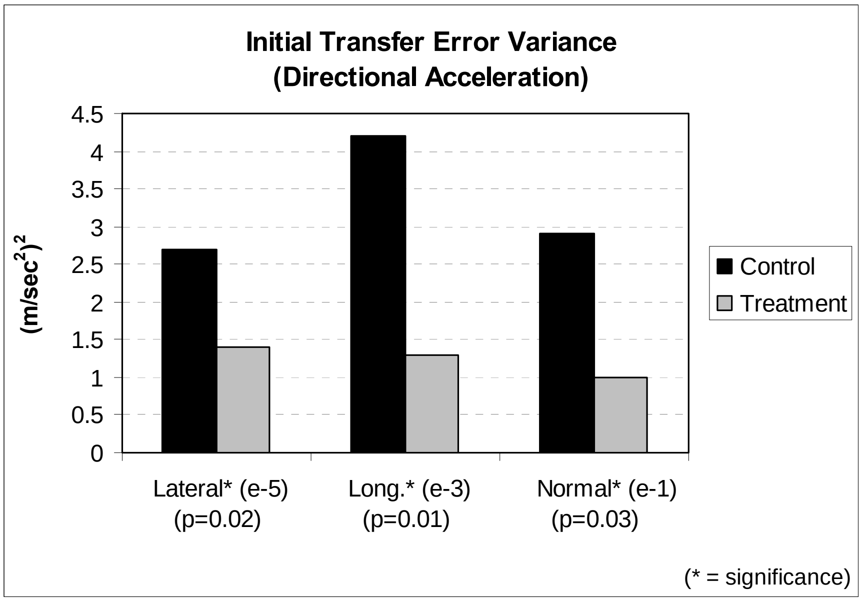


Figure 4.14 Initial acceleration variance analysis

Finally, during the onset of stress, the two groups were significantly different using the variance in roll rate ($F(1,28)=5.14$, $p=0.031$, $\beta=0.41$) and pitch rate ($F(1,28)=5.50$, $p=0.026$, $\beta=0.39$; see Figure 4.15).

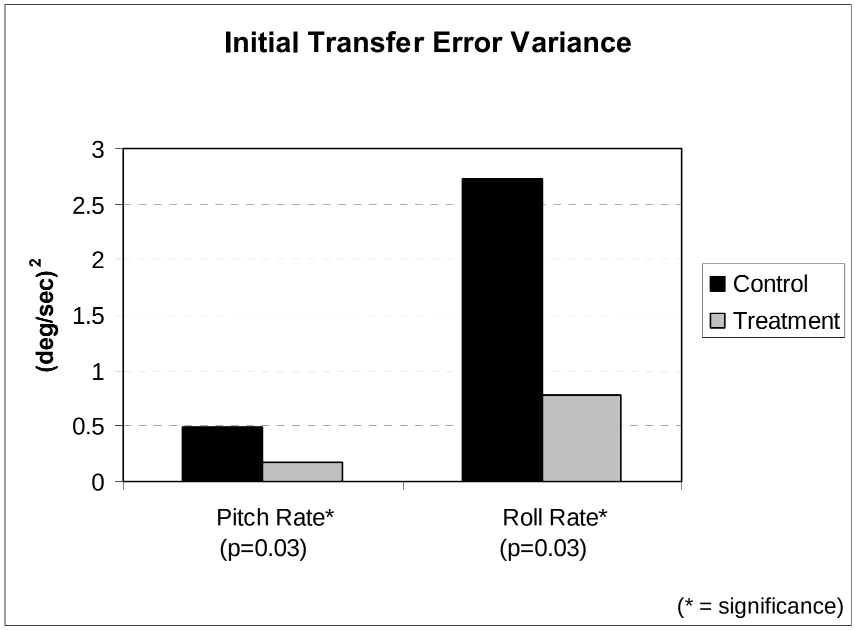


Figure 4.15 Initial rate-of-rotation variance analysis

CFI Performance Evaluations

As stated above, the CFIs in this study were SMEs. In addition, an ANOVA determined that there was no significant difference between the two CFIs ratings ($p>0.05$) implying that they reliably and consistently rated participants' performance. The treatment group performed significantly better than the control group for all five queries and the overall session average (see Figure 4.16):

Query 1: $F(1,28)=5.03$, $p=0.033$, $\beta=0.42$

Query 2: $F(1,28)=6.20$, $p=0.019$, $\beta=0.33$

Query 3: $F(1,28)=6.42$, $p=0.017$, $\beta=0.32$

Query 4: $F(1,28)=5.79$, $p=0.023$, $\beta=0.36$

Query 5: $F(1,28)=5.81$, $p=0.023$, $\beta=0.36$

Average: $F(1,28)=7.47$, $p=0.011$, $\beta=0.25$

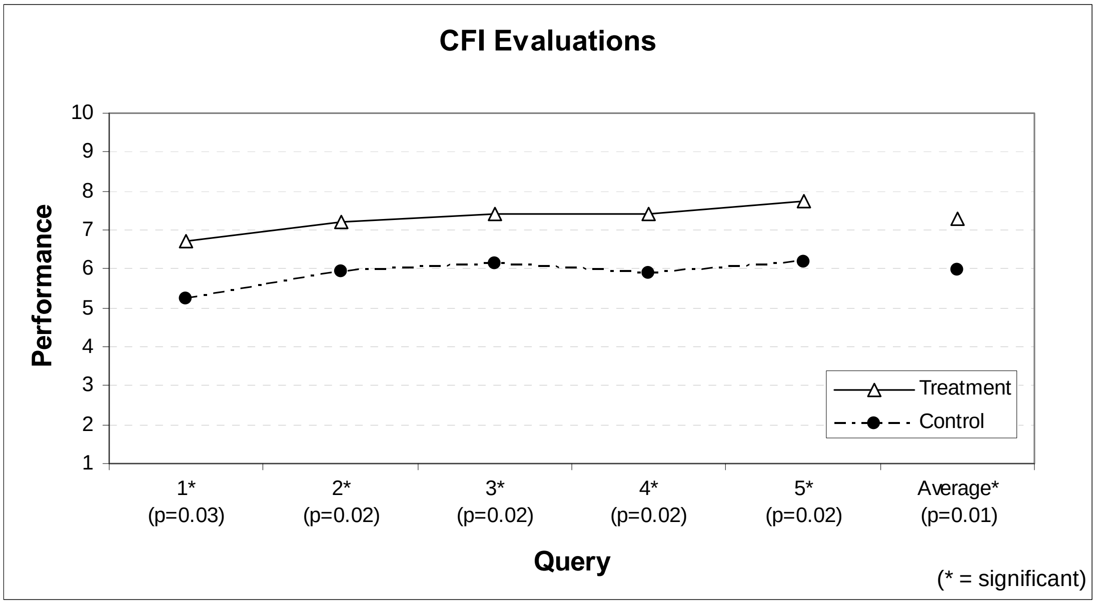


Figure 4.16 CFI performance evaluations

Transfer Effectiveness Ratio

Again, the variance measures were used to calculate the percent transfer and TER for each variable. Table 7 shows these results:

Table 7. Experiment 2 TER

Variable	**% Transfer**	**TER**
Pitch	18.3%	.28
Roll	50.0%	.53
Pitch Rate	59.7%	.93
Roll Rate	96.7%	1.5
Lateral Acceleration	84.4%	1.35
Longitudinal Acceleration	49.2%	.78
Normal Acceleration	59.0%	.90

2. Physiological Measures

Electrocardiogram

Two participants' physiology data was lost during flight due to a failure in the encoder/recorder. The removal of these two participants in addition to the two previously removed participants resulted in a sample size of 28. As a reference for HRV throughout the entire flight, Figure 4.17 shows the two groups' combined LF PSD.

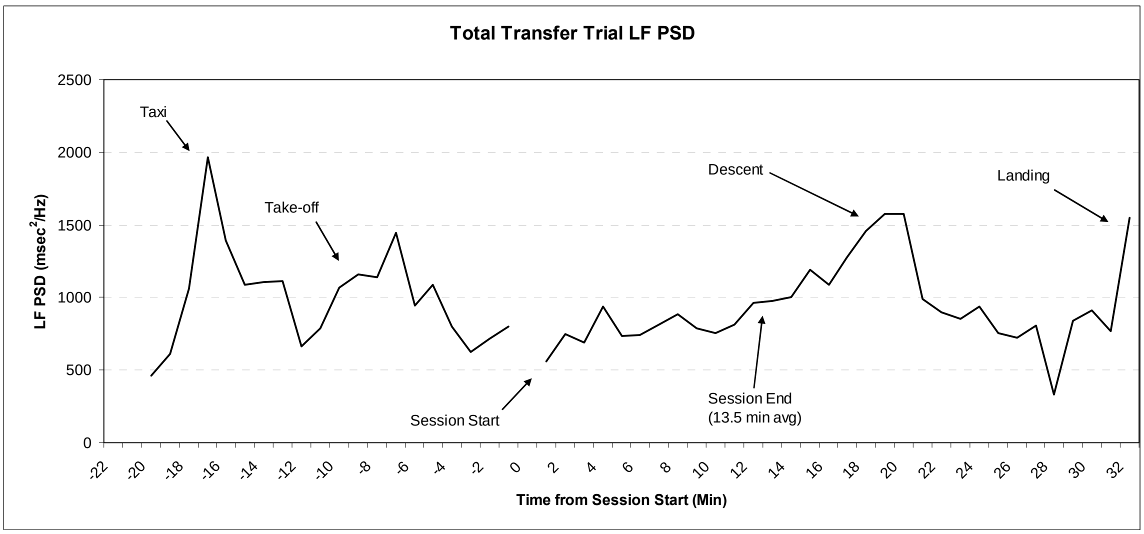


Figure 4.17 Entire flight LF PSD

There was no significant difference between groups when measuring criterion HRV relative to each individual's baseline HRV ($p>0.05$, see Figure 4.18).

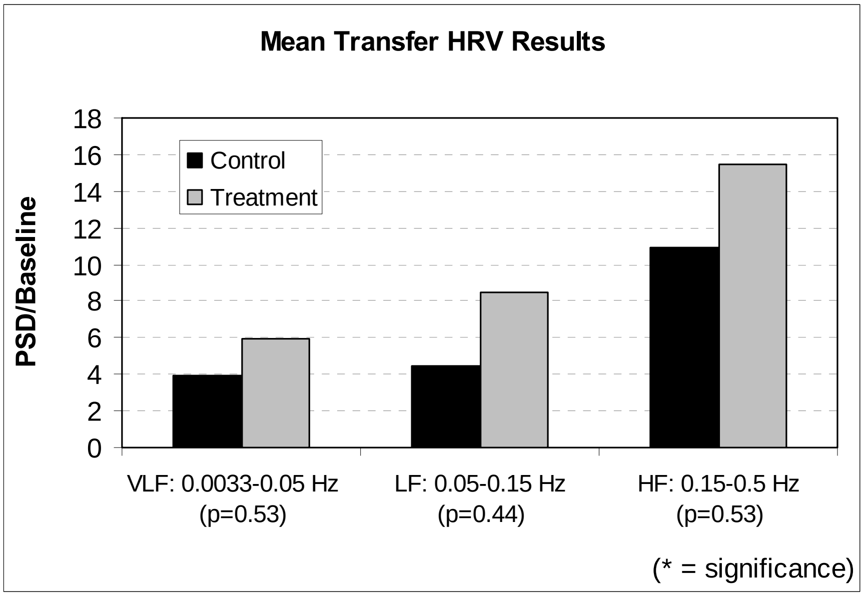


Figure 4.18 Mean transfer results for each PSD band

A step-wise regression was performed to determine other covariate factors other than condition that may be influencing HRV in the transfer task. Condition (C; $t(24)= -2.16$, $p=0.041$), baseline VLF PSD (B; $t(24)=2.72$, $p=0.01$), and gender (G; $t(24)=2.39$, $p=0.02$) explained a significant proportion of the variance in transfer VLF PSD (VLF; $R^2=0.43$, $F(3,24)=5.94$, $p<0.01$):

$$VLF = 514.23 - 107.40C + 0.23B + 193.47G + \varepsilon$$

A step-wise regression also determined that a significant proportion of the variance in transfer LF PSD (LF; $R^2=0.47$, $F(3,24)=7.01$, $p<0.01$) was attributed to condition (C; $t(24)=1.96$, $p=0.062$), baseline LF PSD (B; $t(24)=3.12$, $p<0.01$), and gender (G; $t(24)=2.63$, $p=0.01$):

$$LF = 1023.49 - 219.52C + 0.61B + 476.69G + \varepsilon$$

Electrodermal Response

No significant differences were found between the two experiment groups when using SC as a measure of stress ($p>0.05$). A step-wise regression was performed to determine if other covariate factors other than condition influenced SC in the transfer task. The model indicated that baseline SC (B; $t(24)=4.39$, $p<0.01$), video gaming experience (VG; $t(24)=2.84$, $p<0.01$), and fear of flying (FoF; $t(24)=1.46$, $p<0.01$) all explained a significant proportion of the variance in transfer SC ($R^2=0.55$, $F(3,24)=9.58$, $p<0.01$):

$$SC = 7.21 + 0.73B - 0.78VG - 1.45FoF + \varepsilon$$

Subjective Queries

There was no significant difference between groups for the baseline query ($p>0.05$). The control group indicated a higher subjective stress than the treatment group throughout the entire transfer task (see Figure 4.19), although these differences were not statistically significant ($p>0.05$). The stress reported by the control group was at times more than two times the stress reported during the baseline.

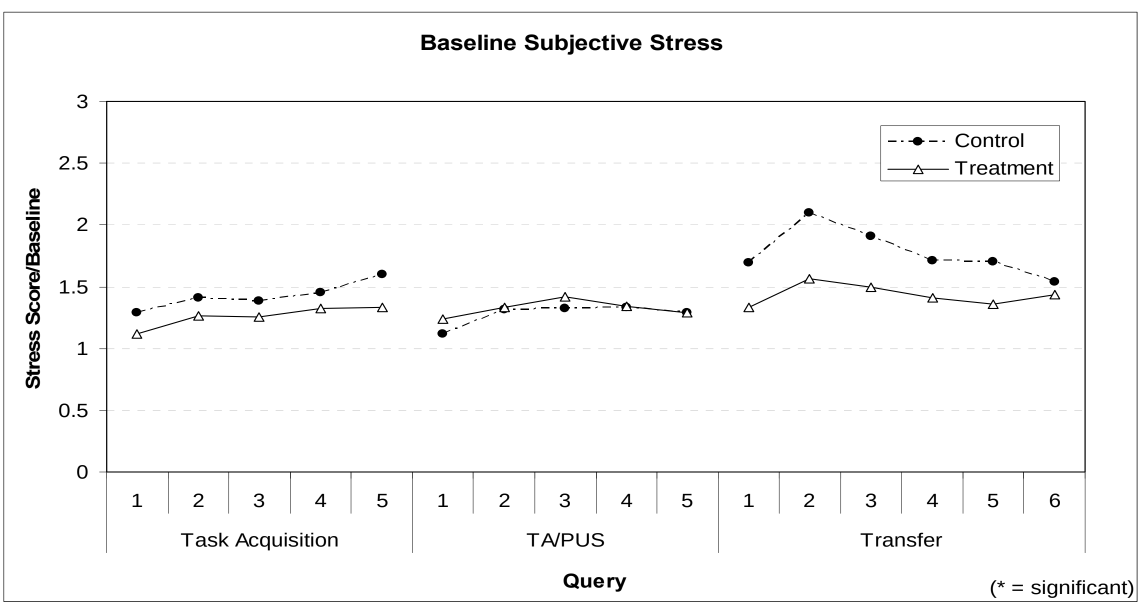


Figure 4.19 Average subjective stress scores for all three sessions

E. EXPERIMENT 2 CONCLUSIONS

The results of Experiment 2 indicate that VE stress training does generalize to a novel task and novel stressor. The novice participants in this study benefited from the artificial, incongruent stressor received in training while exposed to the real-world stress associated with flying an aircraft for the first time in simulated instrument conditions. These results empirically validates Overton's (1964) state-dependent learning theory—introducing stress in training improved performance in a stressful transfer task.

