NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

EFFECTS OF UAV SUPERVISORY CONTROL ON F-18 FORMATION FLIGHT PERFORMANCE IN A SIMULATOR ENVIRONMENT

by

Eric L. McMullen
Brian Shane Grass

March 2013

Thesis Advisor: Ji-Hyun Yang
Thesis Co-Advisor: Quinn Kennedy
Second Reader   Joseph Sullivan

This thesis was performed 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 March 2013	3. REPORT TYPE AND DATES COVERED Master's Thesis	
4. TITLE AND SUBTITLE EFFECTS OF UAV SUPERVISORY CONTROL ON F-18 FORMATION FLIGHT PERFORMANCE IN A SIMULATOR ENVIRONMENT		5. FUNDING NUMBERS	
6. AUTHOR(S) Eric L. McMullen and Brian Shane Grass			

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
  Naval Postgraduate School
  Monterey, CA 93943–5000
		8. PERFORMING ORGANIZATION REPORT NUMBER	

9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A		10. SPONSORING/MONITORING AGENCY REPORT NUMBER	

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. IRB Protocol number NPS.2012.0093-IR-EP7-A.

12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release distribution is unlimited	12b. DISTRIBUTION CODE A

13. ABSTRACT (maximum 200 words)

Continual advances in technology, along with increased cockpit workload— particularly the shift from two-seat to single-seat fighters to save money and reduce risk to life—push the limits of human mental capacity. Additionally, there is interest within the military aviation community to integrate Unmanned Aerial Vehicle (UAV) control into the cockpit in order to expand force projection capability.

This study compared the effects on formation flight performance of two different secondary tasks, specifically a traditional secondary task such as target prosecution with an electro-optical Forward Looking Infra-Red (FLIR) pod, and a futuristic secondary task such as UAV supervisory control.

A total of 34 military fighter aviators volunteered to fly three five-minuteF-18 simulator sessions in close formation with no secondary task, and then treated with each of the two secondary tasks.

Results provided clear indication that the futuristic task was significantly more challenging than the traditional task, and that both secondary tasks significantly increased the average mean following distance and variance compared to the undistracted flying baseline scenario. Additionally, we found no evidence that increased flight experience (total flight hours) significantly improved performance of the prescribed primary task when treated with the futuristic task distraction.

Knowledge gained from the results could contribute to improved crew resource management (CRM) and pilot workload management as well as flight safety resulting from the modification of flight procedures based on known effects of distractions in the cockpit.

14. SUBJECT TERMS Simulator, aviation, workload, task prioritization, distraction.			15. NUMBER OF PAGES 127
			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

THIS PAGE INTENTIONALLY LEFT BLANK

EFFECTS OF UAV SUPERVISORY CONTROL ON F-18 FORMATION FLIGHT PERFORMANCE IN A SIMULATOR ENVIRONMENT

Eric L. McMullen
Lieutenant Commander, United States Navy
B.S. & B.A. Virginia Tech, 1996

Brian Shane Grass
Major, United States Army
B.S., Texas Agricultural and Mechanical University, 1995

Submitted in partial fulfillment of the
requirements for the degree of

**MASTER OF SCIENCE IN
MODELING, VIRTUAL ENVIRONMENTS, AND SIMULATION (MOVES)**

from the

**NAVAL POSTGRADUATE SCHOOL
March 2013**

Authors: Eric L. McMullen
                  Brian Shane Grass

Approved by: Ji Hyun Yang
                  Thesis Advisor

                  Quinn Kennedy
                  Thesis Co-Advisor

                  Joseph Sullivan
                  Second Reader

                  Christian Darken
                  Chair, MOVES Academic Committee

                  Peter J. Denning
                  Chair, Department of Computer Science

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

Continual advances in technology, along with increased cockpit workload—particularly the shift from two-seat to single-seat fighters to save money and reduce risk to life—push the limits of human mental capacity. Additionally, there is interest within the military aviation community to integrate Unmanned Aerial Vehicle (UAV) control into the cockpit in order to expand force projection capability.

This study compared the effects on formation flight performance of two different secondary tasks, specifically a traditional secondary task such as target prosecution with an electro-optical Forward Looking Infra-Red (FLIR) pod, and a futuristic secondary task such as UAV supervisory control.

A total of 34 military fighter aviators volunteered to fly three five-minute F-18 simulator sessions in close formation with no secondary task, and then treated with each of the two secondary tasks.

Results provided clear indication that the futuristic task was significantly more challenging than the traditional task, and that both secondary tasks significantly increased the average mean following distance and variance compared to the undistracted flying baseline scenario. Additionally, we found no evidence that increased flight experience (total flight hours) significantly improved performance of the prescribed primary task when treated with the futuristic task distraction.

Knowledge gained from the results could contribute to improved crew resource management (CRM) and pilot workload management as well as flight safety resulting from the modification of flight procedures based on known effects of distractions in the cockpit.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF FIGURES

xii

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

AAR	After Action Review
AH	Attack Helicopter
AT-FLIR	Advanced Tactical Forward Looking Infra-Red
BN	Bombardier Navigator
CAS	Close Air Support
COP	Common Operational Picture
CRM	Cockpit Resource Management
FAC(A)	Forward Air Controller (Airborne)
FLIR	Forward Looking Infra-Red
fps	feet per second
FRS	Fleet Replacement Squadron
ft	feet
GPS	Global Positioning System
HLA	High Level Architecture
IRB	Institutional Review Board
kt	knot
kts	knots
MIT	Massachusetts Institute of Technology
MPCD	Multi-Purpose Control Display
MRAP	Mine Resistant Ambush Protected
NAS	Naval Air Station
NPS	Naval Postgraduate School
OE	Operational Environment
POC	Performance Operating Curve
RESCHU	Research Environment for Supervisory Control of Heterogeneous Unmanned Vehicles
SA	Situational Awareness
SAWSI	Semi-Autonomous Wingman Supervisory Interface
sd	standard deviation
SFWTI	Strike/Fighter Weapons and Tactics Instruction

section	A flight of two aircraft
TacOPS	Tactical Standard Operating Procedures
TEL	Transporter Ejector Launcher
TFH	Total Flight Hours
TMT	Trail Making Test
TOFT	Tactical Operational Flight Trainer
UAV	Unmanned Aerial Vehicle
USV	Unmanned Sea Vehicle
UUV	Unmanned Underwater Vehicle
WSO	Weapons Systems Operator

ACKNOWLEDGMENTS

David, please know that your love, and the love of Christ through you, is what has brought me through this. I only hope you know how much it means to me and that I can return it to you!

EXECUTIVE SUMMARY

Evaluating the effects of secondary tasks, or distractions, on the performance of a primary task has been previously well studied in an attempt to better understand the impacts of divided attention. However, no known studies have been conducted to assess the impact of distractions in a military aviation/cockpit environment. Constant advances in technology and increasing cockpit workload, particularly the shift from two-seat to single-seat fighters to save money and reduce risk to life, push the limits of human mental capacity. With the rapid increase in use of Unmanned Aerial Vehicles (UAVs) over the past decade, there is an interest within the military aviation community to integrate this capability into the cockpit in order to expand firepower/range/options.

This study compared the effects of two different secondary tasks on the formation flight performance. The two secondary tasks correspond to a traditional secondary task, such as target prosecution with an electro-optical Forward Looking Infra-Red (FLIR) pod, and a futuristic secondary task, such as Unmanned Aerial Vehicle (UAV) supervisory command and control.

A total of 34 military aviators, U.S. Naval aviators and one U.S. Marine Corps pilot, with varying levels of flight experience, volunteered to fly three five-minute F-18 simulator sessions in close formation while presented with two secondary tasks in order to evaluate the effects of those distractions on following distance performance. The simulators used for this experiment were high-fidelity F-18C/D/E/F Tactical Operational Flight Trainers (TOFT), currently used to conduct training and maintain proficiency of active duty Navy/Marine Corps pilots.

Results provided clear indication that a futuristic task such as a UAV supervisory interface is significantly more challenging than a traditional task and that both secondary tasks statistically significantly increased the average mean following distance and variance compared to the undistracted flying baseline

scenario. Additionally, no evidence was found that increased flight experience (total flight hours) significantly improved performance of the primary task of formation flight when participants were presented with a distraction.

The integration of a futuristic secondary task (UAV supervisory interface) into the simulator cockpit was successful and well received by participants, but requires further development to be a viable combat multiplier. Knowledge gained from the analysis of performance differences could contribute to improved crew resource management and pilot workload balancing as well as flight safety resulting from the modification of flight procedures based on known effects of distractions in the cockpit.

I. INTRODUCTION

A. OBJECTIVE-RESEARCH QUESTIONS

This thesis will explore the changes in formation flight performance and physiological measures that occur as pilots perform primary (i.e., formation flight) and varying secondary tasks/distractions in the cockpit environment. Close formation flight in the range of 50–200 ft. will be used as the primary cognitive loading, and two separate operationally relevant secondary tasks will be required. The two secondary tasks correspond to *traditional* secondary task such as target prosecution with an electro-optical Forward Looking Infra-Red (FLIR) pod, and *futuristic* secondary task such as Unmanned Aerial Vehicle (UAV) supervisory command and control. This thesis will mainly compare effects of the two different secondary tasks on the formation flight performance. Additionally, physiological measures such as heart rate, respiration rate, and posture will be compared between the two types of the tasks.

The objective of this thesis is to investigate the following questions:

1. What is the difference in performance of the primary flight task when performing a traditional secondary task versus a potential future UAV supervisory task?

2. Does pilot experience and proficiency/currency indicate their primary task performance when performing a future secondary task?

3. (EXPLORATORY) What is the difference in physiological measures such as heart rate, respiration rate, and posture when performing the traditional versus future secondary tasks?

B. BACKGROUND OF PROBLEM

The concept of distractions degrading primary task performance has been investigated in numerous studies, but rarely concerning the impacts of secondary tasks while operating military aircraft (Thomas & Wickens, 2001). One frequently referenced distraction study, by Strayer, Drews, and Couch, investigated the effect of cell phone usage while driving, and is currently used as a Crew

1

Resource Management (CRM) case study by U.S. Naval Aviation to instruct crews in the dangers and consequences of the effects of task distraction (see Figure 1) (Strayer, Drews, & Couch, 2006).

Figure 1.   PatrolSim high-fidelity driving simulator
(From General Electric I-Sim)

In their experiment, Strayer compared the net performance decrement of drivers using cell phones versus that of drivers under the influence of alcohol while operating a high-fidelity driving simulator. They discovered that drivers using a hands-free cell phone device followed a lead car through a preset city-driving environment at a greater distance and at a slower speed than did undistracted drivers. The 'cell phone group' also braked later and more abruptly. These metrics were then collected on drivers performing the same primary task while legally intoxicated, and results compared to those of the cell phone group. The analysis showed that the net effect of the two types of distractors was similarly detrimental to competent operation of the vehicle. This is an important finding that may have similar effects for operators of military vehicles and weapons platforms.

C.   SCOPE

This research specifically addresses the effects of relevant to the military because all vehicle operators from pilots, submarine drivers, and ship operators to tank and Mine Resistant Ambush Protected (MRAP) vehicle drivers are all required to perform primary manual tasks in execution of U.S. military operational missions. These duties are often joined with additional requirements, operational distractions, and relevant but secondary tasking, including, but not limited to maintaining situational awareness by monitoring Global Positioning System (GPS) moving map displays, communicating with friendly forces, and employing weapons systems. This experiment will provide insight to the impact that these distractions have on the performance of the primary task.

An example of relevance for the surface Navy is the helmsman who steers the ship and operates the throttles. Under normal conditions, this task is tedious yet requires sustained attention to maintain correct heading and manual focus for manipulation of a steering wheel. However, there are times when this individual is required to wear a headset and communicate with other watch standers as a secondary responsibility and/or respond to questions from the conning officer, for whom they have to recall information or reference displays, further dividing their attention.

For special evolutions such as entering or leaving port, conning alongside another ship, or man overboard drills, the Navy recognizes that this simple task is too important to allow the expected diminished performance caused by distractions and secondary tasks. Thus, additional watch stations are manned to share the responsibilities during special evolutions. This research aims to bridge the gap between studies conducted in the maritime and ground vehicle arenas over to the aviation domain and provide empirical data and analysis to answer the question: What is the difference in performance of the primary flight task when performing a traditional secondary task versus a potential future UAV supervisory task?

3

D. INTEGRATION OF UAV SUPERVISORY CONTROL AS A FUTURISTIC SECONDARY TASK

The military aviation community is increasingly incorporating Unmanned Aerial Vehicles (UAVs) into the Operational Environment (OE) in order to leverage their (typically) higher endurance and lower human (i.e., pilot) risk during the conduct of missions. The military continues to research methods for integrating control of UAVs within the cockpit as a potential force multiplier and to increase situational awareness. No known data currently exists on the effect UAV supervisory control might have on the performance of other important aircrew tasks, so this study will provide valuable insight for this area.

Although our literature review indicates the potential for evaluating cockpit workload through additional modalities incorporated into secondary tasks to improve information processing, such as Wickens leveraged during his 2005 study on control of multiple UAVs, we will limit our experiment to the current system modalities found in the military aviation community. Wickens references the term "workload" as defined by the relationship between resource supply and task demand based on his 1995 study with Sarno (Sarno & Wickens, 1995; Wickens, 2005). If cognitive demand exceeded supply (or capability) there might be an opportunity to detect a change in performance of task(s).

While Wickens study utilized a more direct UAV control workstation, we expect the cognitive workload required to understand the situational awareness and tactical aspects of employment to remain the same when controlling UAVs through supervisory control methods as well. Supervisory control methods greatly simplify manual control of UAVs, permitting the operator to concentrate more on higher level decision making such as tactical routing, threat assessment, and target selection rather while the lower level control such as maintaining course, altitude, and airspeed are handled automatically by the UAV. Mental workload, however, is still expected to be a limiting factor for determining how many UAVs an operator can manage at any given time.

Rodas and Veronda (2011) studied Unmanned Vehicle (UV) operator performance through a between-subject design with three team sizes; high, medium, and low, with 9, 7, and 5 UVs within each team, respectively, with combinations of aerial and surface and underwater vehicles (UAVs, USVs, and UUVs). Although the Rodas' study data is still being analyzed, she has determined a baseline of 5 UVs would be used to compare and analyze performance. Her study utilized the Research Environment for Supervisory Control of Unmanned Vehicles (RESCHU) developed by the Massachusetts Institute of Technology (MIT) as a test bed for evaluating operator performance since there are currently no other systems currently available (see Figure 2).
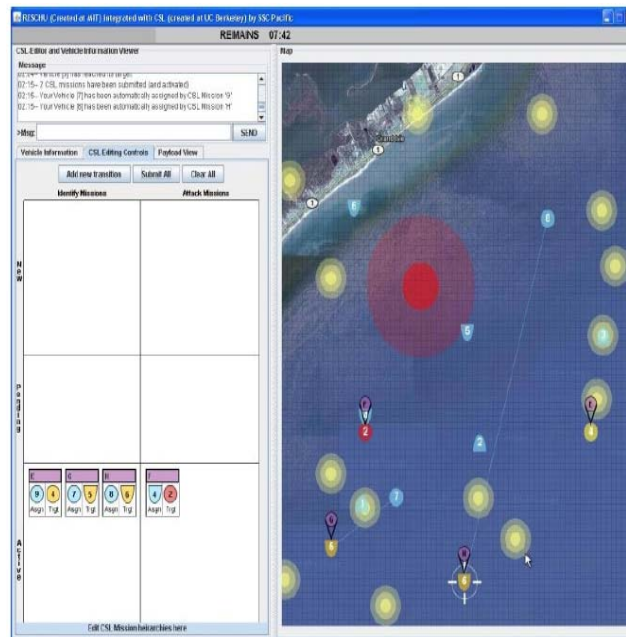
Figure 2.   Research Environment for Supervisory Control of Unmanned-Vehicles (RESCHU) (From Rodas, 2011)

Rodas (2011) utilized RESCHU, a JAVA based simulation in which the operator can control a predefined number of UVs by assigning waypoints to routes, and ordering specific actions for each platform (conduct reconnaissance

or attack a target, for example). No study of implementation of RESCHU within the cockpit has been conducted, and our experiment would be the first to assess its feasibility.

Our study will also use a variant of RESCHU (NPS-developed, called SAWSI: Semi-Automated Wingman Supervisory Interface) to integrate a UAV operator scenario in the cockpit environment. The U.S. military has investigated the integration of UAVs for years, even going so far as placing separate flight controls in the AH-64 (Attack Helicopter) copilot/gunner station in order to pilot a UAV. However, no control systems are currently fielded for employment by tactical units. A further study of UAV supervisory control inside a tactical cockpit environment would provide valuable insight into its impacts on situational awareness as well as the effects it may cause as a result of increased task distraction.

In summary, this study is based on the concept that secondary tasks will result in direct and measurable changes in performance in the primary task of flying the aircraft within professional standards. Additionally, varying the secondary tasks assigned in this experiment will investigate the workload differences between traditional and futuristic tasks compared to baseline conditions (no secondary task). This will hopefully provide insight into the impact of these secondary tasks on cognitive workload.

Through analysis of following distances throughout a programmed flight profile with specifically assigned secondary tasks easily replicated across a pool of test subjects, we expect to provide further insight into their impact on performance of formation flight. If we are able to quantify the impact of secondary tasks, then this study will potentially provide valuable data for design of crew resource management task loading and/or improved integration of UAV supervisory control within the cockpit, a goal which the military has been pursuing for many years.