The incongruence between the training and transfer stress also reaffirm Han Selye's (1936) general adaptation syndrome. The participants in Experiment 2 benefitted from cold pressor strain because they produced the same generalized biological responses to both the cold pressor and the strain of flight. Selye's general adaptation syndrome and the supporting Experiment 2 empirical data suggests that trainees can benefit from any stressor applied in training if the transfer task also is stressful.

Finally, these results suggest that the biological responses to stress can be thought of as structure features and the flying context (simulator versus flight) can be considered a surface feature (Driskel et al. 2001). Experiment 2 determined that the context of a training program is not nearly as important as the underlying training principles and pedagogy.

Experiment 1 and 2 results indicate that the stress training presented in these studies did transfer to their respective transfer tasks. However, without manipulating the task training in these experiments, it is still largely unknown to what extent the task skills transferred from the flight simulator to the transfer tasks. Therefore, a third empirical study was devised to test the transfer of task skills from a non-stressful VE flight simulator to flight.

THIS PAGE INTENTIONALLY LEFT BLANK

V. EXPERIMENT 3: TRANSFER OF VE FLIGHT TRAINING

Failure occurs in the context of an overall, national debacle, caused by systemic problems that fall into three distinct but related categories: failure to anticipate, failure to learn and failure to adapt.

— General T. Michael Moseley, USAF Chief of Staff

A. OVERVIEW

A dearth of literature investigates true transfer of training from virtual environments to the real, physical world. As Rose et al. (1998) point out in Chapter II, few studies have empirically determined "what sort of training shows transfer, in what conditions, to what extent, and how robust the transferred training proves to be" (Dishman, Nakamura, Garcia, Dunn, & Blair, 2000, p. 122). Even less empirical evidence exists for the efficacy of flight simulators to the extent that they prepare pilots for in-flight procedures. Therefore, a third study investigated how much a short exposure to a flight simulator prepares a novice student pilot for piloting an aircraft.

The results of Experiment 3 fill a critical void in flight simulator transfer research. Although there is little doubt that flight simulators can prepare a student for flight, whether or not a short, 40 minute exposure will affect both performance, physiology, and subjective appraisal of stress in a transfer task is unknown. In addition, modern aircraft "glass panel" displays make the construction of realistic flight simulators more feasible—however, the transfer of training from these simulators to flight is largely unknown.

A group of participants was recruited and assigned to a "no training" group (no simulator training) and compared to the control group in Experiment 2 (simulator training but no stress training). To prevent confusion, the Experiment 3 control group that received no training will be referred to as the "no-training" group in text and figures; the Experiment 2 control group will continue to be called "control."

B. METHODOLOGY AND EXPERIMENT DESIGN

Participants

Twelve participants were recruited from the pool of active duty military members at both the Naval Postgraduate School and the Defense Language Institute, Monterey, California. The same surveys and requirements as in Experiment 2 were used.

Procedure

The Experiment 2 procedures were followed including the video (see Appendix E2). Both groups were given instruction regarding the simulator controls and displays. The no-training group did not receive any flight simulator time while the control group followed the Experiment 2 simulator training protocol (see Appendix E2). The two groups then performed the same flying transfer task as in Experiment 2.

Hypotheses

The Experiment 3 hypotheses are as follows:

$$H_0: \quad p_T = p_C, \quad a_T = a_C$$
$$H_1: \quad p_T > p_C, \quad a_T < a_C$$

where,

p_T: transfer performance of VE training group

p_C: transfer performance of no-training group

a_T: transfer arousal of VE training group

a_C: transfer arousal of no-training group

C. DATA ANALYSIS

The same data analyses were performed as in Experiments 1 and 2. Blocking was not performed during Experiment 3 due to logistical constraints (see Results). ANOVAs with an alpha level of 0.05 were conducted on all performance, physiology, and subjective stress rating response measures with "condition" used as a predictor variable (IV).

D. RESULTS

One outlier was removed from the Experiment 3 results. This participant was the same outlier from Experiment 2 who was removed due to a nine day delay between the simulator training and the transfer task. The total sample size was 27, with 15 in the control group and 12 in the no-training group.

The 27 participants included 22 males and 5 females, and their ages ranged from 19 to 37 years old with a mean age of 25.7. Mean ratings for gaming experience, personal computer experience, fear of heights, fear of flying, and motions sickness were 2.9, 4.3, 2.2, 1.4, and 1.9, respectively.

There are some characteristics of the Experiment 3 design that need to be noted. Due to logistical and financial constraints, Experiment 3 was given a lower priority than the other two experiments. Therefore, data collection for the Experiment 3 no-training group did not commence until all Experiment 2 data was collected. Thus, the 12 recruits in the no-training group were not randomly assigned their condition; rather, they were the last 12 volunteers for the entire series of experiments. Furthermore, recruitment for Experiments 2 and 3 initially began on the Naval Postgraduate School campus—a population with a large concentration of military officers. Towards the end of the three months of data collection, a disproportionate number of younger enlisted Defense Language Institute students volunteered as word spread throughout that campus regarding the study. Although no significant age or rank differences were found between the control and no-training group ($p>0.05$), the two groups do appear demographically different (see Figure 5.1). Due to the above constraints, many of the following insignificant results did show significant regressions with many contributing covariates other than condition.

Figure 5.1 Age distribution of participants

1. Performance Measures

Aircraft Telemetry Data

The only telemetry variable that resulted in significant findings was total transfer task mean heading variance ($F(1,25)=6.37$, $p=0.018$, $\beta=0.33$). The no-training group demonstrated significantly more mean transfer heading error than the control group. All other telemetry variables showed no difference between the control and no-training groups ($p>0.05$).

CFI Performance Evaluations

There was no significant difference between the two groups for any of the five CFI performance queries nor the overall average evaluation ($p>0.05$).

2. Physiological Measures

Electrocardiogram

As mentioned in Experiment 2 results, two participants' physiology data was lost during flight due to a failure in the encoder/recorder. The removal of these two participants in addition to the previously removed participant resulted in a sample size of 25. The no-training group had a significantly lower LF/HF ratio than the control group for the mean of the entire flight ($F(1,23)=5.15$, $p=0.032$, $\beta=0.42$).

No significant findings were found for the transfer task itself.

Electrodermal Response

The no-training group had a significantly higher SC compared to the control group during the total mean transfer task ($F(1,23)=5.94$, $p=0.023$, $\beta=0.36$). The no-training group's average transfer SC was nearly twice the SC observed during their baseline, while the control group's SC was 1.5 times their baseline.

The no-training group also had a significantly higher mean SC for the entire flight ($F(1,23)=5.82$, $p=0.024$, $\beta=0.37$). Figure 5.2 shows the skin conductance throughout the entire flight for Experiment 2 and 3 participants.

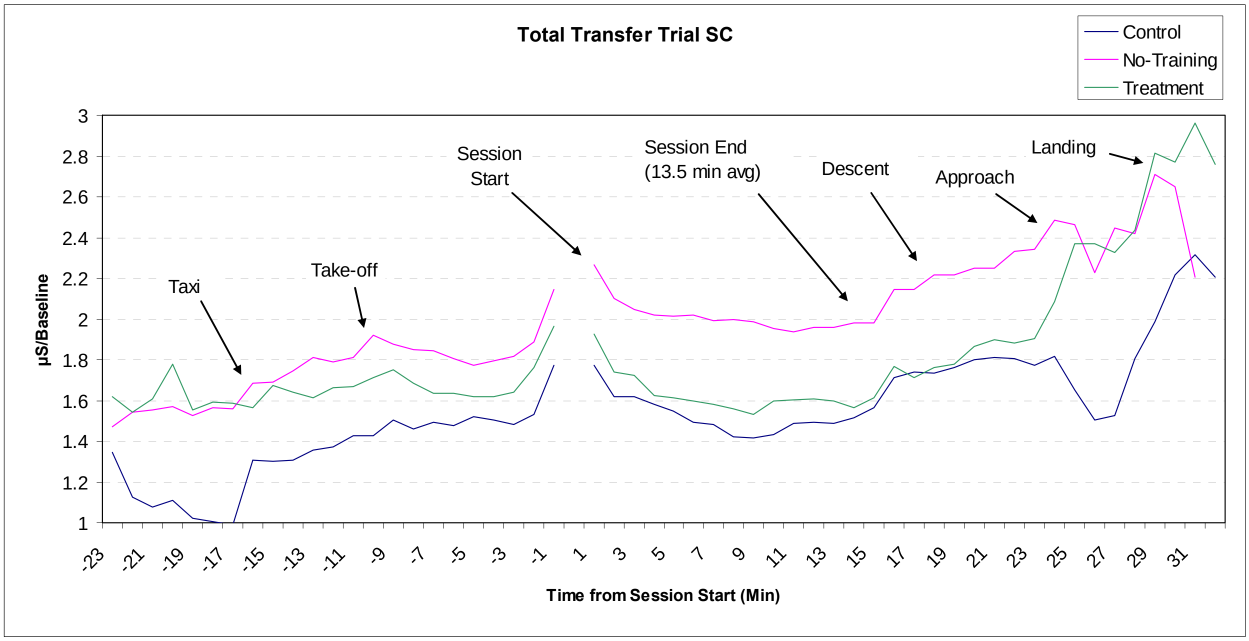


Figure 5.2 SC measures for entire flight

Subjective Queries

The no-training group reported higher subjective stress scores for all six queries compared to the control group, although these differences were not significant (see Figure 5.3). The stress reported by the no-training group was at times more than 2.5 times the stress reported during their baseline.

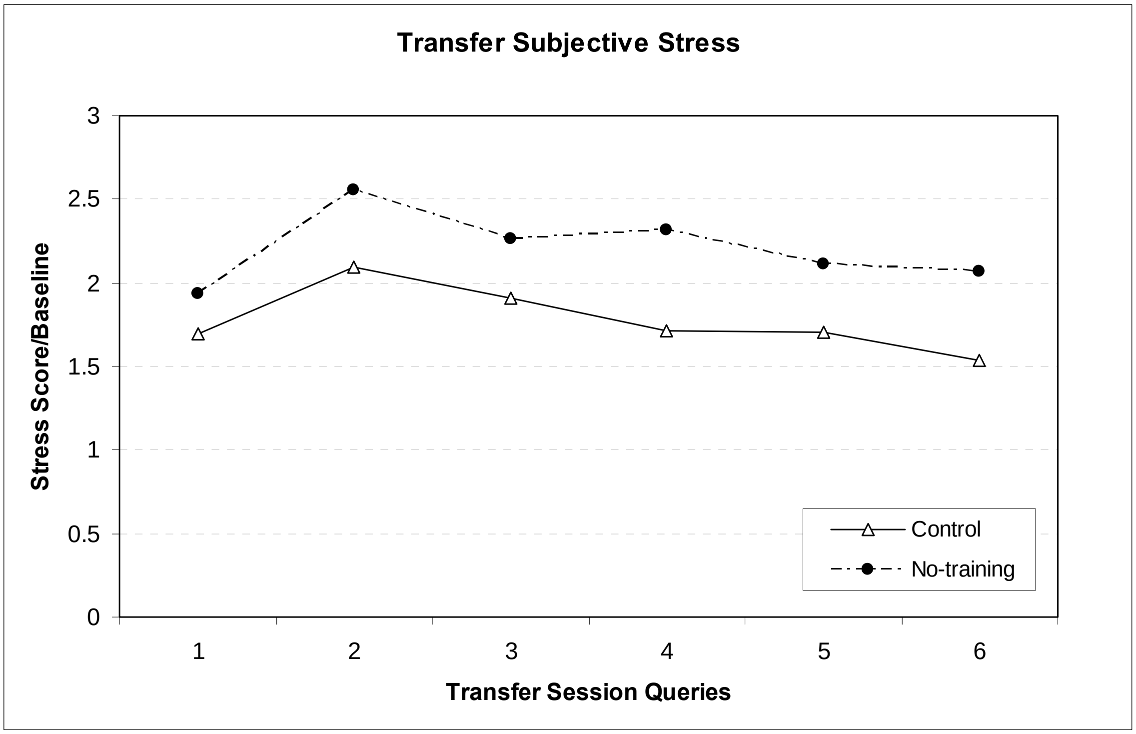


Figure 5.3 Mean subjective stress scores for all subjective stress queries

E. EXPERIMENT 3 CONCLUSIONS

Despite the experiment design limitations in Experiment 3, a slight performance difference can be gleaned from these results—the Experiment 2 control group performed slightly better during the transfer task than the no-training group. These findings are consistent with other training transfer work presented in Chapter II.

The group that received fight simulator training also recorded lower physiological and subjective stress responses during an in-flight transfer task. As pointed out in Chapter II, stress is often defined as the interaction between an individual's perceived resources and the demands of a situation (Beck, 1993; Lazarus & Folkman, 1984; Meichenbaum, 1993; Blascovich, 2007). While Experiments 1 and 2 determined that these perceived resources can be imparted using stress training, Experiment 3 results indicate that perceived resources can also be imparted using skill training. These results support work by Friedland and Keinan (1982, 1984, 1992) which advocates the methodical pairing of both task skill training and stress training in preparation for a stressful transfer task.

The total flight SC results reported in Figure 5.2 along with Experiment 2 HRV results (see Figure 4.17) support work by Hasbrook and Rasmussen (1970). They found that pilots' heart rate increased as they flew closer to the ground on an instrument approach to an airport. Experiment 2 and 3 results confirm these findings using other measures of strain.

VI. DISCUSSION

Knowledge isn't a final destination — something we 'get' and hold on to forever — but is instead a never-ending pursuit.

— *General William R. Looney III,* Commander, Air Education and Training Command

Following the data analysis from each of the three experiments, there are findings that are general to all three experiments. These findings are presented in the following sections, and are organized into both performance and physiology findings. The results of this research also contribute greatly to the development of training pedagogy, and these implications are presented in this chapter. Chapter II describes the relationship between arousal and performance (Yerkes & Dodson, 1908), and results from this research project which contribute to this literature are then presented. Statistical modeling in the three experiments found covariates which effected performance and stress responses other than the experiment condition. Some of these covariates are presented, followed by caveats to the findings in this research project. Finally, future implications and direction for this research project are provided.

A. PERFORMANCE FINDINGS

1. Initial and Total Performance Comparison

The Experiments 1 and 2 performance analyses found similarities between initial transfer performance (criterion and transfer beginning) and total transfer performance. Stress training was effective at improving transfer performance during both the initial onset of stress and throughout the entire transfer task. These results are highly compelling evidence for stress training's efficacy—stress training is not only effective, (1) it works immediately during stress onset, and (2) its advantages stay with a trainee throughout a transfer task. Further research investigating the appropriate duration of stress training in an operational context is needed. Research also is needed that investigates the long-term effects of stress training.

2. CFI Performance Evaluations

Experiments 2 and 3 also explored the effectiveness and feasibility of SME ratings in determining performance. These ratings were sensitive to the research manipulation and were congruent with the objective performance analyses. Trend indications also indicated the measure's sensitivity to a learning curve (see Figure 4.16), although this improvement in performance could be indicative of SME ratings' limitations pointed out in Chapter III (i.e., self-fulfilling prophecy).