II.   LITERATURE REVIEW

A.   SELECTIVE VS. DIVIDED ATTENTION AND MEMORY

Studies have evaluated the impact of selective versus divided attention to reveal how memory is affected by the presentation of information. In selective attention tasks participants are presented with multiple sources of stimuli, or information, and must concentrate on a single source while ignoring others to complete the task, where-as in divided attention tasks participants focus solely on the target stimulus or simultaneously with distractors (Mulligan, 2002). Troyer and Craik (2000), for example, studied divided attention effects on memory, and Mulligan (2002) researched the impact of selective versus divided attention shortly afterward. Both studies revealed that memory is affected by color, shape, and presentation styles of information. Therefore, performance of tasks requiring short-term memory can be affected by the quality and type/method of data distracting the participant. Short-term memory effects are important in instances such as when a pilot monitors a moving map in the cockpit that indicates present position, friendly and enemy positions, as well as potential threats, and requires retention of information despite multiple sources of stimuli. The complexity and level of multitasking required by aviators is generally viewed as high, but the field could significantly benefit from a targeted study on task loading under selective vs. divided attention. Experience could be an important indicator of performance if situation awareness and understanding of the so-called "common operational picture" (or COP) can be correlated with total flight hours.

Another consideration for the pilot is the amount of information that he/she can effectively process at any given time. Wickens (2008) addressed this in his study of multiple resource concept and mental workload. Overloading, or exceeding the mental processing capacity, of a participant will result in a significant decrease in performance when performing two or more simultaneous

tasks. Proctor and Zandt (2008) demonstrate the trade-off in dual task performance graphically through the performance-operating (POC) curve (Figure 3).
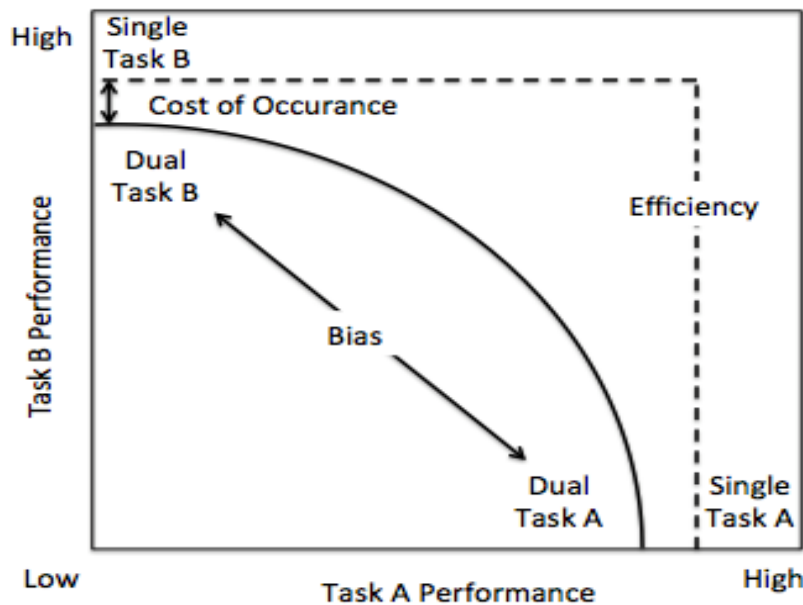
Figure 3. Performance-Operating Characteristic (POC) Curve.
(After Proctor & Zandt, 2008)

Essentially, this hypothetical POC curve illustrates the consequences of divided attention as an operator alternates between Task A and Task B, or Primary and Secondary tasks. Performance of both tasks can be maintained to optimal levels as indicated by the dashed lines when no attentional limitations impact the operator, but decline to sub-optimal levels indicated by the POC curve when the operator is impacted by attentional limitations and attempts to improve performance in either the Primary or Secondary task. While performance of one task can be increased to near optimal level, the operator will never achieve 100% efficiency in both tasks, and their performance of the other task is anticipated to decline rapidly as more and more effort is dedicated to improving performance.

Multi-tasking is a well-researched area, particularly in the aviation domain. Wickens' (2003) research revealed that performance of two simultaneous tasks using different attention resources degraded the performance of each of those tasks. These breakdowns are related to dual-task overloads. The above studies outline the theory of distracted task performance and form the basis of investigation into military application of this concept of the effects of cockpit distractions.

B. TASK PERFORMANCE DURING DIVIDED ATTENTION

Wickens (2008) has also conducted a relevant study on task performance during divided attention from which he has developed the Multiple Resource Theory. The Multiple Resource Theory states that as multiple tasks require overlapping mental or physical resources to perceive, understand, and respond to input, the performance of tasks will be affected after reaching an "overload region," where residual capacity of an individual is exceeded (Wickens, 2008). This theory forms the basis of our research, which will continue investigation into the influences of secondary tasks (distractions) on primary task performance in a tactical flight environment.

C. EFFECT OF INTERRUPTION MODALITY ON PRIMARY TASK

Ratwani and Trafton (2010) examined the effect of interruption modality (visual or auditory) on primary task (visual) resumption to determine which modality was the least disruptive and found that an visual distraction was more disruptive on a primary task. This research attempts to provide insight to a prediction on the impact of distractions in the flight simulator that require the participant to break eye focus on his/her primary task of controlling the aircraft which may be applicable in real flight. Distractors that require the participant to only recall or make a mental computation should have a smaller impact than those that require the participant to look at gages, flip through a reference book, or write something down. Since we are attempting to evaluate the effect of secondary tasks on the primary task of flying, two visual secondary tasks were

chosen as distractors in order to maximize the impact of disruption on the participants' performance. One task, termed the traditional task, was selected from the common missions each participant is already trained to perform, while a second task was designed to model a potential futuristic mission. To that end, this study adds 1) an air-to-ground mission as a traditional task, and 2) a simplified and intuitive touch screen Unmanned Aerial Vehicle (UAV) supervisory interface (termed Semi-Autonomous Wingman Supervisory Interface or SAWSI) as a futuristic secondary task, both of which are described in greater detail in the methods section.

D. SITUATIONAL AWARENESS

While operators of any military vehicle possess inherent control of a number of variables such as direction, speed, acceleration, etc., they must also possess and maintain situational awareness in order to employ its weapons systems. Endsley (1995) defines situational awareness as "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future." She further categorizes three general levels of situation awareness (SA):

Level 1 SA—Perception of elements in the environment (status, attributes, and dynamics of the environment)

Level 2 SA—Comprehension of the situation (understanding the significance of level 1 data and acting upon it)

Level 3 SA—Projection of the future (understanding the implications of level 1 SA and 2 SA)

Figure 4, Endsley's model of situation awareness, explains factors influencing situational awareness. In the extremely dynamic environment as experienced by military aviators that relies heavily on perceiving, processing, and acting upon multiple elements, the division of limited attention and mental workload has direct impact on gaining and maintaining situational awareness.
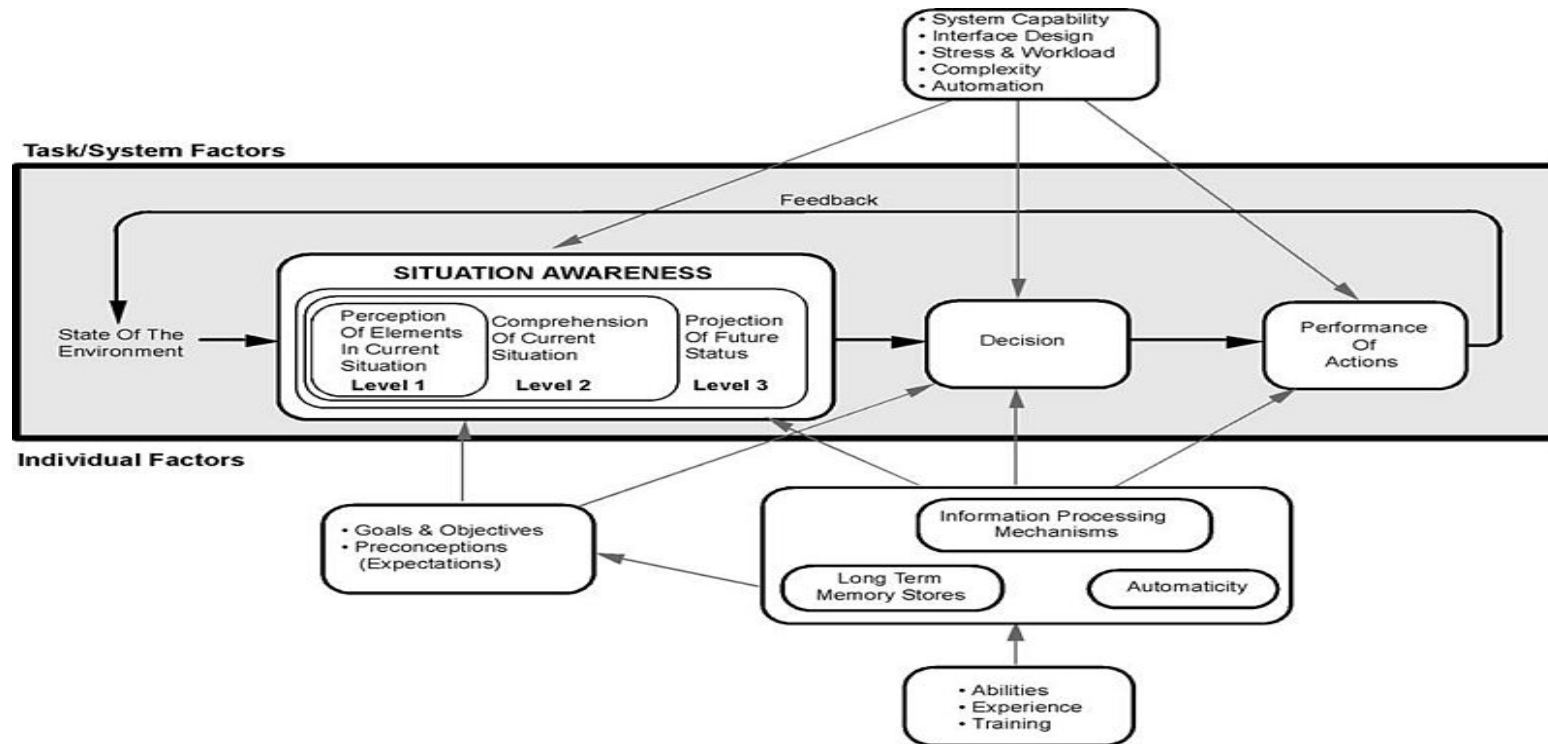
Figure 4.   Endsley's model of Situational Awareness (SA) (From Endsley, 1995)

Endsley (1999) further defines four common elements of SA across all platforms:

1.   Geographical SA—understanding one's own location and surrounding environment, including terrain, facilities, threats, cities, and navigational aids

2.   Spatial SA—perceiving ones flight configuration and performance, including attitude, heading, airspeed, altitude, rates of change, and flight path

3.   Systems SA—knowing the configuration, settings, and or impact of degradation of radios, navigation, weapons, and defensive systems

4.   Tactical SA—locating, identifying, classifying, and understanding the flight dynamics and capabilities of surrounding aircraft (both friendly and enemy)

A combination of all these elements forms the overall SA for military pilots in a tactical environment. Improvements in automation and cockpit design and configuration throughout the years has greatly improved the ease of which pilots sense and correlate information, yet constantly increasing capabilities with ever improving technology continues to push pilot workload to its limit.

Pilots are trained to scan instruments and sensors depending on the primary task they are currently operating in order to avoid attentional limits and working memory capacity. However, there is no established sequence that works best for all pilots. Each must learn to harness the most of their abilities through training and experience what scan patterns and time allocations work best to fly safely, maintain sufficient SA, and be able to employ their weapons systems.

Challenges to situational awareness can be affected not only by the individual characteristics and capabilities of vehicle operators, but also by environmental and systems factors (Endsley, 1999). Stress from physical or social factors, overload/underload of operator/crew tasks, system design and capabilities, complexity of systems, and automation of systems which may leave a pilot "out of the loop" conspire to introduce errors in situational awareness.

E. PILOT PERFORMANCE

While pilot performance for observable standards such as how accurately altitude and heading is held, or number of successful landings, some measures can be more challenging to assess since there is a degree of art rather than pure science to flying. Hitchcock (1999) states performance measures first must be quantifiable. Cockpit workload, while involving physical demand to some degree, largely focuses on the mental and perceptual demands based on the pilot's phase of flight. Hitchcock references a 1984 study that focused on four central questions pertaining to workload:

1. Will a pilot's current workload permit additional tasks?
2. Does the pilot's current workload leave room for dealing with emergencies?
3. Can a pilot's task be modified to reduce workload?
4. Will a new system increase or decrease pilot workload?

These are important considerations when thinking about the addition of a new system (in our case, a UAV supervisory control). Other, better understood and thoroughly studied factors affecting pilot performance include the impacts of rest and fatigue, the physiological stresses of the cockpit environment (g-forces, vibrations, etc.), as well as the general health of the pilot due to individual differences in nutrition, level of physical fitness, age, and use of alcohol or other substances. This research will yield quantitative data that will contribute to the further advancement of understanding cockpit workload and its effects on the pilot's performance of the primary task of flight.

Through further application of Wickens' theory, we hope to objectively evaluate pilot performance of a primary task (maintaining formation flight within close tolerance) while exposed to traditional and potentially future secondary tasks utilizing fleet F-18 simulators as an in situ environment (see Figure 5. Figure 5). By analyzing the performance of the primary task within participants, we will determine what differences, if any, exist between applications of secondary tasks.

Figure 5.   U.S. Navy F/A-18 Tactical Operational Flight Trainer (TOFT)
(From U.S. Navy Archive)

The introduction of secondary tasks as distractors while performing the primary task of maintaining minimum aircraft separation will likely result in similar performance decrements. From Mulligan's study, it can be predicted that distracting questions and tasks that do not pertain to the participant's primary task will influence primary task performance more than those that are relevant the primary task (Mulligan, 2002). The Troyer and Craik (2000) study leads to the prediction that if a person is asked to recall a non-defining attribute of their previous questions, it will have more of an effect than if they were asked to recall a fact. Similarly, the Watkins study considers that if we expect to see a change in flight control performance, the secondary tasks should involve distractions that utilize attention resources that are similar to or overlapping those of the primary task.

Our proposed primary task in this thesis was to have the participant fly the aircraft as closely as possible to the lead aircraft, which uses the resources of

manual stick control and visual monitoring of the lead aircraft and relative distance. A secondary task that overlaps the primary task attention resources could be one of two tasks; traditional and futuristic. A traditional task could be operation of a targeting pod whereas a futuristic task could be controlling one or more UAVs in support of an attack. Both tasks would require the pilot to take their eyes off the lead aircraft for various lengths of time depending on current workload and mental resources required. Wickens' findings support the prediction that the addition of this secondary task will degrade the performance of the primary task.

THIS PAGE INTENTIONALLY LEFT BLANK

III. PILOT STUDY

A. BACKGROUND AND SUMMARY

In approaching this field of research, we wanted to explore the boundaries of where distraction would and would not be a factor. In our home environment of the Naval Postgraduate School however, we have neither access to a large population of qualified and current aviators, nor a realistic simulator in which to check our suspicions. Therefore, we decided to perform a pilot study with the people and materials at hand to begin to assess how performance of a manual primary task was affected by auditory distractions presented in the form of secondary tasks. Volunteer subjects were asked to maintain minimum following distance from a lead aircraft in a desktop flight simulator as their primary task. The main selection criterion for participants was that they had no flight training experience in order to eliminate a potential source of confounding error.

Participants were asked to operate an aircraft simulator for two 5-minute observation periods. The participants were randomly sorted into three groups and each participated in one (unmeasured) orientation period, then two additional periods of data collection. Group A received distractions during the first observation period and Group B received distractors during the second observation period. Group C did not receive any intentional distractions. During the orientation, participants of all groups operated a simulated aircraft in a controlled environment. Participants flew the aircraft in formation behind a lead aircraft of the same type and were asked to maintain the closest possible following distance. That inter-plane distance was the measure of performance for this experiment. Verbal requests for information or other secondary tasks constituted the distractions to primary task performance.

1. Hypothesis

Question: what effect do distractions have on the formation flying ability of non-expert participants?

H_0: there would be no difference in performance of the primary task (aircraft control) with and without the treatment of distractors.

H_A: participants treated with a distraction will fly more erratically than those who are not.

B. METHODOLOGY

1. Participants

The participant population sample pulls from Naval Postgraduate School (NPS) students who possess no prior piloting experience. As this was only a pilot study, no further demographic information was collected. Participants were randomly sorted into three groups (Table 1) with each participant receiving one unmeasured orientation period, followed by two periods of data collection. Group A received distractions during the first observation period and Group B received distractions during the second observation period while group C did not receive distractions during either period.

Table 1.   Pilot Study design of experiment

Group	Session 1	Session 2	# Participants
A	Treatment	No Treatment	3
B	No Treatment	Treatment	3
C	No Treatment	No Treatment	3

2. Apparatus

The experiment was performed on a desktop system running Microsoft Windows 7 and displayed on a two-screen arrangement. Input was provided via

standard keyboard, mouse, and a Logitech 4-axis joystick with reconfigurable buttons as seen in Figure 6. The CPU was an Intel Dual-Core running at 3.65 GHz with 4GB of RAM, and the displays were a pair of Dell 17" class flat screen monitors arranged horizontally with the simulation display window stretched across both screens but slightly offset to allow the center of focus to not be directly on the break between the monitors.

Figure 6.   Pilot Study: Apparatus setup

Software was a standard X-Plane 10 installation set to simulate high-performance fighter aircraft, in this case a pair of F-22 Raptors. The software provided a start-off initial condition for formation flying, but there was a large airspeed differential between the lead and trail aircraft, so the experimenter was required to slow the participant's aircraft and stabilize it at a nominal following distance prior to handing the controls over to the participant for the start of each session.

3. Procedure

Upon commencing each subject's participation, they were given a quick orientation to the controls and then released to "free flight" for several minutes to allow them to acclimate themselves to both the simulator operating environment and the handling characteristics of the aircraft.

Next, the simulator was reset to the initial starting conditions they would see for the remainder of the experiment. The participant's aircraft was reset to a trailing position approximately 1/3 mile behind the lead aircraft, both aircraft in straight and level flight at the same altitude and airspeed (10,000' MSL, 370 knots indicated airspeed, KIAS), as shown in Figure 7 and Figure 8.