3. Variance Measurement in an Aviation Context

Experiment 1 performance results suggests that variance as a performance measure may show promise in an aviation context. These findings also indicate the different effects that stress training may have on both precision and accuracy. The measures using RMSE indicate the treatment group's superior ability to maintain an *accurate* altitude and heading when compared to the control group. Likewise, the participants that received stress training were better able to *precisely* maneuver between clearances when under stress. This finding is not only pivotal for the current research, but it also provides evidence for the use of variance in an aviation context in general. The aviation domain presents many unique challenges to performance evaluation which support the use of a variance measure. These challenges, and justifications for a variance measure, are further discussed here.

One such challenge involves communication breakdowns including incorrect assignment compliance. While piloting an aircraft, and particularly while under instrument flight rules (e.g., radar vectors to a final approach course), controllers continually give course guidance in the form of heading and altitude assignments. Although these assignments are often misinterpreted, errors are often corrected during assignment read-back. If the error is not corrected during read-back, the aircraft's trajectory on a controller's radarscope indicates noncompliance.

In the case of Experiment 1, novice participants were not familiar with reading back clearances (a read-back was not required). In addition, task saturation while complying with simultaneous heading and altitude changes often hindered assignment recall. As Figure 2.5 (and its corresponding discussion in Chapter II) suggests, the stress introduced in these studies also may have increased the probability of a participant missing a clearance assignment (i.e., false negative). Therefore, research participants occasionally flew an incorrect heading and/or altitude, or they may have missed an assignment altogether. In these cases, the experimenter reiterated the correct assignment and waited for correction. Using a RMSE-based measure of performance, these incorrect assignment headings/altitudes were considered error. Furthermore, the amount of error was dependent, not on performance, but on their displacement from the correct assignment.

To illustrate, Figure 6.1 shows two subjects who were both assigned a heading of 270 and 8100 feet. Subject 1 incorrectly maintained a heading of 210 at 8100 feet, while Subject 2 attempted to maintain the correct assignment. Although Subject 1 will produce more error, it appears as if this individual is actually doing a better job of maintaining the incorrect assignment. A variance measure considers performance independent of the assignment, and will score Subject 1 as performing better than Subject 2. However, in the context of this research's objectives (i.e., mitigation of stress following stress training), flying an incorrect heading/altitude may be an indication of stress effects on performance. Assignment noncompliance was not collected, but this could be an effective performance matrix for future stress research.

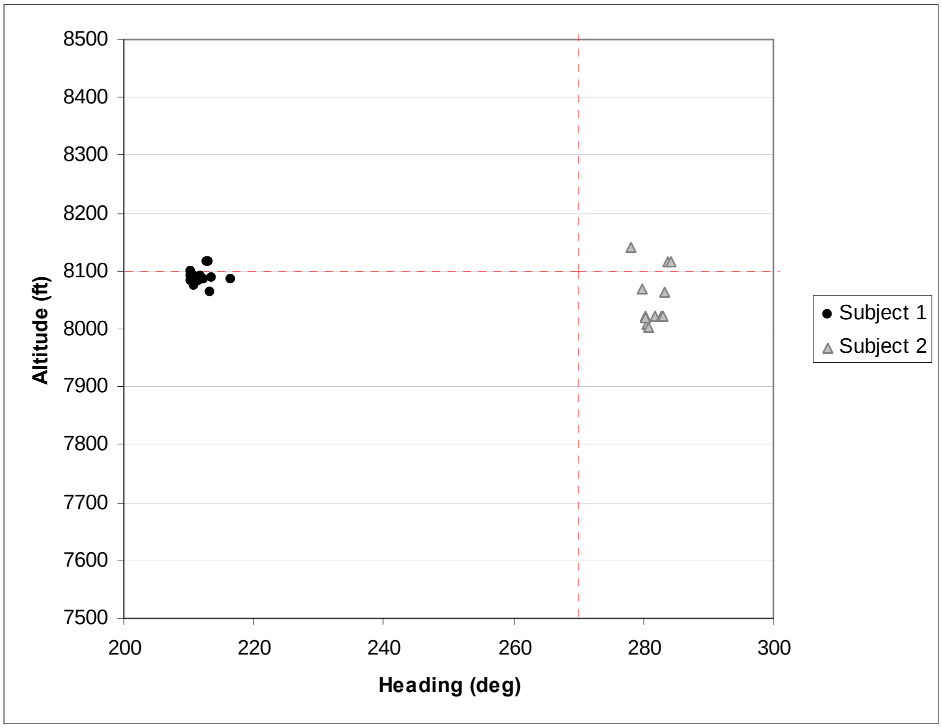


Figure 6.1 Two research participants from Experiment 1

Another challenge of measuring performance in an aviation context is the visual errors presented by analog aviation instruments. These instruments, unlike more modern digital displays, are three dimensional—their needles and icons lay a small distance from the surface of the instrument. This characteristic makes measurement of performance reliant on a viewer's location relative to the instrument. This error is most prominent in the attitude indicator which has a small aircraft icon on top of a moving horizon. The aircraft icon is adjustable and must be set appropriately for the height and distance of the pilot. The variance measure accounts for this error by measuring error from the origin viewed by each individual participant.

Error in the data that is presented by the instruments also is troublesome. For example, the directional gyro precession requires occasional adjustment to the magnetic compass or a GPS receiver. If this and other instruments are not continually monitored for accuracy, the pilot may be receiving inaccurate data from the instruments leading to further performance error. This error is exasperated if the pilot's performance is gathered from independent telemetry data. For example, the altitude that is read on a typical barometric altimeter can be grossly different from true GPS altitude—the altimeter is set to the atmospheric pressure present in a large area of airspace (communicated to a pilot by controllers). The Federal Aviation Regulations require that the altimeter read within 50 feet of the actual airport elevation before take-off. Flying 50 vertical feet from an assigned altitude using the RMSE measure of performance can accumulate an enormous amount of error over time. The variance measure of performance is an effective measure of pilot performance in the presence of the above instrumentation errors.

B. PHYSIOLOGICAL FINDINGS

1. Informed Consent Effects

One unexpected result in Experiment 1 was the significantly higher reported stress by the control group compared to the treatment group during the second task acquisition session. During this session, the treatment group performed the task while exposed to the cold pressor. One possible explanation for this insignificance is the effectiveness of the stress exposure (compared to performance under stress). The stress coping strategies and exposure to the stressor gained during this session may have substantially prepared the participants physiologically for the following PUS session. This finding is consistent with other findings that indicate mental rehearsal is an effective training approach and imagery practice for stress training is effective at reducing performance anxiety (Driskell, Copper, & Moran, 1994). Friedland and Keinan's (1992) TA/SE group further supports this explanation with a quick rise in that group's training performance (see Figure 2.9).

The control group's anxiety in the second task acquisition session may also be explained by the informed consent form. Having read the form which described being subjected to stress, the control group experienced an anticipatory stress response prior to stress exposure. This theory is best demonstrated by an increase in the control group's subjective stress scores throughout the second task acquisition session and a significant difference for the fourth query (see Figure 3.27). In other words, the thought of being subjected to a stressor was more stressful than the actual stressor itself. Due to these observations, the informed consent forms were changed for the following experiments. This change was to prevent the knowledge gained in these forms from affecting experiment trial stress levels, particularly in the control conditions.

In Experiment 2, the treatment and control groups received two different informed consent forms (see Appendices C2 and C3). The only difference in these two forms was that the treatment group was informed they would be subjected to stress, while the control group was not. Even the title of the experiment was changed on the two forms, so the control group did not know that the study pertained to stress in any way. The two groups then followed identical research procedures leading up to and including the baseline physiological recording (video), the subjective stress baseline query, and the task acquisition session. This circumstance allowed the opportunity to perform an empirical analysis comparing the effects of informed consent.

When comparing the two Experiment 2 groups, baseline SC was not effected by the informed consent form ($p>0.05$). Lack of findings when comparing SC baselines is not surprising given the individual differences associated with EDR measures; these measures must be normalized to each individual's baseline, and this analysis attempts to directly compare the actual baselines. During the baseline video, the treatment group did indicate higher HRV in all three PSD bands, although these differences were not statistically significant ($p>0.05$, see Figure 6.2).

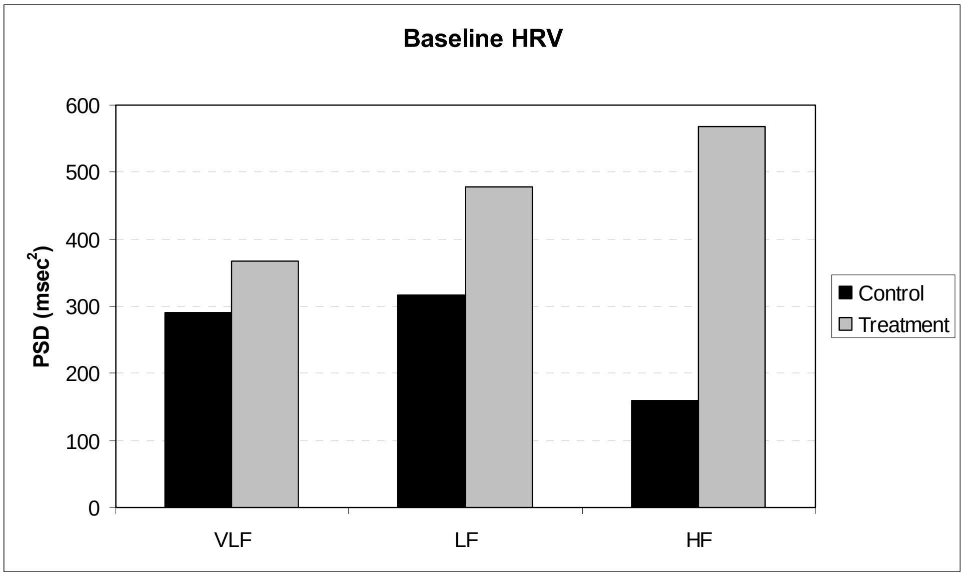


Figure 6.2 Informed consent HRV analysis

Multivariate analyses indicated that video game experience, personal computer experience, age, and rank were all predictors ($p<0.05$) of both baseline VLF ($R^2=0.46$, $F(4,22)=4.86$, $p<0.01$) and baseline LF ($R^2=0.48$, $F(4,22)=5.28$, $p<0.01$) HRV. Rank and age were predictors ($p<0.01$) of baseline HF ($R^2=0.34$, $F(2,24)=6.43$, $p<0.01$). This analysis indicates that in addition to informed consent, random assignment and blocking are important considerations when measuring HRV.

The subjective stress baseline query did not indicate a difference between the treatment and control group ($p>0.05$), although a power analysis indicated an inadequate sample size to test this manipulation ($\beta=0.94$). Trend data indicated a slightly higher baseline score for the treatment group compared to the control group (see Figure 6.3). The identical task acquisition session also indicated higher reported stress scores for the treatment group, although these differences were not significant ($p>0.05$). Further research is needed to address the importance of informed consent verbiage, in addition to any other prior knowledge.

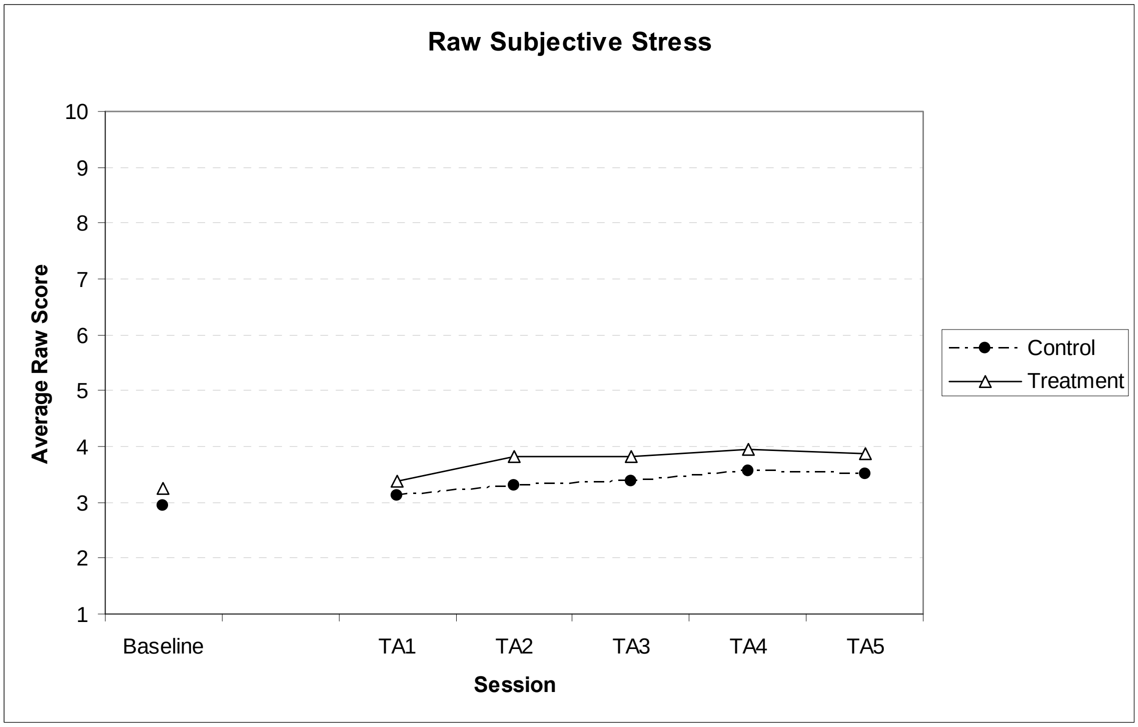


Figure 6.3 Informed consent subjective stress analysis

The subjective stress ratings provide further findings which provide evidence of the general anticipatory response to participating in the study. During the baseline query, participants were asked to rate their current stress level. Most participants took this query seriously and thought momentarily how they felt before giving a rating. During the first minute of the Experiment 1 task acquisition session participants were asked again to rate their stress level. Eighty six percent of participants reported either the same or lower stress rating. This drop in subjective reports indicates that the participants experienced more strain while waiting for the experiment to begin than actually performing in the experiment.

2. Sampling Effects on Stress Results

One aspect of all three experiments that may have decreased overall effect sizes is the eagerness of volunteers to participate. In order to entice a relatively large sample size, attractive recruitment flyers were both posted and distributed electronically (see Appendix I). These flyers described the experiment as "Fun" and "Free.” Therefore, research volunteers may not have been intimidated or stressed by the idea of piloting an aircraft. In fact, the flyer required that volunteers not have severe fear of flying—this verbiage was used for safety reasons.

3. Electrodermal Response as a Measure of Stress

There was a relative lack of significant findings for SC in all three experiments. As mentioned in Chapter V, one reason for this may be the insensitivity of EDR when the stressor is a cold pressor as the cold pressor may be counterproductive to the sweat glands' productivity. Another plausible explanation is the relatively slow response of the electrodermal system making it insensitive to relatively brief state changes such as acute stressors.

C. TRAINING PEDAGOGY

Comparing the Experiment 2 and Experiment 3 results provides informative overall findings. The Experiment 2 results provide much more compelling evidence of that study’s training protocol effectiveness compared to Experiment 3. Previous research presented in Chapter II suggests the importance of pairing both task acquisition and stress training to optimize a training program’s efficacy (Rose, Attree, Brooks, Parslow, Penn, & Ambihaipahan, 1998, p. 69). However, in the case of training novices for a stressful flight environment, this study suggests that stress training may actually be more important than the task training itself. This theory is surely task dependent, and the results suggest the intuitive nature of flying an aircraft—especially considering the commonalities between flying and driving a car (e.g., turn the control left to go left). These results are further explained by the simplification of the flying task in this study—no rudder pedals were used, only three instruments were taught, and other important flying tasks (communication, navigation, engine monitoring, etc.) were not required. This simplification of the task was to decrease the training time and efficiently test the research objectives.