Figure 7.   Pilot Study: Starting position

Figure 8.   Pilot Study: Representative flight displays

The participants were then instructed to familiarize themselves with the primary task by closing to the nearest possible position on the lead aircraft and then attempting to maintain that distance for the remainder of the experiment (Figure 9). Once the participant was comfortable with the task, the orientation period was complete and the experimenter again reset the simulation to the initial trail position while the participant had a two-minute break.

Figure 9.   Pilot Study: Measured response

What occurred over the next two subsequent periods depended on which data group the participant in. If the participant was assigned to group B, the first data-recording phase was a five-minute session where the participant was only to perform the primary task of maintaining position on the lead aircraft again. At commencement of the session, data was automatically gathered and recorded at 10-second intervals on both planes' position (latitude, longitude, and altitude) for follow-on analysis. For analysis, our primary performance measure of inter-plane following distance was then derived from these measures.

At the completion of the 1st session, the experimenter again reset the simulator to the initial condition, while the participant was given another two-minute break. In the second data capture session (for data group B), the exact same procedure was followed as in the first session, except that the participant was given a series of realistic, short duration secondary tasks to perform whilst continuing to try to perform the primary task as well as possible. These

secondary tasks were the same for each participant and introduced introduce serially at 0:30, 1:30, 2:30, 3:30 and 4:30 minutes from session initiation, respectively. At the conclusion of the session, the participant's involvement in the experiment was complete. The following questions were asked:

0:30–"What are your current longitude, latitude, and altitude?"
1:30–"Stand by to copy target coordinates... Target coordinates are:
    43° 21' 28" N
    108° 46' 26.46" W
    Read back your target coordinates."
2:30–"What is your remaining fuel in minutes?"
3:30–"What two radio frequencies are you monitoring?"
4:30–"What was the first question asked?"

The answer to the first question required the participant to read outputs on the heads up display and data output block, seen in the upper left corner of Figure 8.

Answering questions three and four required the participant to shift to a look around view using keystrokes, maneuver the mouse to read the instruments and multi-function display below, and then return to the heads-up display view shown in Figure 8. A snapshot of this can be seen in Figure 10.

Figure 10. Pilot Study: Look around view

Accuracy and correctness of the answers to these five questions was noted, but not considered during data analysis at this time as the actual

performance of secondary tasking was not a metric of evaluation during the pilot study, only the effect that this secondary tasking has on the performance of the primary task.

For treatment group A the order of the data-collection sessions was reversed, with the treatment occurring in the first measured session. If the participant was in treatment group C, no treatments were administered in either session, and the participant simply performed the primary task for both.

C.   RESULTS

Error! Reference source not found. shows the participants' performances in the flight simulator in terms of *average change in following distance* behind the lead aircraft for treated and untreated sessions. A positive delta indicates that, the participant moved closer to the lead plane, and a negative delta indicates that the participant moved farther away from the lead plane.

Table 2.   Pilot Study: Average following distance delta
(green cells are treated sessions)

Subject No.	Group	Session 1 Av. Delta	Session 2 Av. Delta
1	A	-99.27	6.23
2	A	-25.62	-310.4
3	A	-538.30	39.24
4	B	46.42	-51.16
5	B	-21.94	-129.53
6	B	48.84	19.83
7	C	24.71	14.42
8	C	32.75	91.42
9	C	-702.69	-29.31

1. Session One vs. Session Two Performance

The results of session one and session two data manipulation are as follows: the mean change in following distance was -137.23 feet and -2.15 feet, respectively, and the standard deviation was 280.90 feet and 62.38 feet, respectively.

From rough plots (not included due to lack of statistical significance) of the average deltas of the first session versus the second session, it was apparent that there were some outliers on the negative and positive side for the first and second session respectively. However, further testing with a greater number of participants would be necessary to draw any significant conclusions from that observation.

2. Treated Versus Untreated Performance

Statistical evaluation of the treated and untreated sessions was possible only when the three members of the control group were excluded from data evaluation. This is due to dependence on equal sample size in a blocking situation. Over the 18 observed data collection sessions, 12 were untreated and six were treated. This unequal sample size made blocking impossible. Thus, evaluation was conducted of A group and B group separately, with C group being considered only in session one versus session two data analysis.

The average change in following distance for the untreated sessions was -137.34 feet as measured every 10 seconds. This means that on average, the participants were increasing their distance from the lead plane by 137.34 feet every 10 seconds while the standard deviation for these sessions was 48.84 feet. By comparison, standard deviation during the treated sessions was more than four times greater at 203.41 feet. On first look this indicates, the H_0 is not supported, but requiring a larger dataset to confirm the negative result.

D. DISCUSSION

Analysis of the delta information seems to indicate that the performance of each participant was worse with treatment; however there was no statistically significant difference between the two likely due to the very small sample sizes. No statistically significant difference could be detected between the treated and untreated sessions; however data from all sessions trended toward significance, such that we felt increasing the sample size could be a first step in a more accurate evaluation of the null hypothesis.
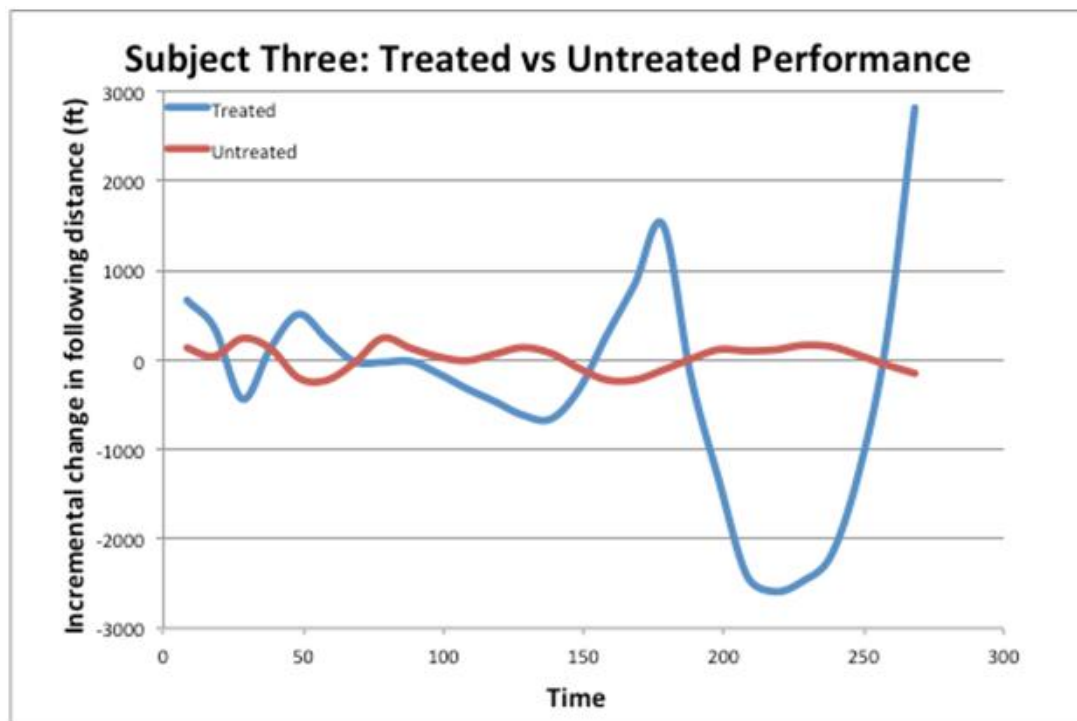
Figure 11. Pilot Study: Treated versus untreated performance

Figure 11 is a plot of the incremental change in following distance in time for one participant's untreated and treated session. From the plot, it can be seen that the red line (untreated delta) assumes a relatively stable adjustment pattern. This is juxtaposed with the blue line, which has much higher peaks and valleys indicating that the subject had a harder time controlling the following distance while being distracted.

Reviewing the protocol used in the treated sessions, we can readily see that each large change in the trend of following distance coincided with timing of a question or task. So, though a small *n* precluded the statistics from being conclusive, this observation was the motivation we received to believe that further study was warranted.

Additionally, since we readily saw the effect of a dedicated control group on our ability to draw conclusions from our data, we were able to apply this lesson learned to a more appropriate design of experiment for our full study.

E.   APPLICATION TO LARGER STUDY

The purpose of this pilot study was to evaluate the idea that performance of a simple primary manual task would be impacted by distractions and secondary tasks. Results concluded that primary task performance was impacted, though not to a statistically significant level. Statistical analysis was unable to prove significant differences in performance between treated and non-treated performance as well as between session one and session two. Subjective observations of treatment versus non-treatment results indicated that including distractors had degrading effect on primary task performance, short-term memory, and allocation of mental resources.

Results of this pilot study indicated significant possibility for successfully conducting a test on a more homogeneous subject population, such as we found for our primary experiment. Therefore, we determined to repeat the experiment with a larger sample size in order to make a more definite conclusion from the data. A population of similarly qualified actual aviators was also preferred to produce more consistent and interpretable data. Other metrics recommended to improve visibility of the effects include those of secondary task performance and biometric response to multiple stimuli. The degradation in performance indicated by these metrics could lead to determining an optimal time for task assignment and management.

IV. METHOD

A. REVIEW OF LOGICAL CONSTRUCT

1. Instantiation of Constructs

Drawing from lessons learned in the pilot study, we decided to continue to use an analogous primary task to Strayer (2006) in having the participants follow a lead aircraft in a designated formation. Further refining of the secondary tasks we chose a traditional ground-target prosecution using the onboard targeting pod, and a potential future task involving the supervision of a small flock of semi-autonomous UAVs.

2. Independent Variables

The independent variable is the factor of distraction, held at three levels: primary task only, primary task plus traditional secondary task, and primary task plus the potential future secondary task.

Additionally, in support of our second research question we consider several *pseudo-independent* variables from our participants' biographical data as candidates to capture the concept of experience level.

3. Dependant Variables

Our primary dependant variable is derived from the performance of the primary task as measured by following distance of the participant's simulated aircraft from the computer-generated lead aircraft entity. Since following distance was a time-varying value in this case, our initial intention was to use both the session mean and standard deviation to capture both the magnitude and relative stability (or lack thereof) of following distance over each session. Other measures could have been used (such as the intra-sample change as used in the pilot study), but that did not seem to produce any further insight into the participants' performance.

Additionally, in support of our exploratory (third) research question, we use biometric data gathered from our participants during the sessions in the form of heart rate, respiration rate, and trunk position.

4. Confounds Held Constant

A large number of potential confounds were held constant by the use of the U.S. Navy simulators. Participants used equivalent cockpits, and due to using the same simulator operator for almost all runs, the virtual lead aircraft followed essentially the same path, altitude and speed profile. Weather and time of day in the simulation was constant and benign, configuration (fuel load, armament, etc.) of each participant's virtual aircraft was the same, and communications were precisely repeated.

Regarding participants, we pre-screened potential aviator participants to be mid-level experience, qualified Hornet pilots. Hornet naval aviators become mission qualified by the Strike/Fighter Weapons and Tactics Instruction (SFWTI) program in a system of levels. Level II qualification equates roughly to a "wing man" qualification where the aviator has been introduced to all the primary tactical missions they can fly as a wingman and demonstrated sufficient proficiency in them. Most aviators complete the Level II syllabus within the first year to year and a half of their tour in their first operational squadron, and this provided a convenient lower experience bound for our acceptable participants. On the upper side we excluded aviators who had started the Level V (or Strike Lead) syllabus as these aviators are senior flyers with a high level of experience and higher than average tactical proficiency. These bounds also conveniently produce relatively consistent experience as quantified by the participant's total number of flight hours.

Between the two simulator models, slight differences in cockpit layout exist which correspond to the analogous differences in the actual aircraft themselves. However, one large difference which will be discussed thoroughly in Chapter V is that that the two models' mission playback systems, which were

28

needed to be able to report the following distance, varied greatly in the stability of their output. The older, Super Hornet system consistently produced "noisy" data, while the newer legacy Hornet system was relatively smooth and stable. This variance inspired production of a smoothing method which is also discussed in Chapter V.

Additionally, a slight possible confound may have been present in the difference in actual ambient room temperature between the two types of F-18 simulators used (F-18C and F-18E/F). The simulator hall in the F-18C area was held about 70°F, while that on the F-18E/F side was noticeably colder at around 65°F, however all these simulator variations will be accounted for in the resultant model as rolled into the simulator model variable.

5. Hypotheses

1. H_{01}: What is the difference in performance of the primary flight task when performing a traditional secondary task versus a potential future UAV supervisory task?

   H_{A1}: There will be significant differences in the execution of the primary task of maintaining safe, effective flight when presented with either a traditional or potential future secondary task.

2. H_{02}: Does pilot experience and proficiency/currency indicate their primary task performance when performing a future secondary task?

   H_{A2}: Pilots of higher proficiency and currency will show less change in performance of the primary task when treated with the operational task distraction of a UAV supervisory task.

3. H_{03}: (EXPLORATORY) What is the difference in physiological measures such as heart rate, respiration rate, and posture when performing the traditional versus future secondary tasks?

   H_{A3}: The subject's heart rate, respiratory rate, and trunk position will be significantly different when performing the traditional versus future secondary tasks.

6. Goal

To test H_{01} we are able to manipulate the level of distracting task while holding constant all confounds in IV.A.4 and measure the dependent variable of following distance.

To test H_{02} we are able to test subjects from a range of demographic backgrounds in the futuristic distracting task while holding constant all confounds in IV.A.4. and measure the dependent variable of following distance.

To explore H_{03} we are able to manipulate the level of distracting task while holding constant all confounds in IV.A.4. and attempted to investigate and record several representative physiologic measures.

B. PARTICIPANTS

The population we are attempting to represent is Strike/Fighter U.S. Naval Aviators, of which there are approximately 750 in the entire Navy who are current in the Hornet (both F-18C/D Hornets and F-18E/F Super Hornets) and assigned to operational units at any given time. We specifically target mid-level pilots as they are the greater portion of this population.

Selection was accomplished through random recruitment of volunteers, and ultimately 34 subjects from operational line squadrons and the west coast Fleet Replacement Squadron (FRS) at Naval Air Station (NAS) Lemoore in California were recruited for this study. This is significant as it represents nearly 10% of the total number of active fighter/attack aviators in the west coast fleet. The participants ranged from 26 to 39 years of age (mean = 31, sd = 3.6 yrs) (see Figure 12 and Figure 13). Total flight hours (TFH) ranged from 400 to 3,200 with an average of 1,475, a median of 1325 and a standard deviation of 693. Total hours in type/model (any F-18 variant) ranged from 200 to 2,500 with an average of 1,147, a median of 1050 and a standard deviation of 615 hours (see Figure 14.). Participants were all male, of which 33 were Naval Aviators, and one was a Marine Aviator.

Age of Subjects
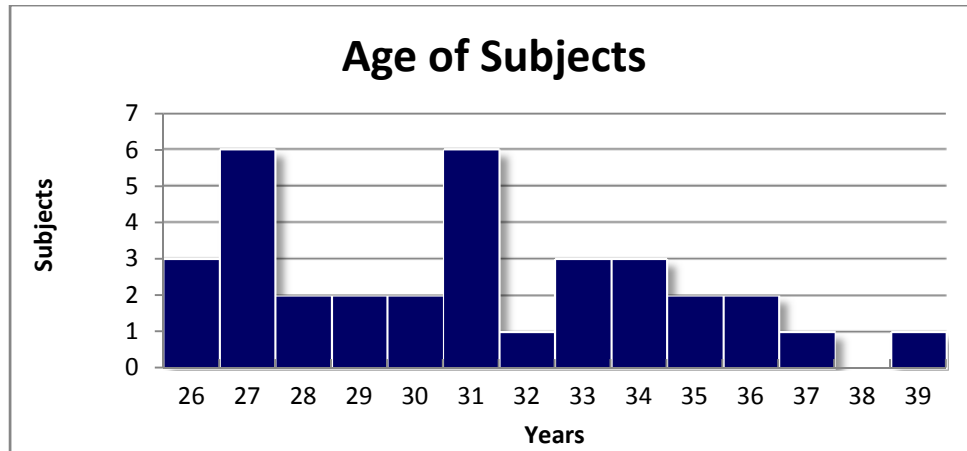
Figure 12.   Subject age distribution

Pilot Rank

Primary F-18 Series
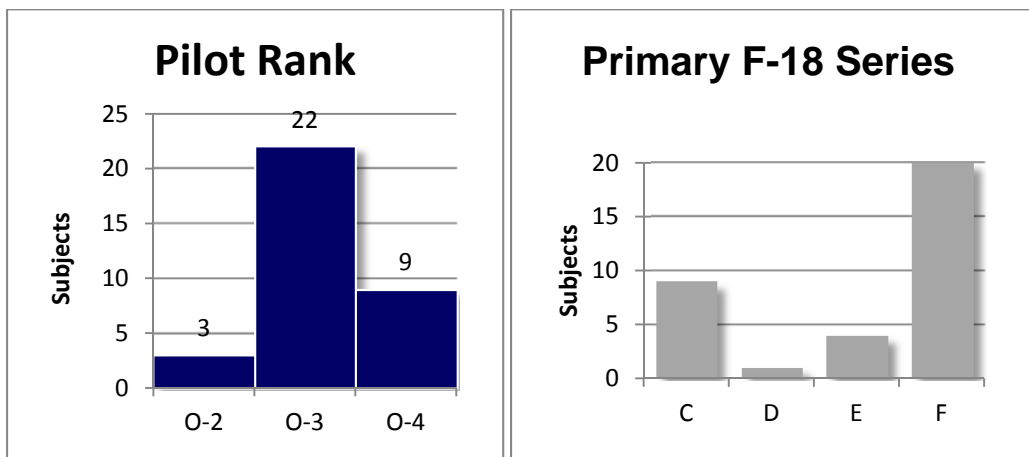
Figure 13.   Participant demographic distribution

To participate in this experiment pilots needed to have completed their Strike Fighter Weapons Tactics Instruction (SFWTI) Level II (wingman) qualification (as described in section A.4), but not have yet started on their Level V (Strike Lead). Volunteers for this study were recruited from the base by an Institutional Review Board (IRB) approved email sent to squadrons at NAS Lemoore and by the authors personally visiting their ready-rooms to solicit participation. This effort was largely successful in that perhaps as much as 20% of our targeted demographic of mid-level experience aviators in that fleet ultimately participated in the study.
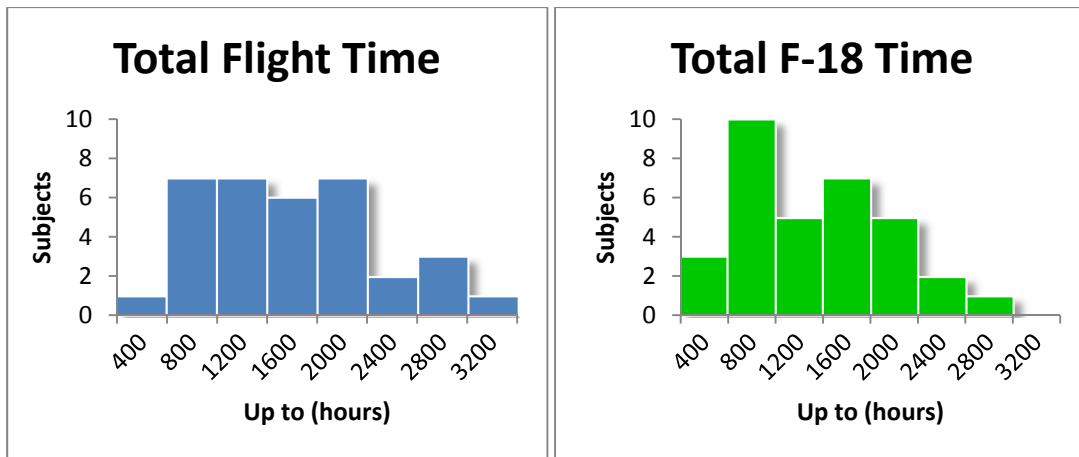
Figure 14.   Participant flight experience distribution

C.   MATERIALS AND APPARATUS

To set the conditions properly for our independent variables and be able to gather data on our independent variables, we leveraged a combination of U.S. Navy flight simulators, a two custom designed tablet computer interfaces, and a pair of commercial "bio-harnesses."