D. RELATIONSHIP BETWEEN AROUSAL AND PERFORMANCE

The results of the first two experiments provide evidence of the syntoxic response suggested by Selye (1974). The stress-trained participants learned to exhibit goal-seeking behaviors (increased performance) in the presence of stressors associated with flying an aircraft. These results also indicate that the structure features trained in a simulator (stress and flight training) generalized to a novel task and stressor (flying an aircraft).

Experiment 3 provided a surprising result with the simulator-trained group showing higher HRV during the transfer task than the no-training group (see Figure 4.18). Conversely, this group demonstrated lower heading error compared to the no-training group. These results demonstrate the intricate balance between cortical arousal and performance as defined by Yerkes and Dodson (1908; see Figure 2.4). For unknown reasons, the no-training group reported relatively low HRV, which may have moved their performance left of optimum on the Yerkes-Dodson curve. The cortical arousal recorded for the simulator-trained group may have provided enough arousal to optimize performance.

E. INDIVIDUAL DIFFERENCES AND OTHER COVARIATES

The participants in this study were inexperienced (zero time piloting an aircraft). As the Yerkes-Dodson Law (1908) suggests, novices may be quickly inundated with stress increasing the efficacy of stress training observed in this study. However, experts may both moderate the effects of stress and require some stress to prevent complacency (see Chapter II, Section C). More research to test the effects of this research on an expert pilot sample is needed.

To reiterate from Chapter II, "the autonomous and sociotopic personality types differ in the kind of stressors to which they are sensitive" (Beck, 1993, p. 345). Given the individual differences associated with stress training, the generalization of results should be limited. Within this current research, individual differences may have limited significant findings. As mentioned in Chapter II, what is stressful for one individual may not be stressful for another.

Other covariates presented also contributed to statistical models. Least surprising was the extent that fear of flying predicted SC measures in Experiment 2. This finding provides further evidence of not only individual differences, but also the task dependency of stress training. Other covariates commonly found in statistical models included gender, age, and military rank. Due to these findings, further research is needed that quantifies the effects these groups have on performance in a stressful environment.

F. CAVEATS

All research participants were military members and relatively young (19-37 years old). This sampling frame was chosen purposely to decrease the variability and individual differences introduced into the experiment design. These demographics are representative of the research target population (military members and aviators). However, for this reason the results of this work should not be generalized to the general population. Further research is needed to address task training and stress training affects on other populations.

In addition, this thesis addressed the feasibility of training novices for stressful situations. Therefore, the affects that stress training may have on experienced operators is unknown. The extent to which stress training should be paired with skill acquisition is unknown and further research in this area is needed.

This research also explored only one stress training protocol. The three-phased stress training approach presented in this work was developed and validated by Friedland and Keinan (1992). However, the effectiveness of other stress training protocols for preparing aviators for stressful flight environments is unknown.

The flying task used in this research was a very simplified aviation task. Throttle control, rudder control, and many other flying sub-tasks (e.g., radio communication, navigation, decision making, regulation adherence, etc.) were not included in the flight simulator or transfer task. This simplification of the task was so that novices could be trained quickly, and research questions could be addressed efficiently. The extent to which these results apply to more realistic and difficult flying tasks is unknown.

The stressful flying task used in Experiments 2 and 3 was relatively benign compared to real-world stressful flight environments. For example, an engine failure or electrical fire may evoke much higher stress responses than the task presented in these experiments. However, conducting highly stressful flight research while ensuring the safety of research participants is a very difficult methodological quandary. One approach to overcome this barrier is simulating unexpected, stressful situations. For example, a highly stressful environment can be created if a research participants is led to believe there is a system malfunction. This predicament ensures a participant's physical safety, although their psychological welfare is still jeopardized.

G. FUTURE RESEARCH

The research presented in this thesis is predicated on the notion that task training is of utmost importance in any training protocol (Caird, 1996; Clark, 2001; Johnston & Cannon-Bowers, 1996). Optimal performance can not be observed in a transfer task if the task skills are not learned in the first place. Only when sufficient skill training is accomplished should other peripheral emotional state training, such as stress training, be incorporated (Friedland & Keinan, 1992; Keinan & Friedland, 1984; Keinan, Friedland, & Sarig-Naor, 1990). However, the literature presented in Chapter II and the experiments conducted in this thesis are simplified partial tasks. These tasks are developed and tested to empirically evaluate specific research questions.

Partial tasks, such as the one tested in this thesis, rarely exist within a training vacuum. These partial tasks are often part of a much larger total training program. For example, student pilots learn many more flying tasks other than navigation during instrument meteorological conditions. This one part-task is paired with many other tasks like landing, take-offs, weather interpretation, communications, etc. Furthermore, a more holistic perspective can be used to view pilot training as a small part-task within the lifecycle of a pilot. Other continuation training, military exercises, and overseas deployments can all be viewed as other part-tasks.

When viewing an entire training program, such as training a pilot, from a global perspective, it is important to determine how stress training should be paired with part-task training. Should it be introduced early as skills are acquired, or should it be introduced prior to known stressful situations (e.g., military deployments)? How should stress training be utilized throughout an entire training lifecycle? How well is stress training retained, and when should reoccurring stress training be provided? These are some of the important operational questions that must be tested prior to the development of affective stress training pedagogy.

There are certain cultural barriers to stress training that also must be addressed. The idea of trainees putting their feet in buckets of ice water will likely not be taken seriously in both current civilian and military communities. Therefore, alternative stressors must be explored which adhere to cultural norms while still effectively introducing stress training.

We need to test other stressors for stress training. Relying on the generalization of stress training literature presented in this thesis (Driskell, Johnston, & Salas, 2001; Selye, 1936, 1956), the actual stressor (i.e., surface features) is not as important as the underlying training principles (i.e., structure features). Therefore, when preparing trainees for stressful events, any stressor introduced in training should theoretically show promise. The degree to which different stressors improve performance and anxiety in transfer tasks is unknown, and more research addressing the effectiveness and feasibility of various stressors is needed.

As seen in this thesis, stress training efficacy also is mediated by individual differences. Although this is a challenge for the development of training pedagogy, some demographic groups seem to respond to stress and stress training in different ways. Chapter II provides evidence that stress response and appraisal are highly reliant on confidence expectancy, organizational success, experience, and anxiety traits attributes (Beck, 1993; Friedland & Keinan, 1992; Meichenbaum, 1993; Otsuka, Onozawa, & Miyamoto, 2006; Parsa & Kapadia, 1997; Saunders, Driskell, Hall, & Salas, 1996). Results from this thesis also determined that gender, fear of flying, and video gaming experience contributed to variance in stress responses during a stressful transfer task (see Chapters IV and V). Further defining demographic groups that respond differently to stress training is needed to develop appropriate training pedagogy.

More work is needed to determine the efficacy of stress training in different contexts. Driskel et al. (2001) point out that the training context can be considered a surface feature and not as important as structure features. However, some transfer contexts are more stressful than others, and also present different emotional and cognitive demands on operators. Therefore, we need to explore the efficacy of stress training in other stressful contexts such as ground combat, firefighting, and air traffic control.

Stress training is presented in this thesis as an alternative to experience-based expertise. By pairing task skill training with stress training, expert performance may be demonstrated in a stressful situation. This high level of performance may be achieved much more efficiently than experience-based expertise. A stress-trained individual may also be better prepared for an unexpected, untrained stressful event than traditional experience-based training approaches. Future work is needed to empirically determine how to optimize stress-training techniques.

VII. CONCLUSIONS

There are two critical points in every aerial flight—its beginning and its end.

— Alexander Graham Bell, 1906

The data from this series of experiments provides compelling evidence for the efficacy of stress training. Related to training transfer definitions presented in Chapter II, Experiment 1 measured how "learning in one context or with one set of materials impacts performance in another context or with other related materials" (Perkins & Salomon, 1992). Experiment 1 determined that stress training in a VE improved trainees' performance in a stressful VE criterion session. The stress training also decreased trainees' physiological response to stressors and decreased their subjective appraisal of stress. This research validated Friedland and Keinan's (1992) stress training protocol in an aviation context.

Experiment 2 expanded on findings regarding the generalization of stress training (Driskell, Johnston, & Salas, 2001; Selye, 1936, 1956, 1974). This experiment answered "whether or not training carried out in a training environment transfers to the equivalent real world situation" (Rose, Attree, Brooks, Parslow, Penn, & Ambihaipahan, 1998). Stress-trained participants in Experiment 2 indicated superior performance and decreased strain during both a novel task and during exposure to a novel stressor compared to a control group. These results are the first evidence that links stress training of this kind to an applicable real-world transfer task.

Finally, Experiment 3 fills a critical void in true transfer-of-training research by testing the extent to which *skills* transfer from a VE to a real-world task. This experiment answered "how skills and knowledge acquired in one domain transfer to another" (Wickens & Hollands, 2000).

Furthermore, the results from these experiments indicate that stress training is effective at increasing performance and mitigating the adverse effects of stress during both the onset of stress and throughout a transfer task. Given these two results, effects of stress training are not limited to only certain portions of a task.

The Sam-to-Sully training paradigm presented in Chapter I explores how expert performance can be gained in a more efficient manner compared to experience-based expertise. Recent media attention has focused on aircraft accidents piloted by relatively inexperienced crews (Chernoff & Laura, 2009), and often these accidents result from inadequate reactions to stressful aircraft situations. The research presented in this thesis presents one potential solution for preparing relatively inexperienced operators for stressful situations by efficiently and effectively bridging the gap between "Sam" and "Sully."

LIST OF REFERENCES

Abel, D. H., & Merryman, R. F. (1987). Training development. In T. S. Andre & A. W. Schopper (Eds.), *Human Factors Engineering in System Design* (pp. 213-242). Write-Patterson AFB, OH: Crew System Ergonomics Information Analysis Center.

Adams, J. A., & McAbee, W. H. (1961). *A program for a functional evaluation of the GAM-83 Melpar Trainer* (No. APGC-TN-61-41). Eglin AFB, FL: U.S. Air Force Air Proving Grounds.

Annett, J. (1998). Training skilled performance. In A. M. Colley & J. R. Beech (Eds.), *Acquisition and Performance of Cognitive Skills*. New York, NY: John Wiley & Sons, Inc.

Army Times. (2007). Editorial: Realistic training a must. Retrieved November 26, 2007, from http://www.armytimes.com/community/opinion/marine_editorial_train_071029/

ASA. (2008). *Federal Aviation Regulations*. Newcastle: Aviation Supplies and Academics, Inc.

Bailenson, J. N., Yee, N., Blascovich, J., Lundblad, N., & Jin, M. (2007). The use of immersive virtual reality in the learning sciences: Digital transformations of teachers, students, and social context. *Journal of the Learning Sciences*.

Baldwin, T. T., & Ford, J. K. (1988). Transfer of training: A review and directions for future research. *Personnel Psychology, 41*, 63-105.

Beck, A. T. (1993). Cognitive approaches to stress. In P. M. Lehrer & R. L. Woolfolk (Eds.), *Principles and Practice of Stress Management* (Second ed., pp. 333-372). New York: The Guilford Press.

Bell, H. H., & Waag, W. L. (1998). Evaluating the effectiveness of flight simulators for training combat skills: A review. *The international journal of aviation psychology, 8*(3), 223-242.

Berntson, G., & Cacioppo, J. (2004). Heart rate variability: Stress and psychiatric conditions. In M. M. A. J. Camm (Ed.), *Dynamic Electrocardiography* (pp. 56-63). New York: Futura.

Blaiwes, A. S., Puig, J. A., & Regan, J. J. (2001). Transfer of training and the measurement of training effectiveness. In R. W. Swezey & D. H. Andrews (Eds.), *Readings in training and simulation: A 30-year perspective*. Santa Monica, CA: Human Factors and Ergonomics Society.

Blascovich, J. (2002). Social Influence within Immersive Virtual Environments. In R. Schroeder (Ed.), *The Social Life of Avatars: Presence and Interaction in Shared Virtual Environments* (pp. 127-145). London: Springer.

Blascovich, J. (2007). Challenge and Threat. In A. J. Elliot (Ed.), *Handbook of approach and avoidance motivation*. Erlbaum.

Bliss, J. P., Tidwell, P. D., & Guest, M. A. (1997). The effectiveness of virtual reality for administering spatial navigation training to firefighters. *Presence, 6*(1), 73-86.

Bourne, P. G. (1970). Psychological aspects of combat. In H. S. Abram (Ed.), *Psychological aspects of stress*. Springfield, IL: Charlies C. Thomas.

Broadbent, D. E. (1958). *Perception and communication*. London: Pergamon Press.

Broadhurst, P. L. (1957). Emotionality and the Yerkes-Dodson Law. *Journal of Experimental Psychology, 54*, 345-352.

Burdea, G. C., & Coiffet, P. (2003). *Virtual Reality Technology* (2nd ed.). Hoboken, New Jersey: John Wiley and Sons, Inc.

Burian, B. K., Dismukes, R. K., & Barshi, I. (2003). The emergency and abnormal situations project. In T. McCarthy (Ed.), *Proceedings of the ISASI 2003 Conference*. Washington, DC.

Bürki-Cohen, J., Soja, N. N., & Longridge, T. (1998). Simulator Platform Motion-The Need Revisited. *International Journal of Aviation Psychology, 8*(3), 293-317.

Caird, J. K. (1996). Persistent issues in an application of virtual environment systems to training. *IEEE Computer Society: Human Interaction with Complex Systems, 3*, 124-132.

Cannon, W. B. (1915). *Bodily changes in pain, hunger, fear and rage*. New York: Appleton.

Carretta, T. R., & Dunlap, R. D. (1998). *Transfer of training effectiveness in flight simulation: 1986 to 1997* (Tech. No. AFRL-HE-AZ-TR-1998-0078). Brooks Air Force Base, TX: Air Force Research Laboratory.

Cheng, E. W. L., & Ho, D. C. K. (1999). A review of transfer of training studies in the past decade. *Personnel Review, 30*(1), 102-118.

Chernoff, A., & Laura, D. (2009). Florida's 'pilot factory'. Retrieved May 28 from http://amfix.blogs.cnn.com/2009/05/28/floridas-pilot-factory/

Clark, R. E. (2001, October 24-25). *Learning from media: What works and what doesn't.* Paper presented at the Symposium on Assessing the Quality of Online Instruction, Monterey, CA.

Clawson, D. M., Miller, M. S., Knott, B. A., & Sebrechts, M. M. (1998). *Navigational training in virtual and real buildings.* Paper presented at the Human factors and ergonomics society, Chicago, IL.

Cobb, S., & Fraser, D. S. (2005). Multimedia learning in virtual reality. In R. Mayer (Ed.), *The Cambridge handbook of multimedia learning*. New York, NY: Cambridge University Press.

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior, 24*, 385-396.

Darken, R. P. (2006). Identifying the limits of training system effectiveness through taxonomies of human performance. *Theoretical Issues in Ergonomics Science, X*(X), xxx.

Darken, R. P. (2009). Identifying the limits of training system effectiveness through taxonomies of human performance. *Theoretical Issues in Ergonomics Science, 10*(3), 231-243.

Darken, R. P., & Banker, W. P. (1998). *Navigating in natural environments: A virtual environment training transfer study.* Paper presented at the Virtual Reality Annual International Symposium, Los Alamitos, CA.

Department of Transportation. (2001). *Instrument Flying Handbook* (No. FAA-H-8063-15).

Dion, D. P., Smith, B. A., & Dismukes, P. (1996). The cost/fidelity balance. *Modern Simulation and Training: The International Training Journal, 2*, 38-45.

Dishman, R. K., Nakamura, Y., Garcia, M. E., Dunn, R. W. T. A. L., & Blair, S. N. (2000). Heart rate variability, trait anxiety, and perceived stress among physically fit men and women. *International Journal of Psychophysiology, 37*, 121-133.