1.   Primary Task Simulators

The U.S. Navy has three primary F-18 bases in the United States, while the Marine Corps. has one, and each has a full set of mission simulators which are capable of realistically portraying a full range of current missions and operating environments. The current system fielded at Naval Air Station (NAS) Lemoore, CA is contracted by L^3 Simulators and provides units for both the F-18C/D ("Legacy" Hornets) and F-18E/F (Super Hornets). Both sets employ highly realistic visual representations and cockpit look and feel, but are not motion-based.

Cockpit simulators in Lemoore are housed in a single large building with two wings, one for each generation of Hornet. Each system is comprised of a cockpit setup, surrounded by a multifaceted, tightly connected arrangement of nine rear-projection pentagonal screens, which itself is all inside a light-baffle enclosure (Figure 15). Each cockpit is then controlled and programmed from a

console located nearby which coordinates the implementation of separate High-Order Language Architecture (HLA) simulations for (1) the aircrew's cockpit and (2) any other entities and the operating "world." At the console, a contract civilian simulator operator programs the simulator environment, and in our study also controlled the other simulated entities, including the lead aircraft for the primary task. The author (Naval Flight Officer) sat at the other position on the console and provided verbal coordination for the sessions, and played the parts of flight lead and Command and Control (C2) via "radio" communication.

Simulators can be tied together to operate in the same scenario to work cooperatively or in opposition, and can even be tied in to other distributed simulations hosted elsewhere. We decided to forgo a manned lead (simulated) aircraft to gain a significantly more stable, repeatable, and consistent lead flight path.

Additional rooms are provided for debrief and replay of events, or can be configured as mission control spaces viewing all relevant tactical system displays and interfaces in real time for up to four separate simulator aircraft. In the replay configuration, the debrief system coordinates playback and synchronization of the two primary mission systems and the implications of this will be covered later in Section F.1.

Figure 15.   F-18C TOFT—entrance panel open
(Training & Simulation Forum, 2012)

a.   *Legacy Hornet Simulator*

For the west coast wings, the newer simulators of the two generations are actually on the legacy side. Generally, visual detail is noticeably higher than the older Super Hornet sims, but simulated aircraft response is similarly accurate between the two as reported by several qualified aviators. As mentioned in A.4, all four simulators are held in one large hall which is kept at a slightly warmer ambient temperature than the spaces on the Super Hornet side. For our study we only used Tactical Operational Flight Trainer (TOFT) 24.

A more significant difference between the sides was that because of the newer system architecture, the legacy hornet sims' playback system was smoother, both in visual representation and in presentation of the data stream we used to gather our primary task measure as described in Section F.1.

b. Super Hornet Simulator

Simulator light enclosures are located in two separate locations, and their control consoles are located in separate, adjacent rooms. Of the four operational simulators during the time of our study, only TOFTs –132 and –133 were used.

On the Super Hornet side the system architecture was older and required a larger portion of system resources to accurately play back the participants' sessions. The effects of this will be discussed in much greater detail in Section F.1.

2. Secondary Task Apparatuses

As described below, the secondary task chosen for our investigation served to provide realistic and tactically relevant distraction from the primary task. Two separate secondary tasks were chosen to represent both current tactical operations and a possible future strike tactic.

a. Traditional Task

Since this distracting task is currently regularly performed by Hornet aircrew, both models of simulator had sufficient means to represent the task within representative cockpit interfaces. The simulators were programmed to represent an Advanced Tactical Forward Looking Infra-Red (AT-FLIR) pod on the left "cheek" station (Figure 16.) and sensor display was on any available Multi-Purpose Control Display (MPCD) in the cockpit, but usually in compliance with squadron and community Crew Resource Management (CRM) and Tactical Standard Operating Procedures (TacSOPs). (Sample display image unavailable due to classification.)

Figure 16.   AT-FLIR on a Hornet (From Federation of American Scientists, 1999) http://www.fas.org/man/dod-101/sys/smart/atflir.htm

b.   SAWSI Interface

A purpose-built "Android" operating system application was written and adapted from the RESCHU research program written for DOS systems. The authors designed an application to portray the intended scenario and depict a Semi-Autonomous Wingmen Control Interface (SAWSI, pronounced "SAW-see") of the type as may be desirable in future years (see Appendix D: SAWSI Scenario Figures). The authors managed a small team of programmers in this effort and developed a "futuristic" tactical distraction task with a deterministic scenario so that the events would be reproducible across all participants.

SAWSI was run on a pair of 32GB Samsung Galaxy tablet computers, model# GT-P7510MA, operating Android version 4.0.4, kernel 3.1.10 and build IMM76D.UELPL. In order to minimize any possible unintended interactions, tablets were operated with Wi-Fi, GPS, auto brightness, and auto screen rotation off, but Bluetooth on so that the exploratory apparatus data could be received as described in the next section. Screen contrast mode was set to "Standard," and screen timeout was set to 30 minutes. Before running the

SAWSI app, all active applications were purged from memory and RAM was manually cleared to ensure smooth operation of the app.

Leveraging the inherent touch in interface of a tablet computer, participants would touch the desired UAV to select it, and then drag a rubber band line to the desired destination. The assigned asset would proceed at a predetermined rate to the prescribed end point and then take up an orbit. After a variable amount of time (3–8 sec) the UAV symbology changes to green to indicate that imagery of the assigned investigation area is ready for the user to view. The user can then tap the UAV symbology to bring up an on-tablet display of the simulated imagery which had just been "gathered" by the UAV. Interaction continues in this vein until the task is completed and the threat symbology is removed. Further detail is provided in Section D.3.

All user interactions with SAWSI (UAV destination assigned, image magnified, strike assigned, etc) are time-stamped to the 1/10th second by event and recorded in a text file which is stored on the tablet. All significant events that the SAWSI simulation completes (UAV reaches destination, etc.) are similarly logged.

3.   Exploratory DV Apparatus

In support of our third research question we employed a pair of Zephyr BioHarness™ BTs wearable physiologic monitoring transmitters to log our participants' heart rate (rated for 25–240 BPM ±1 BPM), respiration rate (rated for 3–70 BPM (±1BPM)), and trunk position (± 180°) (Zephyr Technology, 2010). Data was transferred via Bluetooth link at a 1hz recording rate to the SAWSI tablet and included in the user interaction log file. Once the SAWSI app was started, all data packets it received were logged, whether or not the participant was interacting with SAWSI in any other way. Thus at least 20 minutes of data, or more than 1200 physiologic data points were to be recorded for each participant.

Figure 17.   Zephyr BioHarness (From Zephyr, 2010)
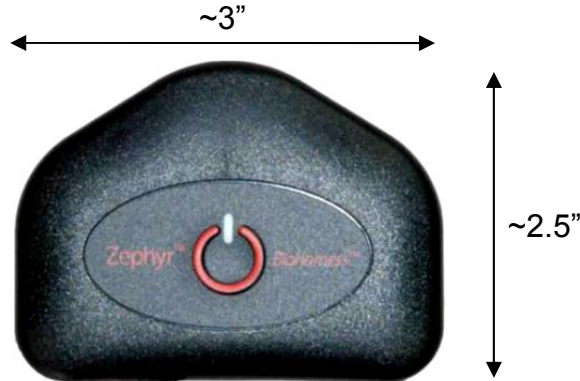
Figure 18.   Zephyr BioHarness transmitter (From Zephyr, 2010)

The harness is constructed with conductive fabric and is worn around the chest just above the bottom of the sternum. It can be fitted to a full range of chest sizes, and is supposed to be able to operate either on top of an undershirt or directly in contact with skin. Unfortunately, our experience showed that when the harness was worn on top of any clothing, data capture was spotty at best and a large portion of our participant's data was incomplete and not reliable.

D.   TASKS

Selected tasks were designed to represent real tactical tasks currently performed by today's Naval Aviators, or in the case of the futuristic task within the realm of feasible while maintaining tactical relevance. Both of the secondary tasks were selected to be of similar modalities to the primary task, as discussed further in Section II.C.

1.   Primary Task Description

As is the case with any flight, safety of flight is always a primary task, so in the case of formation flight maintaining separation from one's lead is of utmost

concern. Participants were asked to maintain "parade position" on the computer generated, sim operator controlled, lead aircraft simulator entity. "Parade" is a standard position defined by a pair of visual checkpoints a following aircraft (or "wing") maintains on the lead aircraft (see Figure 19), and equates to approximately 80' of cockpit-to-cockpit separation between aircraft (see Figure 20). This formation is not standard tactical employment, but in the scenario brief participants were given an explanation at higher classification level as to why this position would be necessary for this mission. For study purposes, this unusually close position was desired as it yields the most rapid visual feedback for the participants and was shown in the pilot study to produce the most consistent results.

Figure 19.   Visual checkpoints for parade position on a Hornet
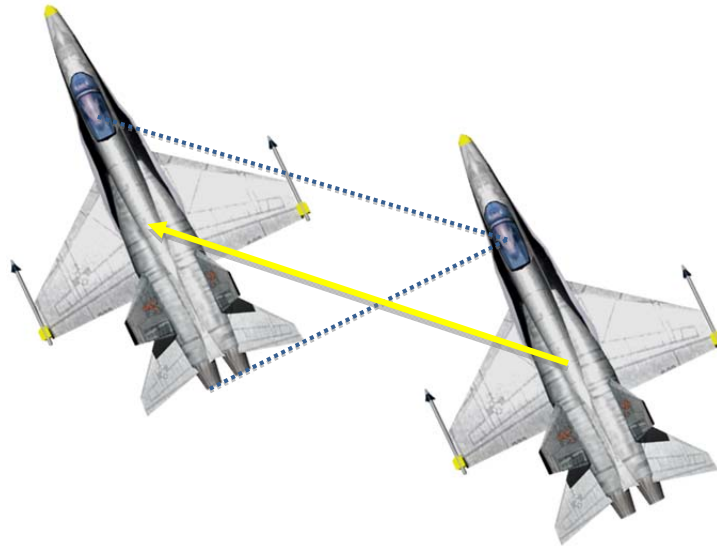
Figure 20.   Inter-plane following distance

The lead aircraft was the *tactical section lead*, so the contract was that he had primary air-to-air threat responsibility, and that wing (the participant's aircraft) would have primary air-to-ground responsibility. This meant that any threat aircraft would be responded to by lead and he would maneuver the section as tactically prudent while wing would maintain position and accomplish whatever other tasks necessary for him to accomplish his tasking. In our scenario there were three such maneuvers.

To give the participants a proper mindset for their actions in this study, we devised a situation in which a section of aircraft was to execute a self-escort strike with radar early warning control provided by an E-2 Hawkeye in an orbit nearby. The section was "feet wet" over the water and would soon cross over into enemy land territory (see marker "A" on Figure 21). Initial conditions were set as in Table 3.

Table 3. Simulation start conditions

Condition	Lead	Wing
Start Position	N36°04.1724', W075°38.5393'	200' out, 080° line of bearing
Altitude	25,000'	20' down
Airspeed	465KTAS/325KCAS	(Same)
Heading	230°	(Same)
Load-outs	6xAMRAAM, 1xGBU-12	AT-FLIR, no centerline tank
Turn Rate	1.5 °/sec—(34° AOB achieved in 3 sec—"Instrument turn rate")	Any
Red	MiG-29	MiG-29
Start Position	N35°14.6098', W076°32.8189'	"Combat Spread" @1nm
Altitude	20,000'	Co-altitude
Airspeed	300KTAS	(Same)
Profile	70nm—Direct Intercept	Cooperative
"Traditional" Ground Target	"Tank"—N35°33.7825', W076°15.8081', 7' El.	
SAWSI target	N 35°44.8907', W076°45.7952', 29' El.	
Bull's-eye (coordination point)	"Rock"—N35°25.9667', W076°51.7333'	

A script ("the script") was used to ensure the highest possible uniformity between participants (see Appendix C. Scenario Script (Timeline)), and included all necessary communications from the participant's lead and the section's controlling aircraft. The first page of the script is for the basic task wherein the lead aircraft performed a basic intercept, prosecution and kill of an enemy section (marker "B" on Figure 21.) while the participant was simply to stay in formation. The first column on the script indicates the scenario elapsed time coinciding with the start of measurement of the session, while the second column reflects the total time elapsed from unfreezing the simulator. The :30 second disparity

allowed for the subject to get stabilized in parade position since testing had shown that immediately after unfreezing the simulator the lead aircraft entity frequently jumped ahead as much as 500′ (relative to the participant's aircraft) before stabilizing at the commanded airspeed.

Markers "C," "D," and "E" on Figure 21. depict the three tactical turns lead made in consummating the intercept, as noted on the script as the three bolded heading changes. The threat ring marked by "F" and ground target marked by "G" were not used in the primary task.

Figure 21.   Scenario orientation pilot briefing sheet

43

2. Traditional Secondary Task Description

The distracting tasks were performed at the same time as the exact same primary task described above. Again working with the script, this time page 2, participants were directed to use their FLIR for finding, fixing, targeting, and reporting a "tank" (marker "G" on Figure 21) whose coordinates were passed to them by radio once the run had started. This required flying tight formation while writing, entering data into the mission computer, reading back the coordinates from the computer to the controller, then operating the AT-FLIR to prosecute the target, all while maintaining awareness of the spatial relationship between the target and his aircraft. Once the target was positively identified, and the lead had completed clearing the air-to-air picture, the wingman directed the section toward the target for further prosecution (marker "E" on Figure 21). Other than being in parade position, this combination of sub-tasks is familiar and well-practiced in the naval strike community and all our participants were very well trained in the particulars of accomplishing such a mission.

In our session, this task required approximately 47 interactions to the end of the session; however this would only have been a portion of the complete task as at the end of the session the target was not yet destroyed. Similarly to the session where only the basic task was performed, in the traditional task the threat ring (marker "F") was not a part of the scenario.

3. Futuristic Secondary Task Description

Our futuristic task was to clear a surface-to-air missile threat to the participant's section using a tablet computer running a custom designed UAV supervisory control interface named SAWSI. The lead aircraft again flew the exact same profile as in the undistracted task, but to successfully complete the futuristic secondary task and allow section be able to proceed to its (notionally previously) assigned target, the participant had to fly in formation while receiving the verbal threat warning, and then interface with SAWSI to clear the threat. This

required no less than 36 distinct interactions with the SAWSI tablet, but depending on how the participant solved the problem that number could be significantly higher.

Figure 22.   SAWSI tablet mounted on the simulator left canopy rail (Author)

A nominal solution to the threat proceeded as follows (see accompanying figures in Appendix D: SAWSI Scenario Figures):

1.   Figure 28. Participant ("user") is presented with a "God's eye" view of his current situation, depicting his section (in white at the bottom) and four yellow diamond symbols representing Investigation Sites (ISs) where the possible threat may be located. The two open blue symbols located between the user's aircraft and the ISs represent reconnaissance UAVs ("recon") which are orbiting at a pre-assigned location and ostensibly operate at high altitude, above of the threat capability. Slashed symbols are attack UAVs ("attack"), also orbiting at pre-assigned locations, and carrying one weapon each. They operate at lower altitudes so are vulnerable once inside the dashed ring. The solid ring represents the threat distance to the subject's section in their current flight profile.

2.	Figure 29. a User touches a recon to select it and drags a "rubber band" line to designate one of the ISs as its destination. Another is selected for the other recon. (User's position is never updated due to system limitations)

3.	Figure 29. b First recon reaches its destination and after a few seconds "gathering imagery," its symbology turns green indicating that an image is ready for the user to view.

4.	Figure 30. User touches the green recon and SAWSI displays a wide-angle view of the location. Buttons at the bottom of the imagery window allow for designating the IS as a "Threat!" or "Safe."

5.	Figure 30. User (wisely) chooses to zoom in on an area of the image to get a closer look by touching that area of the image.

6.	Figure 31. a. SAWSI displays 3x zoomed image. User can continue to zoom in and out on desired areas until they select either the safe or threat button. SAWSI then closes the imagery window.