Dreyfus, H. L. (2002). Intelligence without representation. *Phenomenology and the Cognitive Sciences, 1*(4).

Driskell, J. E., Copper, C., & Moran, A. (1994). Does mental practice enhance performance. *Journal of Applied Psychology, 79*(4), 481-492.

Driskell, J. E., Johnston, J. H., & Salas, E. (2001). Does stress training generalize to novel settings? *Human Factors, 43*(1), 99-110.

Druckman, D., & Swets, J. A. (ed., 1988). *Enhancing Human Performance*. Washington, D. C.: National Academy of Sciences Press.

Fox, S. I. (1993). *Human Physiology* (Fourth ed.). Oxford, England: Wm. C. Brown Publishers.

Friedland, N., & Keinan, G. (1982). Patterns of fidelity between training and criterion situations as determinants of performance in stressful situations. *Journal of Human Stress, 8*, 41-46.

Friedland, N., & Keinan, G. (1992). Training effective performance in stressful situations: Three approaches and implications for combat training. *Military Psychology, 4*(3), 157-174.

Friedman, B. H., & Thayer, J. F. (1998). Anxiety and automatic flexibility: a cardiovascular approach. *Biological Pscyhology, 49*, 303-323.

Gibb, R., & Olson, W. (2005). *Classification of Air Force aviation accidents: Mishap trends and prevention* (No. CI04-1814): Arizona State University.

Goldstein, I. L. (1993). *Training in organizations* (3rd ed.). Monterey, CA: Brooks/Cole.

Grossman, S. P. (1967). *A Textbook of Physiological Psychology*. New York, NY: Wiley and Sons, Inc.

Hancock, P. A., Ward, P., Szalma, J. L., & Ganey, H. C. N. (2002). *Stress and human information processing: A descriptive framework presented in a novel manner*. Paper presented at the 23rd Annual Army Science Conference, Orlando, FL.

Hart, S., & Hauser, J. (1987). In-flight application of three pilot workload measurement techniques. *Aviation, Space, and Environmental Medicine, 58*, 402-410.

Hasbrook, H. A., & Rasmussen, P. G. (1970). Pilot heart rate during in-flight simulated instrument approaches in a general aviation aircraft. *Aerospace Medicine, 41*, 1148-1152

Hastings, D. W., Wright, D. G., & Clueck, B. C. (1944). *Psychiatric experience of the Eighth Air Force, First year of combat*. New York: Josiah Macy, Jr., Foundation.

Hays, R. T., Jacobs, J. W., Prince, C., & Salas, E. (1992). Flight simulator training effectiveness: A meta-analysis. *Military Psychology, 4*(2), 63-74.

Hockey, G. R. J. (1984). Varieties of attentional state: The effects of the environment. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention*. New York: Academic Press.

Hoover, A., & Muth, E. (2004). A real-time index of vagal activity. *International Journal of Human-Computer Interaction, 17*(2), 197-209.

Hoyt, C. L., Blascovich, J., & Swinth, K. R. (2003). Social inhibition in immersive virtual environments. *Presence: Teleoperators and Virtual Environments, 12*(2), 183-195.

Jagacinski, R. J., & Flach, J. M. (2003). *Control Theory for Humans: Quantitative Approaches to Modeling Performance*. Mahwah, NJ: Lawrence Erlbaum.

Jeppesen Sanderson Inc. (1996). *Private Pilot Manual*. Englewood, CO: Jeppesen Sanderson, Inc.

Jeppesen Sanderson Inc. (2006a). *Guided Flight Discovery: Instrument/Commercial*. Englewood, CO: Jeppesen Sanderson, Inc.

Jeppesen Sanderson Inc. (Writer) (2006b). Guided flight discovery: Instrument/Commercial video series. Englewood, CO: Jeppesen Sanderson, Inc.

Jeppesen Sanderson Inc. (2006c). *Instrument Rating FAA Airman Knoledge Test Guide*. Englewood, CO: Jeppesen Sanderson, Inc.

Johnston, J. H., & Cannon-Bowers, J. A. (1996). Training for stress exposure. In J. E. Driskell & E. Salas (Eds.), *Stress and Human Performance* (pp. 223-257). Mahwah, NJ: Lawrence Erlbaum Associates.

Kaempf, G. L., Cross, K. D., & Blackwell, N. J. (1989). *Backward transfer and skill acquisition in the AH-1 flight and weapons simulator* (ARI Research Report No. 1537). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Kalsbeek, J., & Ettema, J. (1963). Scored regularity of the heart rate pattern and the measurement of perceptual or mental load. *Ergonomics, 6*, 306-307.

Kanfer, F. H., & Seidner, M. L. (1973). Self control: Factors enhancing tolerance of noxious stimulation. *Journal of Personality and Social Psychology, 25*, 381-389.

Keinan, G., & Friedland, N. (1984). Dilemmas concerning the training of individuals for task performance under stress. *Journal of Human Stress, 10*, 185-190.

Keinan, G., Friedland, N., & Sarig-Naor, V. (1990). Training for task performance under stress: The effectiveness of phased training methods. *Journal of Applied Social Pscyhology, 20*, 1514-1529.

Klonowicz, T. (1989). Reactivity, experience, and response to noise. In C. D. Spielberger, I. G. Sarason & J. Strelau (Eds.), *Stress and anxiety* (Vol. 12, pp. 123-139). New York: Hemisphere.

Koonce, J. M., & Bramble, W. J., Jr. (1998). Personal computer-based flight training devices. *The international journal of aviation psychology, 8*(3), 877-292.

Kort, B., Reilly, R., & Picard, R. W. (2001). *An affective model of interplay between emotions and learning: Reengineering pedagogy--Building a learning companion.* Paper presented at the ICALT01.

Kozak, J. J., Hancock, P. A., Arthur, E. J., & Chrysler, S. T. (1993). Transfer of training from virtual reality. *Ergonomics, 36*(7), 777-784.

Kroemer, K. H. E., Kroemer, H. J., & Kroemer-Elbert, K. E. (1986). *Engineering Psychology: Physiology bases of Human Factors/Ergonomics* (Vol. 4). New York: Elsevier.

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping*. New York: Springer.

Lintern, G., Roscoe, S. N., Koonce, J. M., & Segal, L. D. (1990). Transfer of landing skills in beginning flight training. *Human Factors, 32*(3), 319-327.

Lintern, G., Sheppard, D. J., Parker, D. L., Yates, K. E., & Nolan, M. D. (1989). Simulator design and instructional features for air-to-ground attack: A transfer study. *Human Factors, 31*(1), 87-99.

Lintern, G., Sivier, J., & Roscoe, S. N. (1990). Display principles, control dynamics, and environmental factors in pilot training and transfer *Human Factors, 32*.

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*(2), 149-169.

Lintern, G., Thomley-Yates, K. E., Nelson, B. E., & Roscoe, S. N. (1987). Content, variety, and augmentation of simulated visual scenes for teaching air-to-ground attack. *Human Factors, 29*(1), 45-59.

Loftin, R. B., & Kenney, P. J. (1994). *Virtual environments in training: NASA's Hubble space telescope mission.* Paper presented at the 16th Interservice/Industry Training Systems and Education Conference, Orlando, FL.

Loomis, J. M., Blascovich, J. J., & Beal, A. C. (1999). Immersive virtual environment technology as a basic research tool in psychology. *Behavior Research: Methods, Instruments, & Computers, 31*, 557-564.

Mackie, R., & Miller, J. (1978). *Effects of hours of service, regularity of schedules, and cargo loading on trucks and bus driver fatigue* (HFR- No. TR-1765-F). Goleta, GA: Human Factors Research, Inc.

Marcotte, T. D., Roberts, E., Rosenthal, T. J., Heaton, R. K., Bentley, H., Grant, I., et al. (2003). *Test-retest reliability of standard deviation of lane position as assessed on a PC-based driving simulator.* Paper presented at the Second International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, Park City, UT.

Matthews, G., Davies, D. R., Westerman, S. J., & Stammers, R. B. (2000). *Human Performance*. Philadelphia, PA: Psychology Press.

Mayer, R. E. (2005). *The Cambridge Handbook of Multimedia Learning*. New York, NY: Cambridge University Press.

McCauley, M. E. (2006). *Do army helicopter training simulators need motion base?* (Tech. Report No. 1176). Monterey, CA: Naval Postgraduate School.

Meehan, M., Insko, B., Whitton, M., & Brooks, F. P., Jr. (2002). *Physiological measures of presence in stressful virtual environments.* Paper presented at the SIGGRAPH, San Antonio, TX, USA.

Meichenbaum, D. (1993). Stress inoculation training: A 20-year update. In P. M. Lehrer & R. L. Woolfolk (Eds.), *Principles and Practice of Stress Management* (Second ed., pp. 373-406). New York: The Guilford Press.

Metalis, S., Bieferno, M., & Corwin, W. (1989). *Is heart rate a valid, reliable, and applicable index of pilot workload in commercial transport aircraft?* Paper presented at the Fifth International Conference on Aviation Psychology.

Moshell, J. M., & Hughes, C. E. (2002). Virtual environments as a tool for academic learning. In K. Stanney (Ed.), *Handbook of virtual environments, design, implementation and applications* (pp. 893-910). Mahwah, NJ: Lawrence Erlbaum Associates.

Mullen, B., Bryant, B., & Driskell, J. E. (1997). Presence of other and arousal: An integration. *Group dynamics: Theory, research, and practice, 1*(1), 52-64.

Nash, E. B., Edwards, G. W., Thompson, J. A., & Barfield, W. (2000). A review of presence and performance in virtual environments. *International Journal of Human-Computer Interaction, 12*(1), 1-41.

NTSB. (2000). Animations of runway incursions from Board Meeting. Retrieved December 11, 2007, from http://www.ntsb.gov/Events/2000/incursion/incur_video.htm

Orlansky, J., String, J., & Chatelier, P. R. (1979). *The cost-effectiveness of military training* (No. AD-POOO168). Alexandria, VA: Office of the Secretary of Defense.

Osgood, C. E. (1949). The similarity paradox in human learning: A resolution. *Psychology Review, 56*, 132-143.

Otsuka, Y., Onozawa, A., & Miyamoto, Y. (2006). Hormonal responses of pilots to training flights: The effects of experience on apparent stress. *Aviation, Space, and Environmental Medicine, 77*(4), 410-414.

Overton, D. A. (1964). State-dependent learning produced by depressant and atropine-like drugs. *Psychopharmacology, 10*(1), 6-31.

Parsa, B. B., & Kapadia, A. S. (1997). Stress in Air Force aviators facing the combat environment. *Aviation, Space, and Environmental Medicine, 68*(12), 1088-1092.

Pepler, R. D. (1958). Warmth and performance: An investigation in the tropics. *Ergonomics, 2*, 63-88.

Perala, C. H., & Sterling, B. S. (2007). *Galvanic skin response as a measure of soldier stress* (Tech. No. ARL-TR-4114): Army Research Laboratory.

Perkins, D. N., & Salomon, G. (1992). Transfer of learning. In *International Encyclopedia of Education* (2nd ed.). Oxford, England: Pergamon Press.

Péruch, P., Belingard, L., & Thinus-Blanc, C. (2000). Transfer of spatial knowledge from virtual to real environments. In C. Freksa, W. Bauer, C. Habel & K.Wender (Eds.), *Spatial cognition II, LNAI 1849* (pp. 253-264). Berlin, Germany: Springer-Verlag.

Pfieper, S. J., & Hammill, S. C. (1995). Heart rate variability: Technique and investigational applications in cardiovascular medicine. *Mayo Clinic Proceedings, 70*, 955-964.

Prather, D. C. (1971). Trial-and-error versus errorless learning: Training, transfer, and stress. *American Journal of Psychology, 84*(3), 377-386.

Reason, J. (1990). *Human Error*. New York: Cambridge University Press.

Regian, J. W., Shebilske, W. L., & Monk, J. M. (1992). Virtual Reality: An Instructional Medium for Visual-Spatial Tasks. *Journal of Communication, 42*(4), 136.

Ricci, B. (1967). *Physiological Basis of Human Performance*. Philadelphia: Lea & Febiger.

Rizzo, A., Morie, J. F., Williams, J., Pair, J., & Buckwalter, J. G. (2005). *Human emotional state and its relevance for military VR training.* Paper presented at the 11th International Conference on Human Computer Interaction.

Roscoe, S. N. (1971). Incremental transfer effectiveness. *Human Factors, 13*(6), 561-567.

Roscoe, S. N., & Williges, B. H. (1980). Measurement of transfer of training. In S. N. Roscoe (Ed.), *Aviation Psychology*. Ames: Iowa State University Press.

Rose, F. D., Attree, E. A., Brooks, B. M., Parslow, D. M., Penn, P. R., & Ambihaipahan, N. (1998). *Transfer of training from virtual to real environments.* Paper presented at the European Conference on Disability, Virtual Reality, and Associated Technologies, Skovde, Sweden.

Rose, F. D., Attree, E. A., Brooks, D. M., Parslow, P. R., Penn, P. R., & Ambihaipahan, N. (2000). Training in virtual environments: Transfer to real world tasks and equivalence to real task training. *Ergonomics, 43*(4), 494-511.

Rosenbaum, M. (1980). Individual differences in self-control behaviors and tolerance of painful stimulation. *Journal of Abnormal Psychology, 89*, 581-590.

Saba. (2001). *Next generation e-Learning and the road to human capital development and management*. Redwood Shores, CA: Saba Software, Inc.

Salas, E., Bowers, C. A., & Rhodenizer, L. (1998). It is not how much you have but how you use it: Toward a rational use of simulation to support aviation training. *The international journal of aviation psychology, 8*(3), 197-208.

Salzman, M. C., Dede, C., & Bowen Loftin, R. (1999). A model for how VR aids complex conceptual learning. *Presence, 8*(3), 293-316.

Saunders, T., Driskell, J. E., Hall, J., & Salas, E. (1996). *The effects of stress inoculation training on anxiety and performance* (Final). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Selye, H. (1936). A syndrome produced by diverse nocuous agents. *Nature, 138*, 32.

Selye, H. (1956). *The Stress of Life*. New York, NY: McGraw-Hill.

Selye, H. (1974). *Stress without distress*. New York: New York: American Library.

Seymour, N. E., Gallagher, A. G., Roman, S. A., O'Brien, M. K., Bansal, V. K., Andersen, D. K., et al. (2002). Virtual reality training improves operating room performance. *Annals of Surgery, 236*(4), 458-464.

Sloan, R., Shapiro, P., Bagiella, E., Boni, S., Paik, M., Bigger, J. Jr., et al. (1994). Effect of mental stress throughout the day on cardiac autonomic control. *Biological Pscyhology, 37*(2), 89-99.

Smode, A. F., Hall, E. R., & Meyer, D. E. (1966). *An assessment of research relevant to pilot training* (Tech No. AMRL-TR-66-196). Wright-Patterson AFB, OH: Aerospace Medical Research Laboratory.

Stewart II, J. E. (1994). *Using the backwards transfer paradigm to validate the Ah-64 simulator training research advanced testbed for aviation* (No. AD-A285 758).

Sullenberger, C. (2009, February 5). [Interview with Catie Couric, host of 60 minutes]. Sullenberger recalls moment engines died: Catie Couric interviews the amazing pilot and crew of US Airways flight 1549. Retrieved March 22, 2009, from: CBS News Website: http://www.cbsnews.com/stories/2009/01/30/60minutes/main4764852.shtml.

Tate, D. L., Sibert, L., & King, T. (1997). *Virtual environments for shipboard firefighting training.* Paper presented at the Virtual Reality Annual International Symposium.

Taylor, H. L., Lintern, G., Koonce, J. M., Jefferson, M., Kaiser, R. H., & Morrison, G. A. (1991, 23-26 Sept. 1991). *Simulator scene detail and visual augmentation guidance in landing training for beginning pilots.* Paper presented at the Aerospace Technology Conference and Exposition, Long Beach, CA.

Taylor, H. L., Lintern, G., Koonce, J. M., Kunde, D. R., Tschopp, J. M., & Talleur, D. A. (1993). *Scene content, field of view and amount of training in first officer training.* Paper presented at the 7th International Symposium on Aviation Psychology, Columbus.

Thorndike, E. L., & Woodworth, R. S. (1901). The influence of improvement in one mental function upon the efficiency of other functions. *Psychological Review, 8*, 247-261.

Vaden, E. A., & Hall, S. (2005). The Effect of Simulator Platform Motion on Pilot Training Transfer: A Meta-Analysis. *International Journal of Aviation Psychology, 15*(4), 375-393.

van Ravenswaaij-Arts, C., Kollee, L., Hopman, J., Stoelinga, G., & van Geijn, H. (1993). Heart Rate Variability. *Annals of Internal Medicine, 118*(6), 436-447.

Wagner, R. J., & Campbell, J. (1994). Outdoor-based experiential training: Improving transfer of training using virtual reality. *Journal of Management Development, 13*(7), 4-11.

Waller, D., Hunt, E., & Knapp, D. (1998). The transfer of spatial knowledge in virtual environment training. *Presence, 7*(2), 129-143.

Welford, A. T. (1973). Stress and performance. *Ergonomics, 16*(5), 567-580.

Wickens, C. D., & Hollands, J. G. (2000). *Engineering psychology and human performance* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Wickens, C. D., Lee, J. D., Liu, Y., & Becker, S. E. G. (2004). *An Introduction to Human Factors Engineering* (2nd ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

Wilson, P. N., Foreman, N., & Tlauka, M. (1997). Transfer of spatial information from a virtual to a real environment. *Human Factors, 39*(4), 526-531.

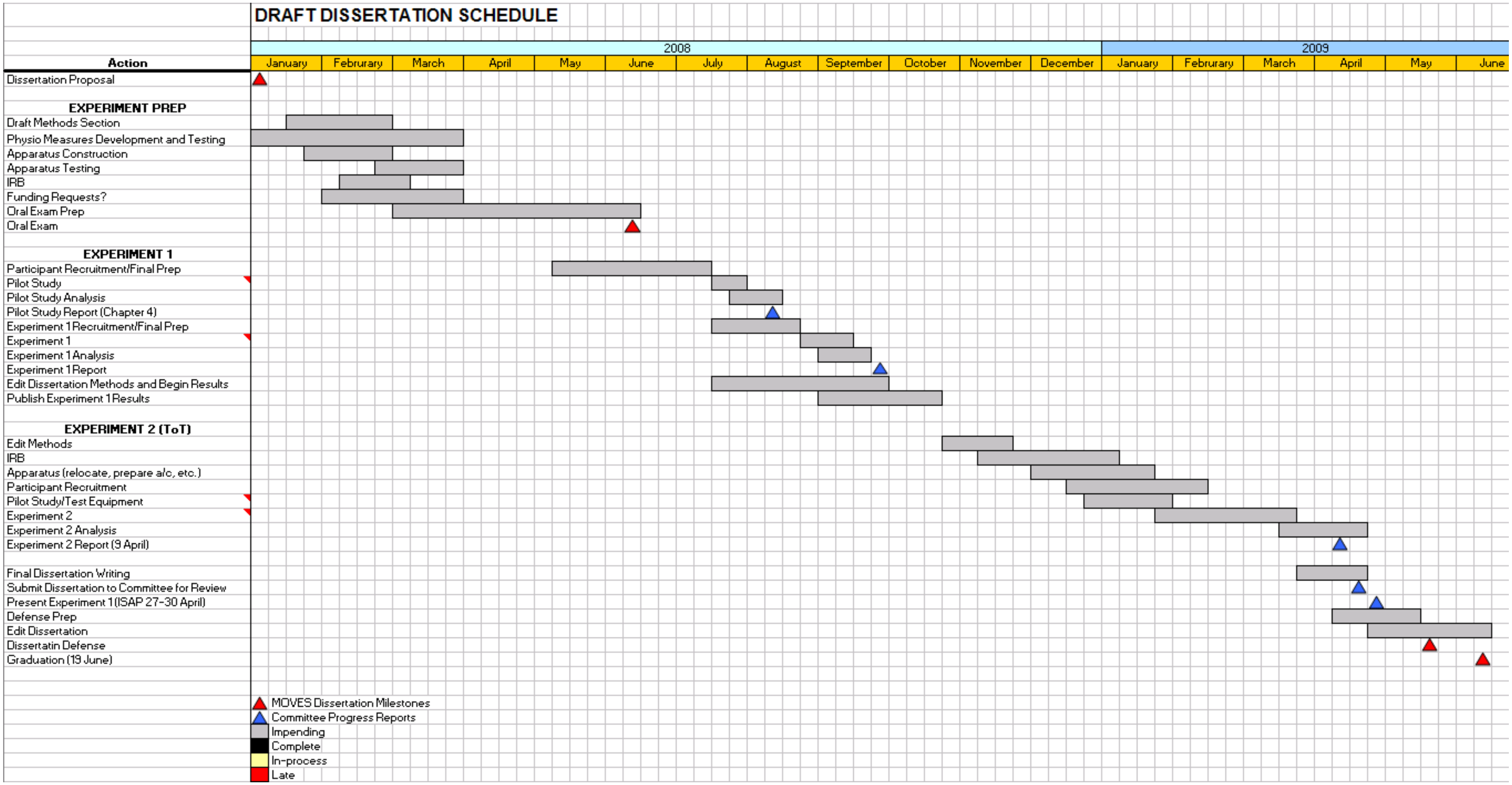
Witmer, B. G., Bailey, J. H., & Knerr, B. W. (1996). Virtual spaces and real world places: Transfer of route knowledge. *International Journal of Human-Computer Studies, 45*, 413-428.

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments, 7*(3), 225-240.

Wylie, C. D., Shultz, T., Miller, J. C., Mitler, M. M., & Mackie, R. R. (1996). *Commercial motor vehicle driver fatigue and alertness study: Project report* (No. FHWA-MC-97-001). Goleta, CA: Essex, Corp.

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology, 18*(459-482), 459.

APPENDIX A: RESEARCH SCHEDULE



THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX B: PARTICIPANT SURVEY

Experimenter Use			
Ex.:		ID:	
Cond.:		Gender: M / F	
C.V.A.:	L 20/____	Handedness: L / R	
	R 20/____	Age:	

For the following questions, use this scale:

1	2	3	4	5
None		**Occasional**		**Frequent**

1. Rate your video gaming experience: ☐

2. Rate your PC experience: ☐

3. Rate your flying experience: ☐

   If you answered anything other than "1" for "flying experience," please explain:
   ____
   ____
   ____

4. Rate your Air Traffic Control experience: ☐

   If you answered anything other than "1" for "ATC experience," please explain:
   ____
   ____
   ____

5. Rate your overall health (check one):
   __ Very Good
   __ Good
   __ Average
   __ Poor
   __ Very Poor

6. Do you have any heart conditions? If yes, explain:
   ____
   ____
   ____

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX C1: INFORMED CONSENT FORM (EXPERIMENT 1)

Introduction. You are invited to participate in a study entitled Stress Effects on Transfer from Virtual Environment Flight Training to Stressful Flight Environments.

Procedures. The purpose of this research is to investigate the effects that stress training has on stressful flight operations in order to mitigate the human factors preconditions to aircraft accidents. Stress training implementation strategies are also investigated in order to develop pedagogy pertinent to stress training. Participants will be asked to perform real flying tasks in a virtual flight simulator while under the stress of cold. The cold stressor will be applied to participants' feet.

My participation in this experiment will last approximately one and a half hours. The experiment will consist of verbal instructions, practical instructions, training sessions, and a final criterion trial. During stress exposure, my foot will be submerged in 3 inches of cold water.

I understand that prior to my participation, physiological sensors will be connected to my body. These sensors will measure my body's responses to the tasks I will perform. The data collected by these sensors will not be analyzed for any purpose other than this study, and any findings in this study can in no way disqualify me for military service.

Due to the immersive nature of the simulator which may introduce slight simulator sickness and the discomfort associated with the cold stressor, I will be asked before, during and after the experiment if I feel ill in any way. Potential illnesses I may experience include nausea, dizziness, headache, numbing of the limbs, and skin discomfort. If at any time I experience any of these symptoms, I will notify the experimenter immediately.

Risks and Benefits. In addition to the risks mentioned above, this study will include no more risks than those associated with using a computer or playing a computer-based video game. I am aware of these risks, and I will receive no benefits for participating in this study other than gaining knowledge concerning piloting an aircraft.

Compensation. I understand that no tangible compensation will be given. I understand that a copy of the research results will be available at the conclusion of the experiment by contacting the PI (memccaul@nps.edu, 831-656-2191) or the experimenter (ckmccler@nps.edu, 719-649-3062).

Confidentiality & Privacy Act. I understand that all records of this study will be kept confidential and that my privacy will be safeguarded. No information will be publicly accessible which could identify me as a participant. I will be identified only as a code number on all research forms/data bases. My name on any signed document will not be paired with my code number in order to protect my identity. I understand that records of my participation will be maintained by NPS for three years, after which they will be destroyed.

Voluntary Nature of the Study. I understand that my participation is strictly voluntary, and if I agree to participate, I am free to withdraw at any time without prejudice. If I am uncomfortable participating at any time during the experiment I may withdraw without any retribution.

Points of Contact. I understand that if I have any questions or comments regarding this project upon the completion of my participation, I should contact the Principal Investigator, Dr. Michael McCauley, 831-656-2191, memccaul@nps.edu. Any medical questions should be addressed to Col George Patrin, MD, USA, (CO, POM Medical Clinic), (831) 242-7550, george.patrin@us.army.mil. Any other questions or concerns may be addressed to the IRB Chair, LT Brent Olde, 656-3807, baolde@nps.edu.

Statement of Consent. I have been provided with a full explanation of the purpose, procedures, and duration of my participation in this research project. I understand how my identification will be safeguarded and have had all my questions answered. I have been provided a copy of this form for my records and I agree to participate in this study. I understand that by agreeing to participate in this research and signing this form, I do not waive any of my legal rights.

______________________________	______________
Participant’s Signature	Date
______________________________	______________
Researcher’s Signature	Date

APPENDIX C2: INFORMED CONSENT FORM (EXP 2, CONTROL)

Introduction. You are invited to participate in a study entitled Training Transfer from Virtual Environment Flight Training to Flight Environments.

Procedures. The purpose of this research is to investigate the effects that training has on flight operations in order to mitigate the human factors preconditions to aircraft accidents. Participants will be asked to perform flying tasks in a virtual flight simulator, followed by performing similar flying tasks in a real aircraft.

My participation in this experiment will last approximately two hours. The experiment will consist of verbal instructions, practical instructions, training sessions, and an aircraft flight.

I understand that prior to my participation, physiological sensors will be connected to my body. These sensors will measure my body's responses to the tasks I will perform. The data collected by these sensors will not be analyzed for any purpose other than this study, and any findings in this study can in no way disqualify me for military service.

I may experience slight simulator sickness due to the immersive nature of the simulator and disorientation during flight. I will be asked before, during and after the experiment if I feel ill in any way. Potential illnesses I may experience include nausea, dizziness, headache, and skin discomfort. If at any time I experience any of these symptoms, I will notify the experimenter immediately.

Risks and Benefits. In addition to the risks mentioned above, this study includes the risks associated with using a computer or playing a computer-based video game. I will also operate a general aviation aircraft under the supervision of a certified flight instructor. To mitigate some of the risks associated with piloting an aircraft, I will not be allowed to manipulate the aircraft's controls below 2,000 feet above the ground (including take-offs and landings). In addition, if at any time the aircraft approaches an unsafe attitude, the instructor will immediately regain control. The benefits of my participation include receiving both ground and flight aircraft training.

Compensation. I understand that no tangible compensation will be given. The flight training I receive is valued at approximately $100. I understand that a copy of the research results will be available at the conclusion of the experiment by contacting the PI (memccaul@nps.edu, 831-656-2191) or the experimenter (ckmccler@nps.edu, 719-649-3062).

Confidentiality & Privacy Act. I understand that all records of this study will be kept confidential and that my privacy will be safeguarded. No information will be publicly accessible which could identify me as a participant. I will be identified only as a code number on all research forms/data bases. My name on any signed document will not be paired with my code number in order to protect my identity. I understand that records of my participation will be maintained by NPS for three years, after which they will be destroyed.

Voluntary Nature of the Study. I understand that my participation is strictly voluntary, and if I agree to participate, I am free to withdraw at any time without prejudice. If I withdraw from the experiment while in flight, the instructor will immediately return to the airport. If I am uncomfortable participating at any time during the experiment I may withdraw without any retribution.

Points of Contact. I understand that if I have any questions or comments regarding this project upon the completion of my participation, I should contact the Principal Investigator, Dr. Michael McCauley, 831-656-2191, memccaul@nps.edu. Any medical questions should be addressed to Col George Patrin, MD, USA, (CO, POM Medical Clinic), (831) 242-7550, george.patrin@us.army.mil. Any other questions or concerns may be addressed to the IRB Chair, LCDR Paul O'Connor, 656-3864, peoconno@nps.edu.

Statement of Consent. I have been provided with a full explanation of the purpose, procedures, and duration of my participation in this research project. I understand how my identification will be safeguarded and have had all my questions answered. I have been provided a copy of this form for my records and I agree to participate in this study. I understand that by agreeing to participate in this research and signing this form, I do not waive any of my legal rights.

____________________________________ ____________________

Participant’s Signature Date

____________________________________ ____________________

Researcher’s Signature Date

APPENDIX C3: INFORMED CONSENT FORM (EXP 2, TREATMENT)

Introduction. You are invited to participate in a study entitled Stress Training Transfer from Virtual Environment Flight Training to Stressful Flight Environments.

Procedures. The purpose of this research is to investigate the effects that stress training has on stressful flight operations in order to mitigate the human factors preconditions to aircraft accidents. Stress training implementation strategies are also investigated in order to develop pedagogy pertinent to stress training. Participants will be asked to perform flying tasks in a virtual flight simulator while under the stress of cold. The cold stressor will be applied to participants' feet. Participants will then be asked to perform similar flying tasks in a real aircraft.

My participation in this experiment will last approximately two hours. The experiment will consist of verbal instructions, practical instructions, training sessions, and an aircraft flight. During stress exposure, my foot will be submerged in 3 inches of cold water.

I understand that prior to my participation, physiological sensors will be connected to my body. These sensors will measure my body's responses to the tasks I will perform. The data collected by these sensors will not be analyzed for any purpose other than this study, and any findings in this study can in no way disqualify me for military service.

I may experience (1) slight simulator sickness due to the immersive nature of the simulator, (2) discomfort associated with the cold stressor, and (3) disorientation during flight. I will be asked before, during and after the experiment if I feel ill in any way. Potential illnesses I may experience include nausea, dizziness, headache, numbing of the limbs, and skin discomfort. If at any time I experience any of these symptoms, I will notify the experimenter immediately.

Risks and Benefits. In addition to the risks mentioned above, this study includes the risks associated with using a computer or playing a computer-based video game. I will also operate a general aviation aircraft under the supervision of a certified flight instructor. To mitigate some of the risks associated with piloting an aircraft, I will not be allowed to manipulate the aircraft's controls below 2,000 feet above the ground (including take-offs and landings). In addition, if at any time the aircraft approaches an unsafe attitude, the instructor will immediately regain control. The benefits of my participation include receiving both ground and flight aircraft training.