7.	Figure 31. b. If user selects safe, IS symbol changes to a green circle (indicating safe) and recon can now be assigned to another destination.

8.	Figure 32. a. After proceeding similarly through the first three ISs, no matter what order they were visited in, deterministically the last one held the actual threat. User opens the last ISs imagery.

9.	Figure 32. b. User sees an area which may contain a threat and touches it to zoom in.

10.	Figure 33. a. Zoomed imagery shows a pair of (surface-to-air) SA-XX missile Transporter Erector Launchers (TELs). User selects "Threat!" and image is closed.

11.	Figure 33. b. IS symbol is replaced by a red and yellow diamond (indicating threat)

12.	Figure 34. a. User reassigns extra recon out of area and assigns first attack to the threat. Attack proceeds inbound for weapon delivery and support.

13.	Figure 34. b. First attack loses contact (shot down) and symbol changes to gray.

14.	Figure 35. a. Second attack is assigned. Attack proceeds inbound.

15.	Figure 35. b. Second attack is shot down. SAWSI logic is deterministic in this field and the first two attacks get shot down each time. Threat battery is now "resetting for next launch."

16.	Figure 36. a. Third attack is assigned.

17. Figure 36. b. Third attack supports its weapon to impact, it turns to return to its previous orbit, and it symbology turns green to signify weapon delivery.

18. Figure 37. a. User taps on recon orbiting overhead the threat to bring up imagery, taps to zoom, and now sees one of the TELs destroyed, one still remaining. User selects "Threat!" again and imagery closes.

19. Figure 37. b. User assigns final attack to threat.

20. Figure 38. a. Fourth attack supports its weapon to impact, returns to its previous orbit, and it symbology turns green.

21. Figure 39. As an alternate technique, several users chose to assign two attacks at once (called "double-tapping" in the community). After the first two attacks are lost, the second two are assigned to the threat together.

22. Figure 39. Both support to impact, and their symbology turns green.

23. Figure 40. a. In either method, after the second delivery, user taps on the recon overhead, zooms in, and sees both TELs hit. User selects "Safe."

24. Figure 40. b. Image clears, threat symbology turns green, and threat rings are removed. Task complete.

E. PROCEDURE

Participants' total involvement was about one hour, with each receiving a pre-event brief and orientation, followed by the four simulator runs, and then a post-event debrief. One researcher conducted the pre-briefs and post briefs, while the other just conducted the experiments runs from the simulator console.

1. Pre-Event Brief

Upon arrival, participants first were given IRB-approved Informed Consent Forms prior to any other briefing. Next, they were asked to complete the Pre-Study Demographic Survey (Appendix A. Pre-Study Demographic Survey) that also served to confirm their qualification for the study. Following that each participant was asked to complete a "test of current alertness state" by undergoing the Trail Making Test, originally named in the Army Individual Test of General Ability (Tombaugh, 2004). This test provided a good standardization test

for the participant's current cognitive capacity in terms of visual scanning, complex attention, psychomotor speed, and mental flexibility, exactly at the time of test.

After a brief overview of the procedure, participants were given a mission brief detailing the primary task and traditional secondary task, using Figure 21. The futuristic secondary task was then introduced by means of a brief slide presentation and a hands-on training scenario imbedded in SAWSI and initiated from the splash screen (see Figure 28 a, Appendix D: SAWSI Scenario Figures). Participants were given as much time as they wanted to familiarize themselves with the operation of SAWSI, and researchers ensured all questions were answered prior to proceeding on to the experiment sessions. Target imagery in the practice scenario was different than in the operational scenario, but the weapon system and TEL shape was the same, thus giving the participant a good idea of what to expect during the experiment.

Finally, each participant was then fitted with one of the two BioHarnesses, and proper Bluetooth connectivity with the training tablet was confirmed to ensure proper logging.

2.   Event

Once complete with orientation, participants were randomly assigned to either A or B treatment order as described below. A within-subjects design was used, so each participant flew each of the three (measured) scenario types plus a familiarization period as described in Section D.

The five-minute warm-up flight, for which the baseline scenario (primary task only, Script 1) was run, served to acclimate them to the dynamics of flying formation off a computer generated entity. This "Session 0" was not measured, and was only used to facilitate participant familiarity.

Next, the first of the three measured five-minute sessions was run where each subject again performed only the primary "baseline" task (again, script 1)

with no secondary task. Each session was recorded for later play-back in an After Action Review room to gather the primary task response data. Since the lead aircraft would perform identical tactics in each run, script timing, communications, and heading changes were carefully controlled to provide the most uniform experience possible both from session to session and between participants. After each session a short 1.5–3 minute break occurred while the simulator operator reset the equipment and scenario.

Secondary tasking for the participants' second measured five minute session was selected randomly from either the traditional task or the futuristic task, while lead again performed the same air-to-air mission profile. Finally, the third session then involved the secondary tasking that was not performed in the second session. During each of the last two sessions, the computer generated lead aircraft entity performed exactly the same A/A mission as in the first session and the subject was reminded to continue to maintain parade position to the best of their ability while at the same time performing the appropriate secondary task.

Upon completion of the third session, researchers entered the simulator enclosure and personally verified the end state of the SAWSI simulation, then reset the tablet for the next participant.

3. De-Brief

Following the simulator runs, each participant was asked to complete the Post-Task Survey shown in Appendix B. Post-Task Survey

F. MEASURES

In gathering required data for the three research questions, both manual and automatic logging and capture methods were employed.

1. Primary Task

The Hornet and Super Hornet simulators have no provision for exporting numerical data (such as distance between two aircraft over time) from a flight

session, so all sessions were recorded and then played back on video monitors in an After Action Debrief Room to gather the data manually. Primary response variable data on inter-plane following distance was produced by setting up a pairing report on screen for distance between the lead aircraft and the participant's aircraft. As the replay ran, the authors read off distance measures (in nautical miles) every six seconds as the numbers ticked by, and manually entered them into an Excel spreadsheet. The system updated the reported distance at approximately a 2Hz rate, so approximately every 12th distance reported was recorded in the spreadsheet. Recording distance every six seconds yielded 50 data points/session, or a total of 150 distances per participant for all three measured sessions.

2. Secondary Tasks

Since the traditional secondary task was performed on the simulator itself, no data about how the task was performed could be readily gathered during playback. Thus, the researchers simply recorded in real time a binary (Y/N) task completion metric, as determined by what the participant said on the radio and what the researcher observed on the cockpit monitor repeaters at the simulator operator console.

Futuristic secondary task was precisely logged on the SAWSI tablet, and every user interaction was reflected sequentially.

3. Physiologic Data

The SAWSI log also recorded a bio-data string (as introduced in C.3) from the BioHarness on a 1Hz cycle whenever the SAWSI program was running and connected on the Bluetooth link.

V.   ANALYSIS AND RESULTS

A.   DATA COLLECTION AND PREPARATION

Data was collected utilizing five sources: 1) a demographic survey administered to each participant during the pre-flight briefing immediately prior to their simulator session, 2) raw simulator data on each participant's flight performance recorded automatically by the simulator, 3) SAWSI interactions by each participant, 4) biometric data recorded by a BioHarness fitted around the chest of each participant, and 5) a post task survey given to each participant following their simulator session.

1.   Surveys

a.   Demographic Survey Data

Following the initial experiment brief and signature of the research consent form, demographic data was collected on all participants to obtain the experience levels and backgrounds. Data from all surveys was compiled to a Microsoft Excel spreadsheet to aid in further analysis and correlation with simulation data. See Table 7.

b.   Post Task Survey Data

Participants completed a brief survey following completion of their sessions in order to provide perceived workload rankings for each scenario, UAV interface ease of use, and any additional comments regarding the experiment. Likert scales were utilized and compiled to a Microsoft Excel spreadsheet for additional analysis. See Table 7.

2.   RAW SIMULATOR DATA

As comprehensively described in the methods section, raw following distance data was collected from each of the participants by replaying the flights and transcribing following distances from the computer generated lead of each

subject at 6 second intervals into a Microsoft Excel spreadsheet. While the debrief system provided numerous other parameters, including airspeed, altitude, heading, latitude and longitude, etc., our primary dependent variable was limited to just the following distance as measured in the XYZ axis.

No consideration was given to the minor differences in altitude each participant experienced while conducting formation flight with the lead aircraft. Since following distances recorded from the simulator were presented in nautical miles, we applied a simple conversion formula in Microsoft Excel to present the following distances in feet. This provided a far more discriminating measure of performance during close formation flights since the standard, in reality, is on the order of slightly more than a wingspan distance between aircraft (~80 ft).

Out of 102 participant sessions (three 5 minute simulator sessions for each of the 34 participants), only one participant's data (participant #30) was discarded due to clear failure of the pilot to adhere to instructions to maintain close formation flight set forth in the preflight brief. Participant #30's following distance mean and sd for his secondary task sessions were far outside the normal range expected of professional pilots (participant 30's following distance mean and sd were 703.99 and 318.53 compared to group following distance mean and sd of 242.42 and 107.39). The remaining 33 data sets were found to be valid to include for further analysis.

a. Smoothing

Based on the initial plots of following distances for each of the participant's simulator sessions, it was apparent there was a significant difference in the signal noise between the two TOFTs (C/D and E/F) recording quality. The 8 participants' recordings utilizing C model F-18 simulators clearly showed smoother following distance performance compared to the 25 participants who flew the F model F-18 simulator. As mentioned in the methods chapter, based on the researchers' observations of participants during their simulator sessions, the "jumpiness" of the computer generated lead was

52

generated by signal noise due to the fidelity of the recording and playback system, and undoubtedly the cause of the wild fluctuations in the majority of unfiltered following distances.

In reality, the participants, while not always maintaining an ideal trail position in formation, were far more stable in maintain formation than the data indicated. All participants were qualified, proficient and current in the F-18 and routinely flew their respective simulator (C/D or E/F model), an appropriate level of proficiency and steadiness in formation flight is to be expected. Accordingly, to mitigate erratic following distances inherent to the simulator recording and playback system, a smoothing function was applied to remove the noise while attempting to maintain a faithful indication of following distance.

An Excel function (see EQN 1) was written to compare following distances at each 6 second interval with a smoothing distance individually developed for each session to reduce signal noise. By applying the smoothing distance in Formula X below, obvious signal noise was minimized to depict more realistic following distances for each participant (individual smoothing distances ranged between 215 and 535 ft based on signal noise). Below is the smoothing function equation, followed by a detailed explanation.

EQN 1.

Smoothing Function =
IF(AC$3=0,AB9,(IF(AB9>AC$3,IF(AB8>AC$3,MIN(AB7,AC$3),AB8),AB9)))

If (filtering distance = 0) then use the original following distance value.
Else, (if original following distance > the filtering distance)
      Then (if the previous cell following distance > filtering distance,
      Then choose the minimum of either the next previous cell's following distance or the filtering distance.
Otherwise use the original following distance.

First, the smoothing distance for each session was evaluated. If the smoothing distance = 0 ft, that meant the researches believed the participants performance was not influenced by signal noise and therefore all following

distances were maintained. Next, if the selected cells [following distance] value was *less* than that session's non-zero smoothing distance, the original cell value was *still* maintained.

If the following distance was greater than the smoothing distance, the *previous* 6 second interval cell value was compared to the smoothing distance. If that value was less than the smoothing distance, it was carried over to the currently selected cell as a smoothed distance. If that value exceeded the smoothing distance, the minimum of either the previous interval or the smoothing distance was carried over.

Every effort was made to apply a filtering distance to reduce only the obvious signal noise generated by the recording and playback system, and not artificially improve the participant's performance.

Figure 23 provides an example of a clipped portion of the data from participant #6's baseline session to demonstrate the resulting application of the smoothing distance function applied to the raw following distance value in cell AB9 and based on a 300 ft smoothing distance (indicated by cell AC3 circle).

Since cell AB9's value was less than the smoothing distance of 300 ft, the original following distance was maintained. However, for cell AB11, since the following distance of 564 ft exceeds the smoothing distance, it is compared to the previous interval's value. Since AB10's value is 180 ft, which is less than the 300 ft smoothing distance, that value is carried over to AB11 as a smoothed distance. We believe the application of this smoothing distance methodology has mitigated most of the system-generated noise. Not only did it *not* artificially improve the participants' apparent performance, we believe it successfully unmasked a more accurate representation of their sessions. This observation is based on both direct monitoring of all participants' sessions, which were far smoother than the data indicates, as well as the professional experience of the aviators/researchers who conducted the experiment.

54

Figure 23. Example data and smoothed distance

Experienced military aviators are unlikely to make a basic error in judgment by conducting formation flight with high opening and closure rates (closure rate is the absolute difference in velocity between the lead and trail aircraft). The lowest applied filtering distance was 210 fps, or 21 kts, while the highest applied filtering distance was 535 fps, or 53 kts, which far exceeds the usual 25 kt maximum closure speed usually seen in formation flight. Applying these criteria to the graphs of following distances over time clearly revealed obvious signal noise. It was therefore relatively easy to observe the influence of small changes in filtering distance to reduce noise. Again, following analysis, it was found that the raw data was in fact significant enough to prove our hypothesis without filtering. However, we felt that the filtered data more accurately measured actual performance.

Table 4 summarizes the filtered session mean and sd of following distances for each session of all participants, with the exception of the outlier, participant #30, on which our statistical analysis was conducted.

Table 4.   Summary of smoothed following distance mean and
standard deviation for participants

Participant #	Filtered Session Following Distances (ft)					
	Baseline mean	Baseline sd	Traditional mean	Traditional sd	Futuristic mean	Futuristic sd
1	255.38	85.45	414.92	205.83	803.28	558.06
2	172.72	66.70	308.39	123.85	245.80	111.54
3	177.65	74.57	249.18	93.76	256.08	81.29
4	162.88	65.31	284.59	111.21	159.65	55.71
5	201.65	83.22	192.47	72.60	289.92	133.35
6	170.00	56.50	160.08	52.35	224.94	60.97
7	236.59	41.14	295.06	67.88	372.16	85.24
8	148.67	36.95	190.71	62.11	185.53	52.08
9	168.98	64.83	252.71	63.05	249.88	89.08
10	149.41	54.82	183.49	50.78	205.06	62.79
11	144.75	72.51	211.41	95.95	199.10	63.20
12	151.57	85.46	180.94	87.82	217.57	90.93
13	183.06	100.11	230.35	105.48	230.35	87.64
14	110.71	22.59	120.59	42.79	129.48	69.89
15	159.41	150.18	337.18	208.53	656.52	474.01
16	336.00	400.96	316.12	120.42	419.88	324.64
17	192.47	74.64	273.61	80.63	371.06	109.12
18	164.24	73.57	340.27	106.10	312.08	107.14
19	146.35	72.53	220.24	103.20	246.12	70.83
20	142.24	66.72	221.88	85.26	240.24	100.64
21	120.94	54.48	146.71	64.54	209.88	70.47
22	139.65	41.71	198.12	84.99	227.18	80.74
23	228.35	97.87	270.00	112.57	370.32	109.09
24	267.76	129.82	472.00	337.03	762.48	558.66
25	127.76	25.29	125.41	28.23	180.48	81.40
26	122.35	26.81	260.71	122.33	153.60	43.11
27	102.47	37.14	118.82	36.62	167.16	76.55
28	176.00	52.19	257.53	96.66	311.65	72.69
29	114.35	45.51	226.35	103.24	175.06	88.37
31	136.71	52.64	172.71	100.16	130.98	65.53
32	135.76	56.71	170.59	69.69	163.10	88.34
33	130.82	46.62	210.49	79.98	187.41	46.13
34	109.65	45.74	240.16	119.90	220.12	72.08

Figure 24 and Figure 25 depict plots of following distances over time for one participant's sessions both before and after applying the filtering function. For this particular data set, a smoothing function was applied only to the baseline data set, since it clearly exhibited signal noise or jumpiness contributed by the low fidelity of the simulator playback system. Comparison of both raw and

filtered plots indicates the smoothing function retained the same general pattern of data, and reduced the level of signal noise as expected.
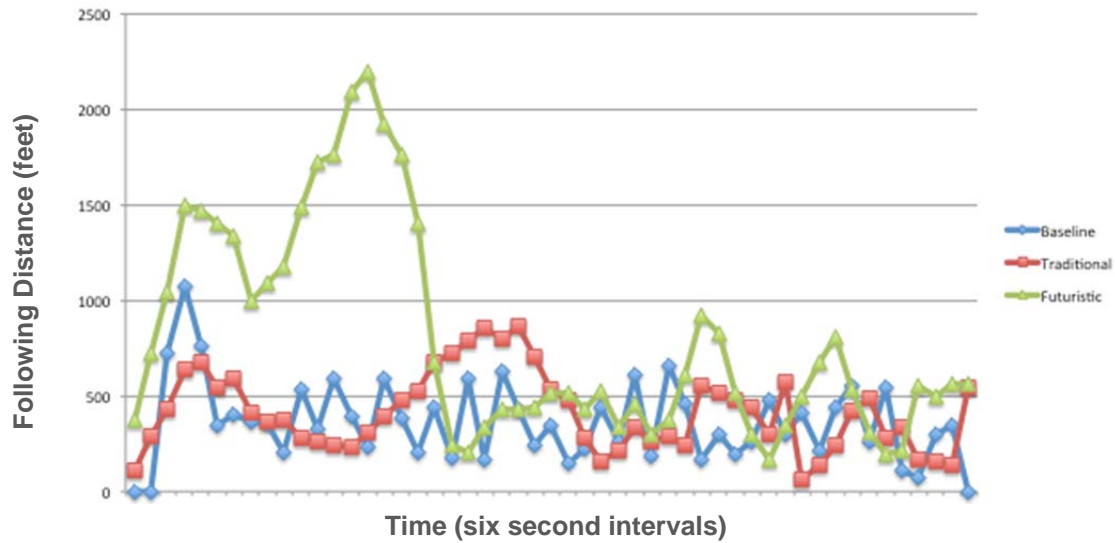
Figure 24. Unfiltered data plot of time vs. following distance by scenario
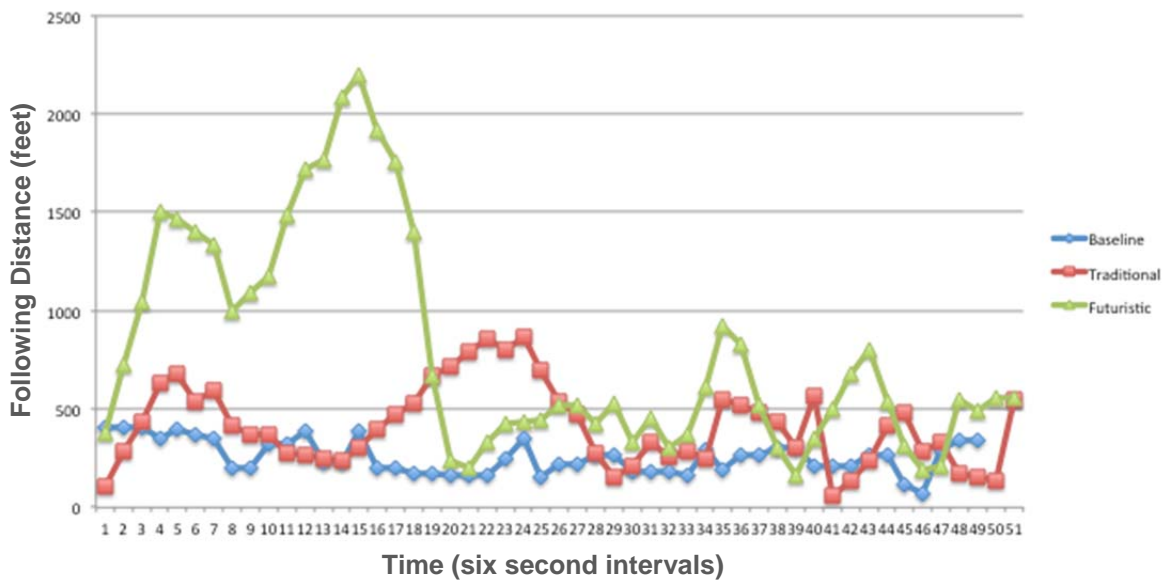
Figure 25. Filtered data plot of time vs. following distance by scenario

3. SAWSI DATA

Data input for futuristic [UAV Supervisory Control] task was recorded to evaluate participant completion of the secondary task. This data was treated as binary: either all targets were destroyed or they were not. Based on this criteria for mission accomplishment, 76% of all participants completed the futuristic secondary task. Results were then compared with following distance to determine correlation.

4. BIOHARNESS DATA

While pretesting and early sessions utilizing the BioHarness appeared to indicate correct logging of bio-parameters (heart rate, respiration, etc.), the data was discovered to be inconsistent and unreliable. Unfortunately no bio-data could be salvaged and correlated to the simulator data.

5. STATISTICAL ANALYSES USED

To evaluate the effect of secondary tasks on following distances, the paired-t analyses was used, except in cases in which the assumption of normality was not met. In these cases, the Wilcoxon signed rank test was implemented. Additionally, demographic data from the pre-flight survey was used to identify influences on following distance mean and sd based on correlation coefficients. An alpha level of .05 (1 tailed) was used for all analyses testing hypotheses.

B. PRELIMINARY ANALYSES

1. PARTICIPANTS

The 34 participants in this experiment were all volunteer Active Duty U.S. military aviators based at Naval Air Station (NAS) Lemoore, CA where they fly F/A-18 Tactical Operational Flight Trainers (TOFTs) (either the C/D model or E/F model). With the exception of one U.S. Marine pilot, all participants were U.S. Naval Aviators in the ranks from O-2 to O-4 with the majority O-3's, and ranging in age from 26 to 39 years old. The least experienced aviator had 400 total flight hours and the most experienced had 3200 total flight hours. All aviators were

volunteers from the operational and training squadrons at NAS Lemoore. At this point in their careers, all aviators had achieved a minimum Strike Fighter Weapons and Tactics level II Qualification (fully trained and experienced in our primary task), which was a significant discriminator for participating since our objective was not to evaluate their individual formation flight proficiency, but what effect, if any, would secondary tasks (i.e., cockpit tasks or distractions) have on the quality of their formation flight. Table 5.   provides descriptive statistics and relevant demographic characteristics of the participants.

Table 5.   Summary of demographic characteristics

	Median	Mean	SD
Age	31 yrs	30.9 yrs	3.6 yrs
Total Flight Hours (TFH)	1250 hrs	1475.4 hrs	693.3 hrs
F-18 Actual Hours	1000 hrs	1146.8 hrs	615.2 hrs
Hours Flown Last 30 Days	10 hrs	9.4hrs	5.8 hrs
Rank (Grade)	3 x O-2's / 22 x O-3's / 9 x O-4's		

2.   RESULTS OF POST TASK SURVEY

Following each participant's simulator session, a post task survey was administered as described in the method chapter. The post task survey employed a Likert scale to assess the perceived workload during each of the three simulator sessions, as well as how challenging the SAWSI interface was to employ. The survey also offered each participant an opportunity to make any additional comments concerning the conduct of the experiment and their thoughts on how to best employ a potential UAV supervisory control inside the cockpit.

Figure 26 presents a diverging stacked bar chart summarizing participant Likert ratings of perceived cockpit workload grouped horizontally by scenario. Each bar indicates by color the percentages of responses from participants.

Responses are centered on zero, which reflects an "ideal" workload. Understandably, participants found the baseline scenario relatively easy compared to the sessions with traditional and futuristic secondary tasks, with nearly all baseline responses indicating the scenario was "boring" (35%) to "low" (44%) workload, with an additional 17% judging somewhere in between "boring" and "low," and only one respondent (3%) claiming an "ideal" workload. However, both traditional and futuristic sessions clearly presented a more challenging perceived workload to participants, and indicated remarkably similar distributions to one another. Approximately 45% of all participants perceived the traditional and futuristic scenarios "low" to "ideal" workload, while approximately 55% perceived the scenarios as "ideal" to "too high" or "challenging." Only one participant in the traditional scenario judged his workload as "too high" and only one participant in the futuristic scenario judged his workload as "challenging." This data clearly provides preliminary support for hypothesis 1.
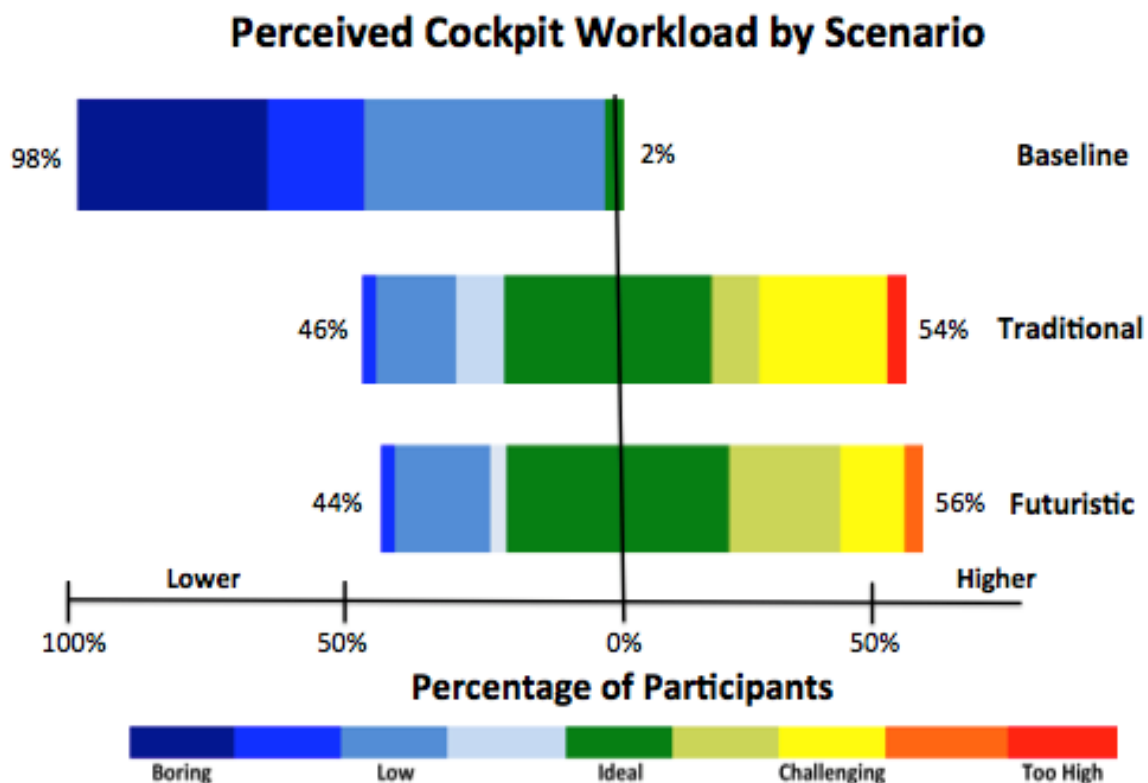
Figure 26. Perceived cockpit workload by scenario

C.   HYPOTHESIS TESTING

1.   Hypothesis 1 Results

Hypothesis 1 investigated the differences in consistency of following distances between pilots flying close formation with and without secondary tasks. As previously discussed, the secondary tasks selected for this experiment involved realistic tasks, requiring divided attention between flying with eyes "outside" the cockpit and operation of weapons systems through interfaces with eyes "inside" the cockpit.

H_{01}: There will be no significant differences in the execution of the primary task of maintaining safe, effective flight when presented with either a traditional or potential future [UAV supervisory] secondary task.

H_{A1}: There will be significant differences in the execution of the primary task of maintaining safe, effective flight when presented with either a traditional or potential future secondary task.

Result: We tested the hypothesis by comparing the sd following distance and the mean following distance across simulator sessions. Results revealed a significant difference in sd's of following distances between traditional and futuristic sessions, in which participants in the traditional task showed greater consistency in following distance than when in the future secondary task. However, significant p-values between baseline and traditional tasks (p=.000032) and baseline and futuristic tasks (p=.000061) were found, and overwhelmingly indicated secondary tasks have a deleterious impact on the consistency in maintaining following distance, the primary task of flying. Table 6 outlines these results.

A similar pattern of results was found for the mean following distance, in which participants had shorter mean following distance means in the traditional than futuristic task, and had the shortest mean following distance in the baseline

61

task (see Table 6). Raw data results are included in Table 8 Appendix E. Data Table, however, the results were equally impressive: only 2 of the 9 *t*-test results failed to indicate significance.

Table 6.   Paired *t*-test results for filtered data

	t-test results for all models					
	Filtered Data					
	AVE			SD		
	Base vs Trad	Trad vs Future	Base vs Future	Base vs Trad	Trad vs Future	Base vs Future
means	166.28 237.99	237.99 281.03	166.28 281.03	74.58 99.86	99.86 128.52	74.58 128.52
t test statistic	7.069370	2.270570	4.890540	2.088480	1.737260	2.623490
p-value	0.000000	0.015021	0.000014	0.022398	0.045979	0.006612

	t-test results for E/F model					
	Filtered Data					
	AVE			SD		
	Base vs Trad	Trad vs Future	Base vs Future	Base vs Trad	Trad vs Future	Base vs Future
means	163.30 233.32	233.32 257.37	163.30 257.37	62.82 91.48	91.48 100.16	62.82 100.16
t test statistic	7.609630	1.310680	4.634060	4.903950	0.568350	2.016520
p-value	0.000000	0.101185	0.000053	0.000027	0.287538	0.027534

	t-test results for C/D model					
	Filtered Data					
	AVE			SD		
	Base vs Trad	Trad vs Future	Base vs Future	Base vs Trad	Trad vs Future	Base vs Future
means	181.85 252.60	252.60 354.99	181.85 354.99	111.33 126.06	126.06 217.17	111.33 217.17
t test statistic	2.259410	2.045960	2.416250	0.302220	2.090600	1.718870
p-value	0.029191	0.040003	0.023173	0.385632	0.037452	0.064664

2.   Hypothesis 2 Results

Hypothesis 2 investigated the effect of total flight hours and currency on following distances, specifically with respect to the futuristic scenario in which participants operated a UAV supervisory console.

H_{02}: No difference will be observed in performance of the primary task when treated with an operational task distraction of a UAV supervisory task between pilots of higher proficiency and currency (currency being how much they've flown lately).

H_{A2}: Pilots of higher proficiency and currency will show less change in performance of the primary task when treated with an operational distraction of a UAV supervisory task.

Naturally, we expected following distance performance to improve relative to experience (i.e., participants would maintain following distance with more consistency/less variance). Interestingly, we found a bimodal pattern of increasing following distance averages at approximately the 600 and 1800 total flight hours. This discovery may be explained by lack of currency resulting from periods of extended shore duty with limited opportunity to fly in order to maintain proficiency

Figure 27's graphs mean following distances by total flight hour of participants for each scenario and revealed total flight hours alone lowered the average following distances only very slightly: ~25ft for the baseline scenario to ~35 ft for the traditional and futuristic sessions.
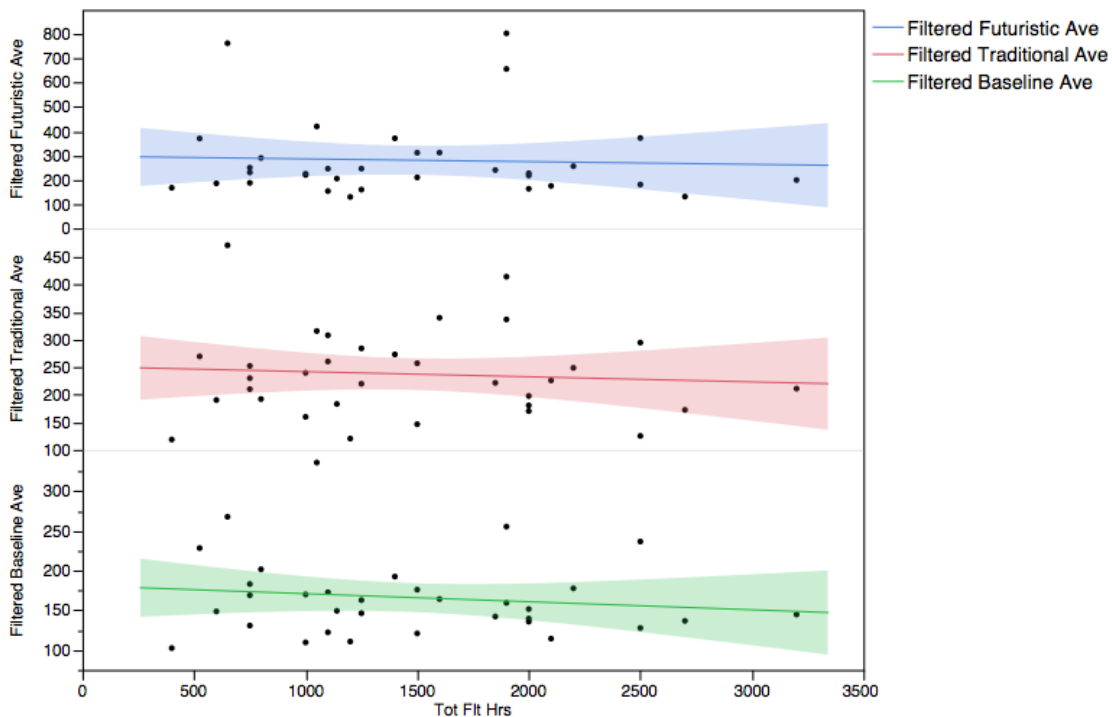
Figure 27. Total flight hours vs. filtered mean following distance

As expected, the inclusion of a secondary task generally increased the average following distances across the board for all participants (in only 3 sessions out of 78 did a participant achieve a *lower* average following distance during a session including a secondary task than they achieved in the baseline with *no* secondary task). However, attempts at finding a satisfactory regression model failed due to unequal variance of data, even after applying power transformations. Therefore, results failed to reject the null hypothesis.

Analysis leads to the conclusion that inclusion of a secondary task has equal effect on performance of the primary task of formation flying when treated with an operational task distraction of a UAV supervisory task regardless of flight experience and currency of the participant.

3.   Hypothesis 3 Results

H_{O3}: The subject's heart rate, respiratory rate, and trunk position will be not different when performing the traditional versus future secondary tasks.

H_{A3}: The subject's heart rate, respiratory rate, and trunk position will be significantly different when performing the traditional versus future secondary tasks.

Due to incomplete recording of BioHarness data, we were unable to investigate hypothesis 3 (exploratory): What is the difference in physiological measures such as heart rate, respiration rate, and posture when performing the traditional versus future secondary tasks? While we fitted participants with the BioHarness in accordance with manufacturer's instructions, and pretested the logging, actual recordings were incomplete and/or widely fluctuating, making analysis of the bio-data impossible.

VI. CONCLUSION AND DISCUSSION

A. DISCUSSION

This research grew out of a collective desire of the researchers to better understand the effects operationally relevant secondary tasks have on the primary task of actually flying fighter aircraft, in particular the difference in performance between a traditional secondary task that military pilots are well trained to conduct and a futuristic task such as a supervisory UAV interface. While many studies have been conducted on UAV control (Wickens 2005; Rodas & Veronda 2011) or on distracted driving (Strayer, Drews & Couch, 2006), very little has been investigated in the high workload environment of the military aviation cockpit. Our experiment leveraged the high-fidelity of military TOFTs and the operational expertise of seasoned Naval Aviators to produce a highly valid result which can now be applied to further research on the management and organization of pilot workload in order to improve safety and mission effectiveness.

1. Flight Task Performance

While the experimental design of tasking participants to fly each simulator session in parade position was not operationally realistic, it was deliberately chosen in order to set the conditions to more effectively measure and evaluate the differences in performance between the baseline, traditional, and futuristic tasks. Our dependent variable, following distance, was selected as the primary indicator of performance since it would require participants to dedicate significant effort to maintaining consistency.