Compensation. I understand that no tangible compensation will be given. The flight training I receive is valued at approximately $100. I understand that a copy of the research results will be available at the conclusion of the experiment by contacting the PI (memccaul@nps.edu, 831-656-2191) or the experimenter (ckmccler@nps.edu, 719-649-3062).

Confidentiality & Privacy Act. I understand that all records of this study will be kept confidential and that my privacy will be safeguarded. No information will be publicly accessible which could identify me as a participant. I will be identified only as a code number on all research forms/data bases. My name on any signed document will not be paired with my code number in order to protect my identity. I understand that records of my participation will be maintained by NPS for three years, after which they will be destroyed.

Voluntary Nature of the Study. I understand that my participation is strictly voluntary, and if I agree to participate, I am free to withdraw at any time without prejudice. If I withdraw from the experiment while in flight, the instructor will immediately return to the airport. If I am uncomfortable participating at any time during the experiment I may withdraw without any retribution.

Points of Contact. I understand that if I have any questions or comments regarding this project upon the completion of my participation, I should contact the Principal Investigator, Dr. Michael McCauley, 831-656-2191, memccaul@nps.edu. Any medical questions should be addressed to Col George Patrin, MD, USA, (CO, POM Medical Clinic), (831) 242-7550, george.patrin@us.army.mil. Any other questions or concerns may be addressed to the IRB Chair, LCDR Paul O'Connor, 656-3864, peoconno@nps.edu.

Statement of Consent. I have been provided with a full explanation of the purpose, procedures, and duration of my participation in this research project. I understand how my identification will be safeguarded and have had all my questions answered. I have been provided a copy of this form for my records and I agree to participate in this study. I understand that by agreeing to participate in this research and signing this form, I do not waive any of my legal rights.

______________________________________ ____________________

Participant’s Signature Date

______________________________________ ____________________

Researcher’s Signature Date

APPENDIX D: COHEN PERCEIVED STRESS SCALE

The questions in this scale ask you about your feelings and thoughts during the last month. In each case, you will be asked to indicate how often you felt or thought a certain way. Although some of the questions are similar, there are differences between them and you should treat each one as a separate question. The best approach is to answer each question fairly quickly. That is, don't try to count up the number of times you felt a particular way, but rather indicate the alternative that seems like a reasonable estimate.

For each question choose from the following alternatives:

0 = Never
1 = Almost Never
2 = Sometimes
3 = Fairly Often
4 = Very Often

	0	1	2	3	4
1. In the last month, how often have you been upset because of something that happened unexpectedly?	O	O	O	O	O
2. In the last month, how often have you felt that you were unable to control the important things in your life?	O	O	O	O	O
3. In the last month, how often have you felt nervous and "stressed"?	O	O	O	O	O
4. In the last month, how often have you felt confident about your ability to handle your personal problems?	O	O	O	O	O
5. In the last month, how often have you felt that things were going your way?	O	O	O	O	O
6. In the last month, how often have you found that you could not cope with all the things that you had to do?	O	O	O	O	O
7. In the last month, how often have you been able to control irritations in your life?	O	O	O	O	O
8. In the last month, how often have you felt that you were on top of things?	O	O	O	O	O
9. In the last month, how often have you been angered because of things that were outside of your control?	O	O	O	O	O
10. In the last month, how often have you felt difficulties were piling up so high that you could not overcome them?	O	O	O	O	O

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX E1: EX. 1 PARTICIPANT INSTRUCTIONS

VERBAL INSTRUCTIONS

Experiment checklist COMPLETE.
Welcome participant (note time). Encourage them to use the restroom.

1. Cell phones OFF.

2. Read/complete the Informed Consent Form.

3. Complete the Participant Survey, Cohen Stress Scale, and Enter Survey.

4. Hook up physiological measures and test.

Following are the verbal instructions which will be read to the participant.

INTRODUCTION:
We are about to begin the first training phase of the experiment. During this time you will be read instructions regarding the use of the provided flight controls and instruments for aviation navigation. The aircraft you will be flying in this experiment is a Cessna 172, and the instruments will look and act like a real Cessna 172.

CONTROLS:
The aircraft you will be flying will be controlled with the yoke. Please practice as I describe each control input. To climb, gently pull back on the yoke. To descend gently push forward on the yoke. To turn right, gently turn the yoke to the right. And to turn left, gently turn the yoke to the left. A banked aircraft produces less lift, and the yoke must also be pulled back slightly to maintain level flight in a turn. To descend during a turn, you may find that the yoke does not need to be pulled back at all, or very little. To climb during a turn, you will need to pull back on the yoke more than a level turn. Please practice these actions now.

Un-pause sim and let them practice on their own for 1 minute.(NO HEADSET.)

After, Sim —> Session.

Do you have any questions regarding the flight controls?

INSTRUMENTS:

The primary flight instruments include the attitude indicator, altimeter, and heading indicator. You will uses these three instruments for the navigation tasks. Let us take time to learn the functions of each instrument. Following this short instruction, you will take a quiz to test you knowledge of this information.

Attitude Indicator: The attitude indicator is your primary reference for changes in pitch and bank. Degrees of pitch are marked on the center of the attitude indicator. The first long line with a number next to it is 10-degrees of pitch, and the next longer line with a number is 20-degrees of pitch. For the purposes of these simulated flights, any pitch change should occur at 10-degrees-up to climb, or 10-degrees-down to descend. The attitude indicator also depicts bank angle using the peripheral marks. The first, small mark is 10-degrees of bank, the second small mark is 20-degrees of bank, and the third mark which is bold is 30-degrees of bank. For the purpose of your flights, you will be asked to maintain 20-degrees of bank for all turns. During all level flight, you will be required to maintain 0-degrees of pitch and bank. Do you have any questions about using the attitude indicator?

Altimeter: The altimeter will accurately indicate the altitude you will be flying at. The small hand indicates thousands of feet, and the large hand indicates hundreds of feet. The current altitude on the altimeter is 7000 feet. During your flights you will be asked to climb or descend to specific altitudes. When told to do so, climb or descend at 10-degrees of pitch to the specified altitude. During all other phases of flight, you must maintain your assigned altitude. For these level portions of flight, the altimeter will be your primary source of pitch information. Do you have any questions about using the altimeter?

Heading Indicator: The heading indicator provides your aircraft's heading relative to the cardinal directions: North, East, South, West (pointing to each). This is your primary instrument for maintaining level bank—if you drift from an assigned heading, you are turning and you must correct. During your flights, you will be instructed to turn a direction (left or right) to a specified heading. For example, "Turn right heading 300." This clearance asks you to turn to the right until you are aligned with the numbers "3-0" on the heading indicator—notice that the last zero is removed from headings on the indicator. During level flight, you must maintain your assigned heading; for these level portions of flight, the heading indicator will be your primary source of bank information. Do you have any question about the heading indicator?

Do you have any questions regarding the instruments?

RECAP:

To recap, during straight and level flight, the altimeter and heading indicator are your *primary* instruments for maintaining pitch and bank, and the yoke will be relatively neutral (although small inputs may be needed). For climbs and descents, use the attitude indicator to maintain 10-degrees of pitch and the altimeter to judge reference to your target altitude while pushing or pulling the yoke in the correct direction. For turns, use the attitude indicator to maintain 20-degrees of bank and the heading indicator to judge reference to your target heading while gently turning the yoke as needed. However, do keep in mind that all three instruments will need to be referenced for all portions of flight.

Do you have any questions regarding the controls, instruments, or how you will use them?

Headsets On.

QUIZ:
You will now take a short quiz regarding the training you just received.

During turns, which instrument will primarily be used to maintain bank?
Attitude Indicator

How many degrees of bank must you maintain during turns?
20

During climbs, which instrument will primarily be used to maintain pitch?
Attitude Indicator

How many degrees of pitch must you maintain during climbs?
10

During straight and level flight, which instrument will *primarily* be used to maintain pitch?
Altimeter

During straight and level flight, which instrument will *primarily* be used to maintain bank?
Heading Indicator

VIDEO:
BioGraph RECORD

You will now watch a short film describing the functions of the instruments I just described to you. The video will also address some instruments and procedures which you have not learned and you are not expected to use for this experiment—keep in mind during the video the three instruments which you have learned, and pay particularly close attention to the operations pertaining to only those three instruments. Please reference your instruments as needed during the film.

BioGraph "V." Play video (:28-5:00). Turn off lamp.

BioGraph "E." Do you have any questions concerning the video?

PRACTICAL INSTRUCTIONS

Audacity — RECORD.

You have now learned about the flight instruments and how they will be used. During this practice session, I will give you heading and altitude instructions similar to what you will receive during the test events. During your maneuvers you are required to maintain the following parameters:

During level flight:
+/- 50 feet altitude
+/- 5 degrees heading

During turns:
20-degrees bank +/- 5 degrees

During climbs:
10-degrees pitch +/- 3 degrees

If you have any questions during this phase of training feel free to ask.

The simulator will begin at a heading of 180 and 7000 feet. Are you ready?

1.) Maintain straight and level flight.

BioGraph — "1"
Sim — UNPAUSE
Read the clearances from the collection sheet.
BioGraph — "E," Sim — PAUSE.

This concludes your training. Do you have any questions regarding these navigation tasks?

SUBJECTIVE RATING ASSESSMENT:
During the following sessions, you will also be asked to rate your stress levels during sessions using a 10-point scale. "1" is not stressed at all, similar to a peaceful, relaxing afternoon. "10" is the most possible stress you can withstand; for this experiment that would be the stress you can withstand before asking to terminate the experiment. The queries will refer to the most recent maneuver that you flew. Please now answer the following sample query for your current stress level:

"Rate your stress on a scale from 1 to 10."

Do you have any questions concerning the stress queries?

TASK ACQUISITION

You are about to begin the first test session. This session will be similar to the previous training session you just accomplished, and it will last approximately 10 minutes. You may ask to have a clearance re-read. Please wait to ask any other questions you may have until after the session, although do let me know if you become ill, are very confused, or need to stop for any other reason.

To review, you are required to maintain the following parameters:

During level flight:
+/- 50 feet altitude
+/- 5 degrees heading

During turns:
20-degrees bank +/- 5 degrees

During climbs:
10-degrees pitch +/- 3 degrees

Do you have any questions before we begin?

The simulator will begin at a heading of 180 and 7000 feet. Are you ready?

1.) Maintain straight and level flight.

BioGraph — "2"
Sim — UNPAUSE
Read the clearances from the collection sheet.
BioGraph — "E," Sim — PAUSE.

This concludes the first test session. Do you need a break before we proceed with the next session?

STRESS EXPOSURE

For participants in the treatment group, PAUSE BioGraph, have them remove their left shoe/sock, check temperature, relocate the bucket, and UNPAUSE BioGraph, and read the following instructions.

For the following test session, you will be asked to perform the flying task during exposure to a stressful cold pressor. The cold pressor will be a cold bucket of ice water that you will put your left foot into. During exposure to the cold pressor, or any stress, it is important to, first, maintain your normal breathing as best as possible. This will help calm and relax you. Next, attempt to focus on the task at hand, and ignore the distractions of the stressor. Finally, pay especially close attention to the performance parameters which you are asked to fly. These will be read to you again.

Now, please place your left foot in the ice water.

"I." Wait one minute.

Notice how your breathing has increased. For the next few minutes focus on slowing and regulating your breathing while attempting to relax.

Wait one minute.

Now, attempt to ignore the stress of the cold and focus your attention on the flying task you have been taught. Please reference your flight instruments as you attempt to do this.

Wait one minute.

Finally, visualize the performance parameters you will have to maintain during the next session. They are:

During level flight:
+/- 50 feet altitude
+/- 5 degrees heading
During turns:
20-degrees bank +/- 5 degrees
During climbs:
10-degrees pitch +/- 3 degrees

Wait one minute.

You may now remove your foot from the ice water and dry it off.

BioGraph — "E," PAUSE. Allow 5 min. for their foot to return to normal temperature.

TA2/PUS

You are about to begin the second test session. This session will be similar to the previous test session you just accomplished, and it will last approximately 10 minutes. You may ask to have a clearance re-read. Please wait to ask any other questions you may have until after the session, although do let me know if you become ill, are very confused, or need to stop for any other reason.

To review, you are required to maintain the following parameters:

During level flight:
+/- 50 feet altitude
+/- 5 degrees heading

During turns:
20-degrees bank +/- 5 degrees

During climbs:
10-degrees pitch +/- 3 degrees

Do you have any questions before we begin?

If a participant is in the treatment group, UNPAUSE BioGraph, and apply the cold pressor now.

The simulator will begin at a heading of 180 and 7000 feet. Are you ready?

1.) Maintain straight and level flight.

BioGraph — "3"
Sim — UNPAUSE
Read the clearances from the collection sheet.
BioGraph — "E," Sim — PAUSE.

This concludes the second test session. Please take a short brake and relax.

BioGraph — PAUSE.

Allow a treatment group participant 5 min. to warm their foot.

CRITERION

If participant is in control group, have them remove their left shoe/sock, and move the bucket closer.
UNPAUSE BioGraph for anticipatory response.

You are about to begin the third and final test session. This session will be similar to the previous test sessions you have accomplished, and it will last approximately 10 minutes. You may ask to have a clearance re-read. Please wait to ask any other questions you may have until after the session, although do let me know if you become ill, are very confused, or need to stop for any other reason.

To review, you are required to maintain the following parameters:

During level flight:
+/- 50 feet altitude
+/- 5 degrees heading

During turns:
20-degrees bank +/- 5 degrees

During climbs:
10-degrees pitch +/- 3 degrees

Do you have any questions before we begin?

Apply the cold pressor now.

The simulator will begin at a heading of 180 and 7000 feet. Are you ready?

1.) Maintain straight and level flight.

BioGraph — "4"
Sim — UNPAUSE
Read the clearances from the collection sheet.
BioGraph — "E," "STOP"
Sim — PAUSE
Audacity — STOP

This concludes the experiment.

Help participant with foot, headset, and electrodes. Read de-brief.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX E2: EX. 2 & 3 PARTICIPANT INSTRUCTIONS

Experiment 3 no-training group does not receive the marked training

VERBAL INSTRUCTIONS

Experiment checklist COMPLETE.
Welcome participant (note time).

1. Cell phones OFF.

2. Read/complete the Informed Consent Form, Participant Survey, Audio Recording, and Entry Survey.

3. Hook up physiological measures.

Following are the verbal instructions which will be read to the participant.

Open calibration screen in x-plane. Set neutral yoke setting.

INTRODUCTION:
We are about to begin the first training phase of the experiment. During this time you will be read instructions regarding the use of the provided flight controls and instruments for aviation navigation. The aircraft you will be flying in this experiment is a Piper Archer, and the instruments will look and act like a real Piper Archer.

CONTROLS:
The aircraft you will be flying will be controlled with the yoke. Please practice as I describe each control input. To climb, gently pull back on the yoke. To descend gently push forward on the yoke. To turn right, gently turn the yoke to the right. And to turn left, gently turn the yoke to the left. A banked aircraft produces less lift, and the yoke must also be pulled back slightly to maintain level flight in a turn. To descend during a turn, you may find that the yoke does not need to be pulled back at all, or very little. To climb during a turn, you will need to pull back on the yoke more than a level turn.

Please practice these actions now using the full range of motion.

Set neutral yoke setting. Un-pause sim and let them practice on their own for 1 minute. (NO HEADSET.)

After, Sim —> Instructions

Do you have any questions regarding the flight controls?

INSTRUMENTS:
The primary flight instruments include the attitude indicator, altimeter, and heading indicator. You will uses these three instruments for the navigation tasks. Let us take time to learn the functions of each instrument. Following this short instruction, you will take a quiz to test your knowledge of this information.