Although we anticipated a normal distribution of average following distances for all participants due to individual experiences and skills, the "textbook" standard following distance for parade formation was not as important as measuring its variation. Variation would provide a clear indicator of the level of distraction secondary tasks were causing.

2. Secondary Task Performance

Secondary tasks were selected to represent realistic actions within the cockpit but also purposefully distract each participant from his primary task of flying. Although completion of secondary tasks was not required or necessarily relevant to this study, a preliminary look into the correlation between the primary and secondary task completion indicates that one is not predictive of the other. This phenomenon may indicate an interesting line of further investigation.

B. HYPOTHESES RESULTS

Results of test for hypothesis 1 provided clear indication that a futuristic task is more significantly more challenging than a traditional task and that both secondary tasks increase the average mean following distance compared to undistracted flying. Plots of the primary response variable vs. time illustrated the unmistakable variation in following distance among most participants, particularly around the three timestamps where scripted turns took place with secondary tasks being processed in the background.

However, for hypothesis 2 there is no evidence that pilot experience and proficiency significantly improves following distance performance when distracted. As experience increased there was a very slight decrease in average following distance, but variation was still present and there were indications that some degree of individual skill (or lack thereof) regardless of experience contributed to effecting performance. Linear regression of mean following distances across all three scenarios show nearly identical and essentially flat slopes across the 400 to 3200 hour range of total flight hours of our participants.

Finally, due to incomplete recording of BioHarness data, the exploratory hypothesis 3 question of whether a difference in physiological measures such as heart rate, respiration, and posture is detectable and correlated to secondary tasks could not be analyzed.

C. LIMITATIONS

Our primary limitation in the experiment revolved around the classified nature of the flight simulators. We attempted to obtain the raw data files from each simulator session in order to perform a more thorough analysis of following distances. However, during pre-experiment site surveys we discovered the recording and playback systems were not designed to provide any form of exportable data output. Only by replaying each session in real time and then manually transcribing following distances were we able to compile sufficient data for analysis. Additionally, we discovered a significant difference in noise induced on the following distance between the C/D and E/F models. Investigation and requests for information revealed the simulators were never intended to facilitate evaluation of formation flight to the degree we desired. However, we conducted a proof of concept that confirmed they were capable of supporting the designed scenarios. While our experiment attracted enough interest from aviators stationed at NAS Lemoore, additional participants may have provided an opportunity to expand our analysis within simulator types. There were 26 participants who flew the E/F model but only 8 participants who flew the C/D model.

D. STRENGTHS

Our experiment was well designed from the standpoint of having a very controlled environment inside the simulator building. The high fidelity flight simulators (TOFTs) provided a fully accredited training platform that was ideal for evaluating all necessary flight parameters. Furthermore, the simulators were supported by a professional full time staff that greatly facilitated the development of highly tailored and scripted scenarios that contributed to reliably repeatable flight profiles as verified during pre-testing. The experiment also provided a natural progression from pilot study in order to further evaluate the effects of pilot workload in the cockpit and was extremely well supported and endorsed by the Strike/Fighter community.

Last but not least, the chain-of-command and simulation center staff were extremely supportive and recognized the value of investigating workload in the cockpit. They facilitated the scheduling of simulators for several sessions of pre-testing as well as the two weeks required for conducting the experiment, and also greatly facilitated coordination throughout the training squadrons for volunteers to participate in the experiment.

E.   IMPLICATIONS OF THE RESULTS

1.   Feasibility of Additional Tasks

The results of our experiment indicate there are potential benefits to quantifying the impacts of secondary tasks on the performance of primary task in order to achieve a balanced workload within the cockpit without sacrificing safety of flight. Analysis indicates the inclusion of a relatively simple but futuristic task in the form of the SAWSI can be rapidly mastered by aviators, but further integration with onboard avionics would be required to leverage the full capability of the features introduced through this interface. The increasing utilization of UAVs throughout the military obviously mitigates the physical threat experienced by manned aircraft, and this SAWSI offers a novel alternative to reducing the threat to pilots by providing the option to engage enemy targets (either air or ground) with a UAV wingman. Furthermore, SAWSI has the potential to extend the capability of pilots to conduct air to ground missions since missions conducted from a locally controlled asset are always preferential to those conducted via Telepresence.

2.   Multi-place Cockpits

Analysis also indicated the addition of any secondary task reduced performance from the baseline undistracted flying scenario, highlighting the utility and value of a second crewmember. As depicted by the POC curve previously discussed, a second crewmember (Weapons System Operator (WSO), Bombardier Navigator (BN), etc.) helps relieve the pilot of any distractors to flying, particularly during high intensity mission phases. Since today's missions

are becoming more complicated, not less, it seems reasonable to predict that our results would have been very different had we been evaluating workload in two-seat aircraft.

F. SUMMARY

Our investigation into the differences in performance of the primary flight task with and without secondary tasks yielded significant results. Clearly the addition of secondary tasks increased the average following distances for the majority of participants, and a futuristic UAV supervisory task resulted in a statistically significant difference in following distance compared to a traditional air to ground task. Regardless of the group examined (all participants, C/D model participants, and E/F participants), participants had consistently shorter mean following distances in the baseline than in the other 2 scenarios. There was also a consistent pattern that participants achieved shorter mean following distances in the traditional versus futuristic tasks: a significant difference for all participants ($p = .015$), a trend towards significance for E/F participants ($p = .101$), and a significant difference for C/D participants ($p = .040$).

Furthermore, there is evidence that adding a secondary task can affect consistency in following distance. Based on standard deviation, analyses between all participants and the E/F participants indicate there was less variability in the following distance in the baseline scenario than in the other two scenarios. Analyses between all participants and the C/D participants revealed that participants also showed less variability in the following distance in the traditional versus futuristic task. Therefore, there is some evidence that adding a secondary task can affect consistency in following distance, and that a more challenging (i.e.; futuristic) secondary task affects the consistency of following distance more so than a traditional secondary task. However, it remains unclear whether these results are due to simulator differences or task differences. A future study that uses only one simulator model or larger samples for both simulators is required to better understand these results.

Although the BioHarness data was incomplete and consequently affected our ability to investigate the exploratory hypothesis, we feel there is potential for further analysis in this area as described in the recommendations section.

G.   RECOMMENDATIONS

We have several recommendations for future work which may be worthy of thesis opportunities for students interested in cockpit workload and the effects of secondary tasks on flight performance.

1) Even though we were unable to optimize our smoothing algorithm for following distance in order to improve the processing of raw following distance data, we believe with additional effort a more accurate model could be developed to better portray the reality of each pilots performance. Even though the raw data itself turned out to be statistically significant, there is further insight to be gained from attempting to mitigate the effects of recording and playback induced noise on F-18 E/F model following distance data.

2) Additionally, we feel it would be interesting to attempt to salvage the incomplete BioHarness data and/or replicate the experiment with a more comprehensively verified BioHarness setup in order to achieve valid recordings from participants. The additional data would no doubt contribute to a better understanding of cockpit workload and the measure of stress on participant performance.

3) A similar study could also be conducted involving Weapon Systems Operators (WSOs) in order to evaluate high workload missions requiring high coordination between crewmembers compared to the same mission with only a pilot. While our study involved relatively low stress tactical maneuvering and secondary tasks, an experiment with increased crew demands could shed light on whether primary task performance remains comparable.

4) With regard to the concept of airborne Semi-Automated Wingman supervisory control, an additional study could be performed using aircrew

supervision of larger flocks of UAVs, or protocols for assignment/reassignment of the UAVs. Both this and the suggested WSO study would help to further define upper bounds of pilot-tolerated workload. While our investigation served to quantify the transfer of concept from the driving and civilian flying realms, we did not investigate varying levels of distraction that these suggested studies could.

5) Finally, although SAWSI seems to have been well received by our participants, further study into refining the interface, adding functionality, and adapting it to actual military use seems to be warranted. Our scenarios included only one target per scenario while in reality a crew could potentially be attacking multiple targets. A dedicated experiment where the crew is supervising multiple flights of UAVs would provide invaluable data on the effect a SAWSI type interface has as a combat multiplier.

THIS PAGE INTENTIONALLY LEFT BLANK

Demographic Survey

Date:

Subject ID:

1. Date of birth: _____

2. Gender: Male   Female

3. Rank: _____

4. Total flight hours: _____

5. Total flight hours in F-18's: _____

6. Hours flown in the last 30 days: _____

7. Hours flown in the last week: _____

8. Approximate total hours flown in the simulator: _____

9. Approximate hours flown in the simulator in the last 30 days: _____

10. Approximate hours flown in the simulator in the last week: _____

11. Hours flown today? _____

12. Last time you flew in an F-18 simulator?

13. Last time you flew formation?

14. How many hours of sleep did you get last night? _____

15. How would you rate your health?

_____Excellent_____Good_____Average_____Poor

16. Are you right handed or left handed?

_____Right_____Left

17. Do you play video games often? If so what type?

_____Flight Sims_____FPS_____Command and Control

18. Do you own or operate a touch screen based tablet or smartphone?

_____Yes_____No

APPENDIX B. POST-TASK SURVEY

Date:

Subject ID:

Please answer the questions below regarding the experiment you just completed, using the following scale to answer the first three questions.

| Complete | Low | Ideal | Challenging | Helmet |
| Boredom | Workload | Workload | | Fire |
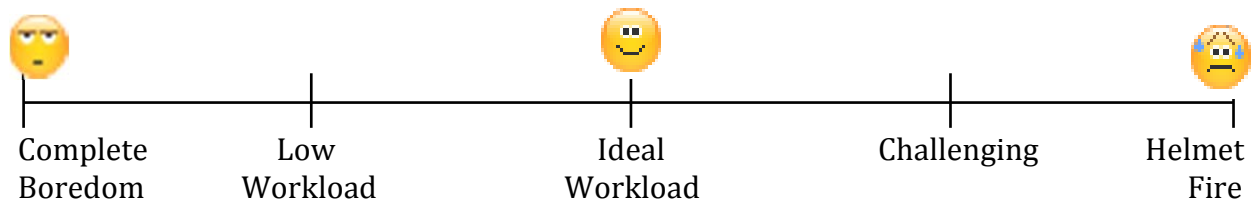
1. For each phase of the experiment, please rate your average level of workload for that section.

Air-to-air only	
FLIR Pod A/G	
UAV Supervisory	

2. How challenging was the UAV supervisory control task?

UAV Control Interface Task	
	Boring About Right Too Much Work

3. Do you have any questions?

4. Any comments or suggestions?

APPENDIX C. SCENARIO SCRIPT (TIMELINE)

IC set: Blue: 25K / 465T/325CAS / Hdg 230 / 500' out, 20' down, 080° line of bearing
            Lead: 6xAMRAAM, 1xGBU-12
            Subject: Pod (1684), no centerline tank
            Start: (N 36° 04.1724', W 075° 38.5393')
            "Instrument" turn rate—<u>1.5</u> ° / sec—(34° AOB, achieved in 3 sec)
      Red: 20K' / 300T / Direct / 70nm / 2xMiG 29, Combat spread 1nm
            Start: (N 35° 14.6098', W 076° 32.8189')
Traditional target location: "Tank" (T-71)—N35° 33.7825', W076° 15.8081', 7' Elev.
SAWSI target (WP 9): N 35° 44.8907', W 076° 45.7952', 29' Elev.
Bull's-eye (WP 20): "Rock"—N 35° 25.9667', W 076° 51.7333'

<RECORD!>
<u>   Basic Task   </u>

0:00	**0:30**	Fight's on, tapes on
0:05	**0:35**	Picture: Group ROCK 129 / 20, track NE Rage commit
0:18	**0:48**	Rages left 20 **Heading- 210**
0:40	**1:10**	Picture: Single Group ROCK 121 / 21, hot, hostile
0:50	**1:20**	Rage 21 contact BRA 232/49
2:00	**2:30**	Rage 21—35 miles
3:12	**3:42**	Rage 21 Fox 3, two-ship
3:15	**3:45**	Rages Crank left **Heading- 200**
4:13	**4:43**	Rage 21 timeout single group
4:20	**4:50**	Single group vanish, picture clean
4:25	**4:55**	Rages continue, right 300 **Heading- 300**
5:00	**5:30**	Knock it off (Rage 21 KIO, *Rage 22 KIO*)

<u>   Task 1</u>
0:00   Laser code 1684

<u>Run</u>
0:00   **0:30**   Fight's on, tapes on

0:05   **0:35**   Picture: Group ROCK 129 / 20, track NE
                         Rage commit

0:18   **0:48**   Rages left 20
                         Heading- 210

0:40   **1:10**   Picture: Single Group ROCK 121 / 21, hot, hostile

0:48   **1:18**   Rage 22, stand by for target coordinates

0:50   **1:20**   Rage 21 contact BRA 232 / 49

1:00   **1:30**   Rage 22, target coordinates as follows: N 35° 33.8201', W 076° 15.8493',
                         elevation: 7'. Tank in the open. Standing by for system read-back.

2:00   **2:30**   Rage 21— 35 Miles

      Upon correct readback: That's a good readback, report target capture when able

3:12   **3:42**   Rage 21 Fox 3, two-ship

3:15   **3:45**   Rages Crank Left
                         Heading- 200

4:13   **4:43**   Rage 21 timeout single group

4:20   **4:50**   Single group vanish, picture clean

4:25   **4:55**   Rages continue, Rage 22 say bearing to target
                         Rages right ____, flow to target
                         Heading- ____ (~300°)

5:00   **5:30**   Knock it off (Rage 21 KIO, *Rage 22 KIO*)

<u>   Task 2</u>

0:00		Target is in WP 9.

0:00   After the fight's on: Locate, fix and prosecute SA-17 with SAWSI. Report threat clear to your lead.

Run

0:00	**0:30**	Fight's on, tapes on
0:05	**0:35**	Picture: Group ROCK 129 / 20, track NE Rage commit
0:18	**0:48**	Rages left 20 **Heading- 210**
0:30	**1:00**	Rage 22, 21, Locate, fix and prosecute SA-17 with SAWSI. Report threat clear
0:40	**1:10**	Picture: Single Group ROCK 121 / 21, hot, hostile
0:50	**1:20**	Rage 21 contact BRA 232 / 49
1:10	**1:40**	22, 21—we're staying outside the threat area here, looks like we should be good with this heading
2:00	**2:30**	Rage 21—35 miles
3:12	**3:42**	Rage 21 Fox 3, two-ship
3:15	**3:45**	Rages Crank left **Heading- 200**
4:13	**4:43**	Rage 21 timeout single group
4:20	**4:50**	Single group vanish, picture clean 22, 21 –Say status threat?
4:25	**4:55**	Rages continue, right 300, flow to target *(if threat clear)* **Heading- 300**
5:00	**5:30**	Knock it off (Rage 21 KIO, *Rage 22 KIO*)