Attitude Indicator: The attitude indicator is your primary reference for changes in pitch and bank. Degrees of pitch are marked on the center of the attitude indicator. The first long line with a number next to it is 10-degrees of pitch, and the next long line with a number is 20-degrees of pitch. For the purposes of these simulated flights you will attempt to maintain zero degrees of pitch at all times. The attitude indicator also depicts bank angle using the peripheral marks. The first, small mark is 10-degrees of bank, the second small mark is 20-degrees of bank, and the third longer mark is 30-degrees of bank. For the purpose of your flights, you will be asked to maintain 20-degrees of bank for all turns. During all level flight, you will be required to maintain 0-degrees of pitch and bank. Do you have any questions about using the attitude indicator?

Altimeter: The altimeter will accurately indicate the altitude you will be flying at. The strip indicates your relative altitude and the window indicates your specific altitude. During your flights you will be asked to maintain a specific altitude, and you will use the altimeter to do so. Do you have any questions about using the altimeter?

Heading Indicator: The heading indicator provides your aircraft's heading relative to the cardinal directions: North, East, South, West (pointing to each). This is your primary instrument for maintaining level bank—if you drift from an assigned heading, you are turning and you must correct. During your flights, you will be instructed to turn a direction (left or right) to a specified heading. For example, "Turn left heading 300." This clearance asks you to turn to the left until you are aligned with the numbers "3-0" on the heading indicator—notice that the last zero is removed from headings on the indicator. A pink "bug," currently located on North, will indicate your target heading. During level flight, you must maintain your assigned heading; for these level portions of flight, the heading indicator will be your primary source of bank information. Do you have any question about the heading indicator?

Do you have any questions regarding the instruments?

RECAP:
To recap, during straight and level flight, the altimeter and heading indicator are your *primary* instruments for maintaining pitch and bank, and the yoke will be relatively neutral (although small inputs may be needed). For turns, use the attitude indicator to maintain 20-degrees of bank and the heading indicator to judge reference to your target

heading while gently turning the yoke as needed. However, do keep in mind that all three instruments will need to be referenced for all portions of flight.

Do you have any questions regarding the controls, instruments, or how you will use them?

Headsets On.
Audio — Connect to mixer.

QUIZ:
You will now take a short quiz regarding the training you just received.

During turns, which instrument will primarily be used to maintain bank?
Attitude Indicator

How many degrees of bank must you maintain during turns?
20

During straight and level flight, which instrument will *primarily* be used to maintain pitch?
Altimeter

During straight and level flight, which instrument will *primarily* be used to maintain bank?
Heading Indicator

VIDEO:
BioGraph — Turn ON

You will now watch a short film showing the instruments I just described to you. The video addresses some instruments and procedures which you have not learned and you are not expected to use for this experiment—keep in mind during the video the three instruments which you have learned, and pay particularly close attention to the operations pertaining to only those three instruments.

BioGraph Mark. Play video. Turn off lamp.

BioGraph Mark. Do you have any questions concerning the video?

PRACTICAL INSTRUCTIONS

You have now learned about the flight instruments and how they will be used. During this practice session, I will give you heading instructions similar to what you will receive during the test events. You will also receive altitude instructions to prepare you for maintaining pitch control. During your maneuvers you are required to maintain the following parameters:

During level flight:
+/- 50 feet altitude
+/- 5 degrees heading

During turns:
20-degrees bank +/- 5 degrees

If you have any questions during this phase of training feel free to ask.

The simulator will begin at a heading of 360 and 5000 feet. Are you ready?

1.) Maintain straight and level flight.

BioGraph — MARK
Sim — UNPAUSE
Read the clearances from the collection sheet.
BioGraph — MARK, Sim — PAUSE.

This concludes your training. Do you have any questions regarding these navigation tasks?

Sim —> Experiment2

SUBJECTIVE RATING ASSESSMENT:
During the following sessions, you will also be asked to rate your stress levels during sessions using a 10-point scale. "1" is not stressed at all, similar to a peaceful, relaxing afternoon. "10" is the most possible stress you can withstand; for this experiment that would be the stress you can withstand before asking to terminate the experiment. The queries will refer to the most recent maneuver that you flew. Please now answer the following sample query for your current stress level:

"Rate your stress on a scale from 1 to 10."

Do you have any questions concerning the stress queries?

TASK ACQUISITION

Sim —> Add "chatter"
EFIS —> Add BRG

You are about to begin the first test session. This session will be similar to the previous training session you just accomplished, and it will last approximately 10 minutes. You may ask to have a clearance re-read. Please wait to ask any other questions you may have until after the session, although do let me know if you become ill, are very confused, or need to stop for any other reason.

To review, you are required to maintain the following parameters:

During level flight:
+/- 50 feet altitude
+/- 5 degrees heading

During turns:
20-degrees bank +/- 5 degrees

Do you have any questions before we begin?

The simulator will begin at a heading of 180 and 5000 feet. Are you ready?

1.) Maintain straight and level flight.

BioGraph — MARK
Sim — UNPAUSE
Read the clearances from the collection sheet.
BioGraph — MARK, Sim — PAUSE.

This concludes the first test session. Do you need a break before we proceed with the next session?

STRESS EXPOSURE

For participants in the treatment group, PAUSE BioGraph, have them remove their left shoe/sock, check temperature, relocate the bucket, and UNPAUSE BioGraph, and read the following instructions.

For the following test session, you will be asked to perform the flying task during exposure to a stressful cold pressor. The cold pressor will be a cold bucket of ice water that you will put your left foot into. During exposure to the cold pressor, or any stress, it is important to, first, maintain your normal breathing as best as possible. This will help calm and relax you. Next, attempt to focus on the task at hand, and ignore the distractions of the stressor. Finally, pay especially close attention to the performance parameters which you are asked to fly. These will be read to you again.

Now, please place your left foot in the ice water.

BioGraph — MARK
Wait one minute.

Notice how your breathing has increased. For the next few minutes focus on slowing and regulating your breathing while attempting to relax.

Wait one minute.

Now, attempt to ignore the stress of the cold and focus your attention on the flying task you have been taught. Please reference your flight instruments as you attempt to do this.

Wait one minute.

Finally, visualize the performance parameters you will have to maintain during the next session. They are:

During level flight:
+/- 50 feet altitude
+/- 5 degrees heading

During turns:
20-degrees bank +/- 5 degrees

Wait one minute.

You may now remove your foot from the ice water and dry it off.

BioGraph — MARK. Allow 5 min. for their foot to return to normal temperature.

TA2/PUS

You are about to begin the second test session. This session will be similar to the previous test session you just accomplished, and it will last approximately 10 minutes. You may ask to have a clearance re-read. Please wait to ask any other questions you may have until after the session, although do let me know if you become ill, are very confused, or need to stop for any other reason.

To review, you are required to maintain the following parameters:

During level flight:
+/- 50 feet altitude
+/- 5 degrees heading

During turns:
20-degrees bank +/- 5 degrees

Do you have any questions before we begin?

If a participant is in the treatment group, apply the cold pressor now.

The simulator will begin at a heading of 180 and 5000 feet. Are you ready?

1.) Maintain straight and level flight.

BioGraph — MARK
Sim — UNPAUSE
Read the clearances from the collection sheet.
BioGraph — MARK, Sim — PAUSE.

This concludes the second test session.

BioGraph — PAUSE.

Allow treatment group to put on sock and shoe.

TRANSFER TASK PREP

We are now ready to begin the transfer task in the aircraft.

This session will be similar to the previous test sessions you just accomplished, although it will take place in a real aircraft.

Now we will review some of the operations regarding the aircraft.

During your flight, you will sit in the front left seat with a yoke in front of you similar to what is in front of you now. A pilot will sit to your right. I will be in the back seat of the aircraft. Do not engage in any conversation with either the pilot or me until after the test session. During the session, you will be read clearances similar to the previous sessions. You may ask to have a clearance re-read. Please wait to ask any other questions you may have until after the session, although do let me know if you become ill, are very confused, or need to stop for any other reason.

It is important that we practice positive change of aircraft control during your flight. While we are en route to the practice area, you must place your hands in your lap. When aircraft control is given to you, the pilot will say, "You have the aircraft," at which time you must repeat, "I have the aircraft" and place your hands on the yoke. Likewise, if the pilot regains control of the aircraft at any time, he will say "I have the aircraft," at which point you must immediately place your hands in your lap and say "You have the aircraft."

Let's practice now:
Place your hands in your lap.

You have the aircraft.

Participant must state "I have the aircraft" and place their hands on the yoke
I have the aircraft.

Participant must state "You have the aircraft" and place their hands on the yoke

On the way to the practice area you will dawn glasses which will prevent you from seeing the instrument panel. Please focus your attention outside the aircraft and do not attempt to view any of the aircraft instruments. After approximately 10 minutes, we will arrive at the practice area where you will be given a second set of glasses which will restrict your vision to a screen identical to the one you have been trained on. Please focus your attention on this display, and do not attempt to view any other instruments or the outside view.

Do you have any questions regarding the flight operations?
1. Pause BioGraph, collect electrodes
2. MNFC Covenant not to Sue, AF 1585, FIF
3. Follow experimenter checklist

APPENDIX F: SUBJECTIVE AROUSAL ASSESSMENT

Stress Survey

Participant ID:
Date/Time:
Experiment:

	1 Pilot	1a	1b	2 Pilot	2	3 Pilot	3
Choose:	☐	☐	☐	☐	☐	☐	☐

Condition:

	TA	PUS1	PUS2 (Experiment 2)
Choose:	☐	☐	☐

TASK ACQUISITION. Rate how MENTALLY stressed you currently are.

	Not at all stressful (1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	Most possible stress (10)
Query 1, Time:	○	○	○	○	○	○	○	○	○	○
Query 2, Time:	○	○	○	○	○	○	○	○	○	○
Query 3, Time:	○	○	○	○	○	○	○	○	○	○
Query 4, Time:	○	○	○	○	○	○	○	○	○	○
Query 5, Time:	○	○	○	○	○	○	○	○	○	○
Overall Session Stress:	○	○	○	○	○	○	○	○	○	○

TA2/PUS. Rate how MENTALLY stressed you currently are.

	Not at all stressful (1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	Most possible stress (10)
Query 1, Time:	○	○	○	○	○	○	○	○	○	○
Query 2, Time:	○	○	○	○	○	○	○	○	○	○
Query 3, Time:	○	○	○	○	○	○	○	○	○	○
Query 4, Time:	○	○	○	○	○	○	○	○	○	○
Query 5, Time:	○	○	○	○	○	○	○	○	○	○
Overall Session Stress:	○	○	○	○	○	○	○	○	○	○

TRANSFER TASK. Rate how MENTALLY stressed you currently are.

	Not at all stressful (1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	Most possible stress (10)
Query 1, Time:	○	○	○	○	○	○	○	○	○	○
Query 2, Time:	○	○	○	○	○	○	○	○	○	○
Query 3, Time:	○	○	○	○	○	○	○	○	○	○
Query 4, Time:	○	○	○	○	○	○	○	○	○	○
Query 5, Time:	○	○	○	○	○	○	○	○	○	○
Overall Session Stress:	○	○	○	○	○	○	○	○	○	○

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX G1: PARTICIPANT EXIT SURVEY (EXPERIMENT 1)

Entry Questions: Rate how you feel in the following areas.

	Not at all ill (1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	Most possible illness (10)
Nauseous:										
Dizzy:										
Blurred vision:										
Tumbling sensation:										
Light-headed:										
Disoriented:										
Headache:										

Exit Questions: Rate how you feel in the following areas.

	Not at all ill (1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	Most possible illness (10)
Nauseous:										
Dizzy:										
Blurred vision:										
Tumbling sensation:										
Light-headed:										
Disoriented:										
Headache:										

The experiment you just participated in was a study to test the effects of stress exposure training on performance in a stressful flight environment. Your participation and the results of this study will greatly help both the military and civilian aviation communities better understand how to train pilots. Your specific results in this study will be kept strictly confidential, and your name will not be associated with any of the data you provided today. If you have any questions or concerns regarding this study at any point in the future, you may contact me (provide a business card). You may also contact me if you would like to see the results of this study.

Do you have any questions regarding this study?

Finally, it is absolutely imperative that you not discuss the details of this study with any other friends, family, or colleagues. Please do not discuss the tasks, procedures, or apparatus to anyone. This information can greatly bias future participants, many of which will be people you know.

Thank you for your participation.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX G2: PARTICIPANT EXIT SURVEY (EXP. 2 & 3)

Entry/Exit Survey

Entry Questions 1: Rate how susceptible you are to the following conditions.

	Not at all susceptible (1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	Highly susceptible (10)
Fear of heights:	○	○	○	○	○	○	○	○	○	○
Fear of flying:	○	○	○	○	○	○	○	○	○	○
Motion sickness:	○	○	○	○	○	○	○	○	○	○

Entry Questions 2: Rate how you feel in the following areas.

	Not at all ill (1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	Most possible illness (10)
Nauseous:	○	○	○	○	○	○	○	○	○	○
Dizzy:	○	○	○	○	○	○	○	○	○	○
Blurred vision:	○	○	○	○	○	○	○	○	○	○
Tumbling sensation:	○	○	○	○	○	○	○	○	○	○
Light-headed:	○	○	○	○	○	○	○	○	○	○
Disoriented:	○	○	○	○	○	○	○	○	○	○
Headache:	○	○	○	○	○	○	○	○	○	○

Exit Questions: Rate how you feel in the following areas.

	Not at all ill (1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	Most possible illness (10)
Nauseous:	○	○	○	○	○	○	○	○	○	○
Dizzy:	○	○	○	○	○	○	○	○	○	○
Blurred vision:	○	○	○	○	○	○	○	○	○	○
Tumbling sensation:	○	○	○	○	○	○	○	○	○	○
Light-headed:	○	○	○	○	○	○	○	○	○	○
Disoriented:	○	○	○	○	○	○	○	○	○	○
Headache:	○	○	○	○	○	○	○	○	○	○

The experiment you just participated in was a study to test the effects of stress exposure training on performance in a stressful flight environment. Your participation and the results of this study will greatly help both the military and civilian aviation communities better understand how to train pilots. Your specific results in this study will be kept strictly confidential, and your name will not be associated with any of the data you provided today. If you have any questions or concerns regarding this study at any point in the future, you may contact me (provide a business card). You may also contact me if you would like to see the results of this study.

Do you have any questions regarding this study?

Finally, it is absolutely imperative that you not discuss the details of this study with any other friends, family, or colleagues. Please do not discuss the tasks, procedures, or apparatus to anyone. This information can greatly bias future participants, many of which will be people you know.

Thank you for your participation.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX H: CFI EVALUATION

ID: _____ **Date:** __________

CFI Evaluation

During each session, the student will be asked to rate their current stress levels. At this time, rate the student's performance using the below scale. If any additional information is needed for a query (e.g., Student was throwing up), indicate that in the notes section. Following each session, rate the student's overall performance, and provide any overall comments if needed. Please do not let the student see this evaluation form.

1	2	3	4	5	6	7	8	9	10
Very Poor			**Average**				**Very Good**		

	Score	Notes
Query 1		
Query 2		
Query 3		
Query 4		
Query 5		
Query 6		

Overall Session Performance:

Comments:

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX I: RECRUITMENT FLYER

Research Volunteers Needed

Fly! Learn! Free!

Volunteers are needed to participate in an exciting, real-world research project. You will spend one hour in a flight simulator and one hour in an aircraft. Sign up fast before slots are filled!

16 Feb to 27 Mar 2009

<u>All Volunteers must meet the following criteria:</u>
--Active duty military less than 40 yrs old
--Zero time piloting any aircraft
--No air traffic control experience
--At least 20/30 corrected vision
--Individuals with severe fear of heights/flying or motion sickness are <u>not</u> qualified

THIS PAGE INTENTIIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California