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX D: SAWSI SCENARIO FIGURES
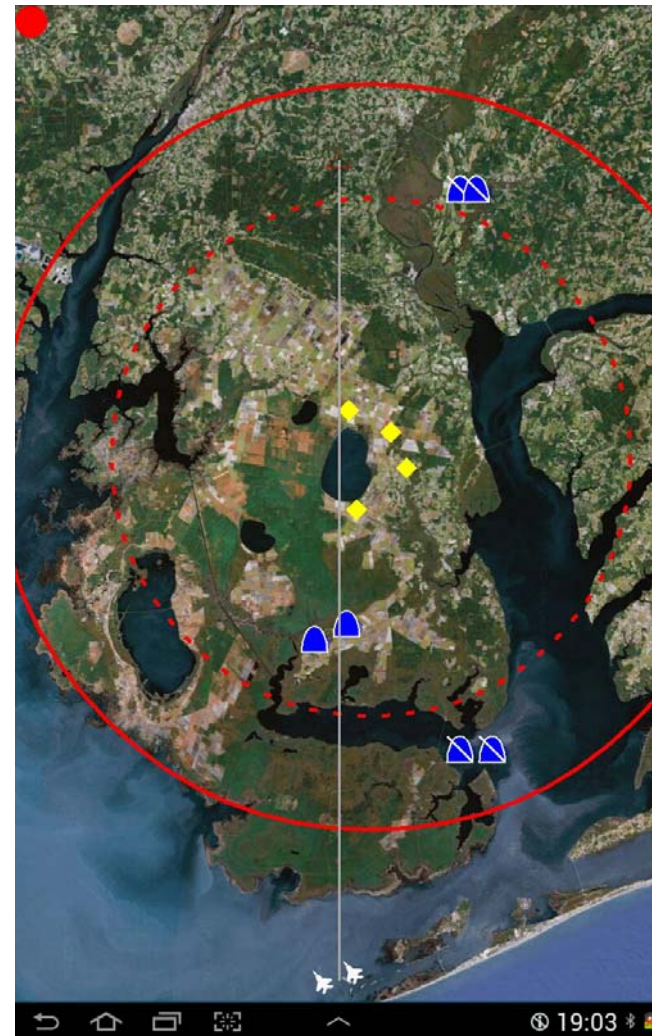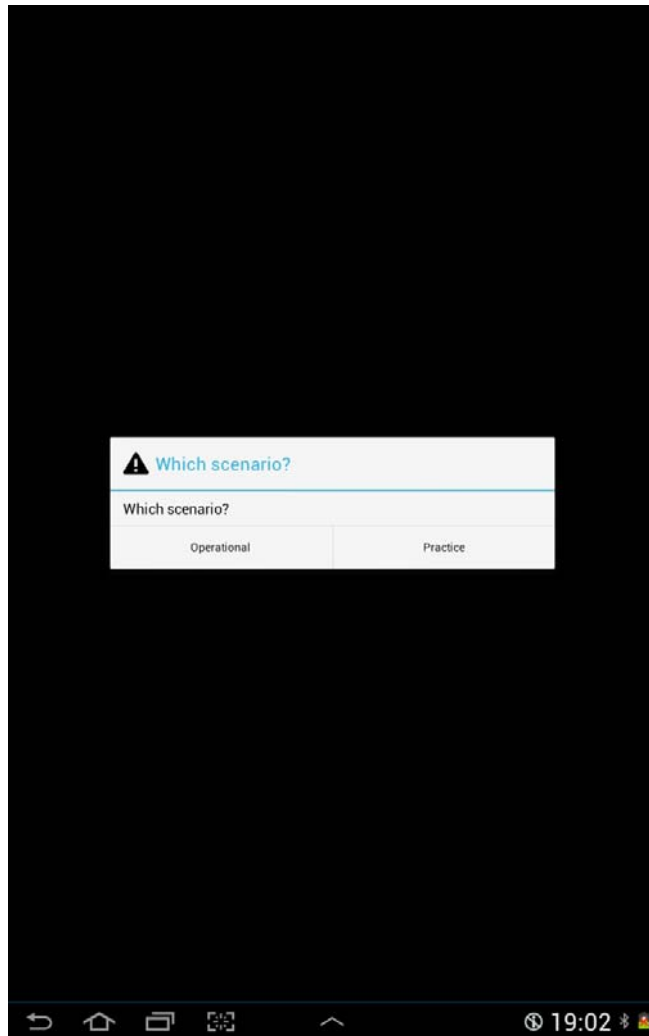
Figure 28.   a. SAWSI splash screen, selection of scenario   b. Initial view: four investigation sites(IS) (yellow diamonds)
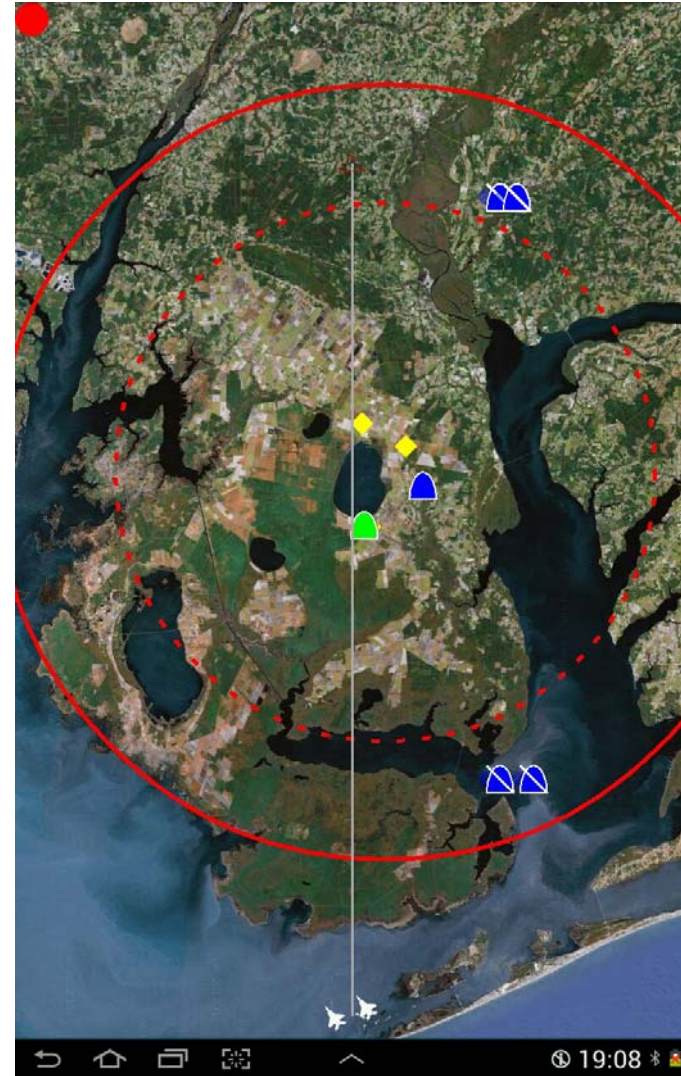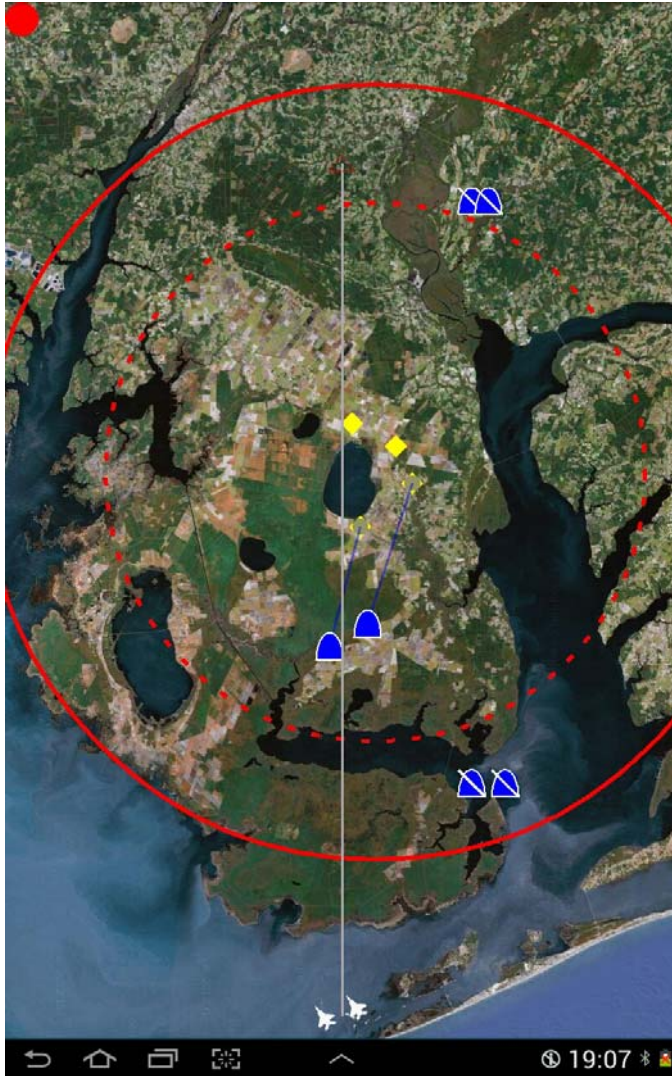
Figure 29.   a. Recon UAVs assigned to first two ISs   b. First recon UAV reaches IS and ready with imagery (green)
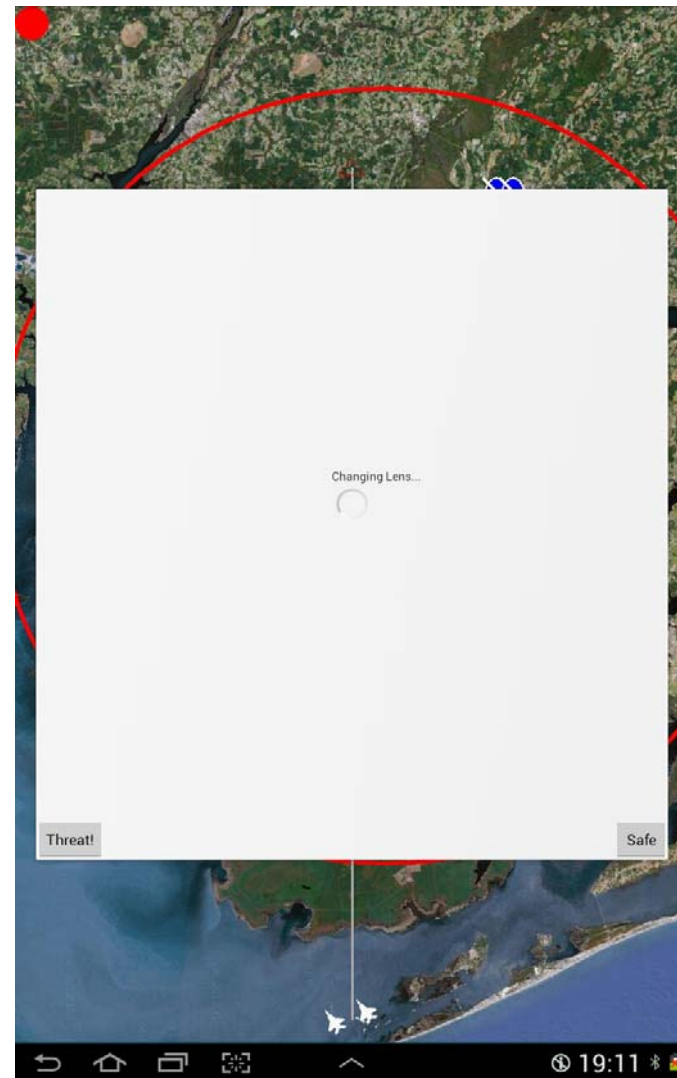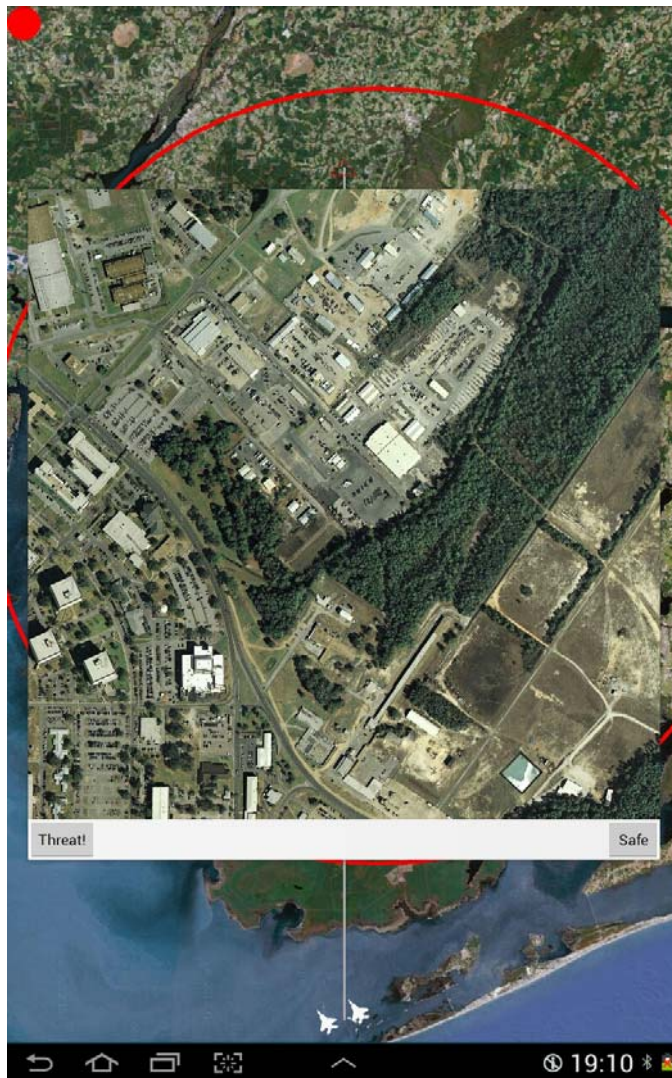
Figure 30.   a. Imagery from first IS, "Threat" or "Safe" options   b. User zooms imagery for closer inspection
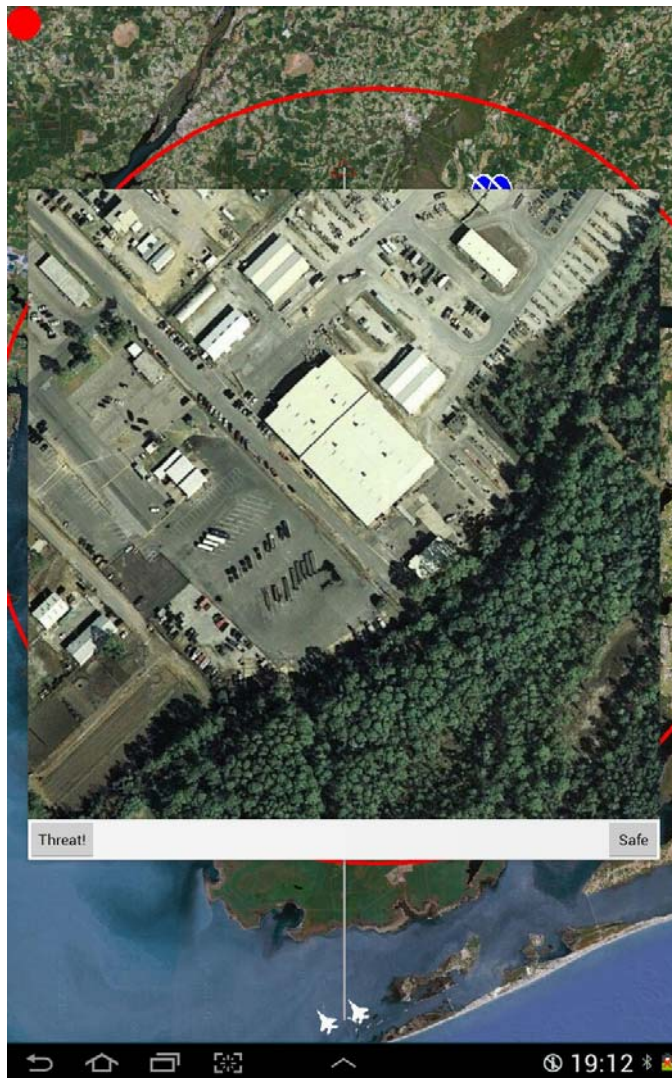
Figure 31.   a. Zoomed area   b. User selects safe, IS icon changed to safe (green circle)
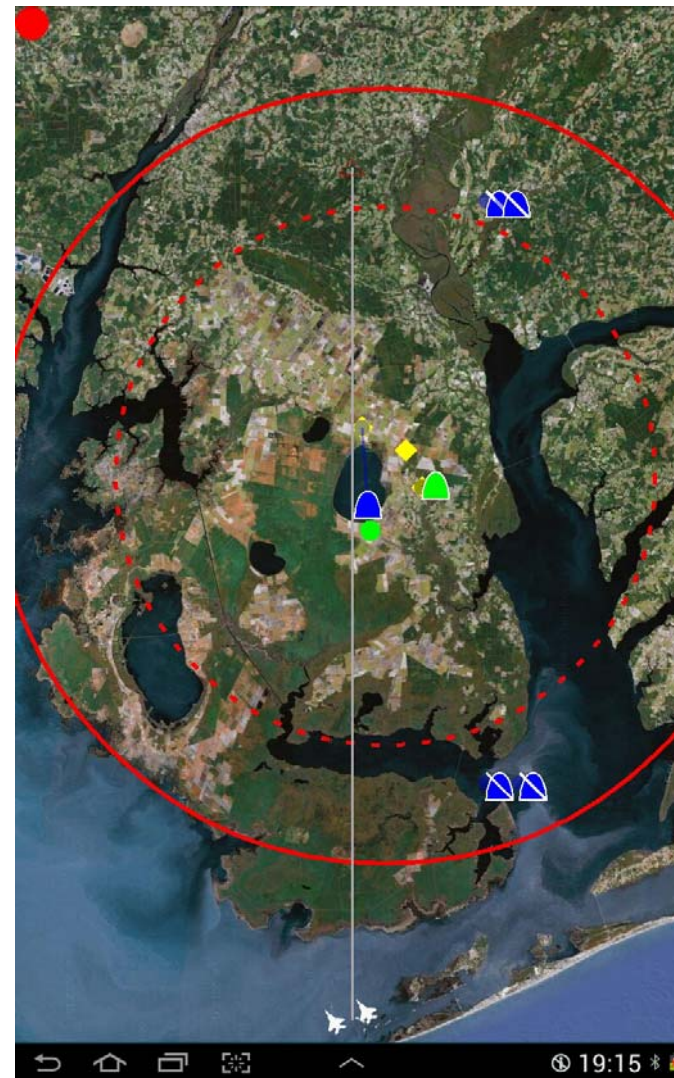
84
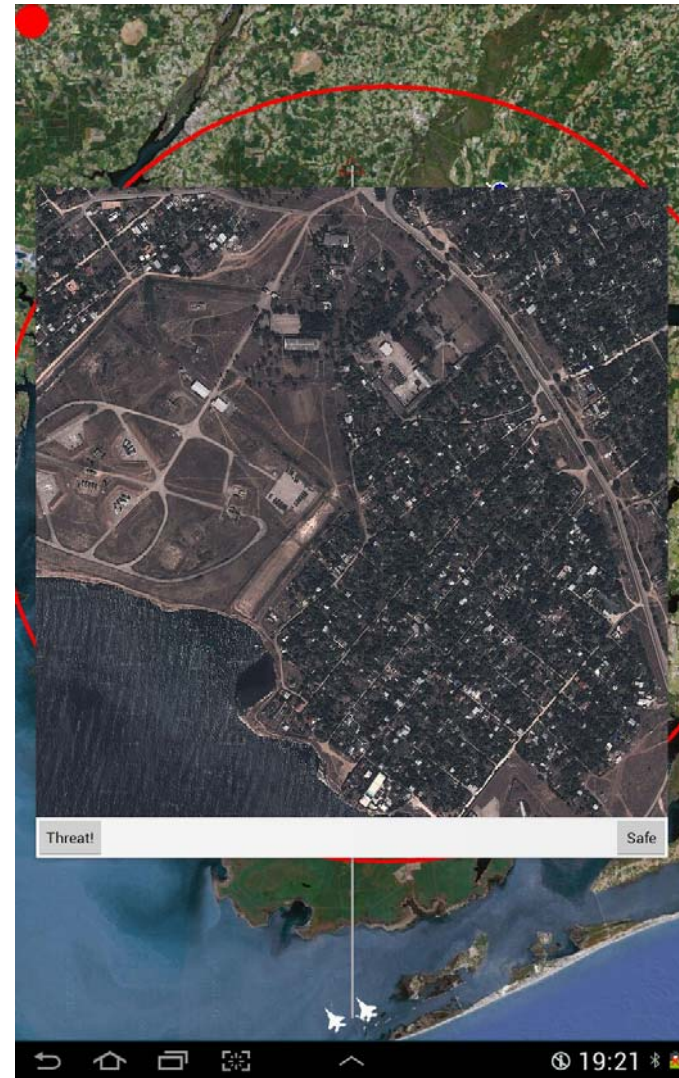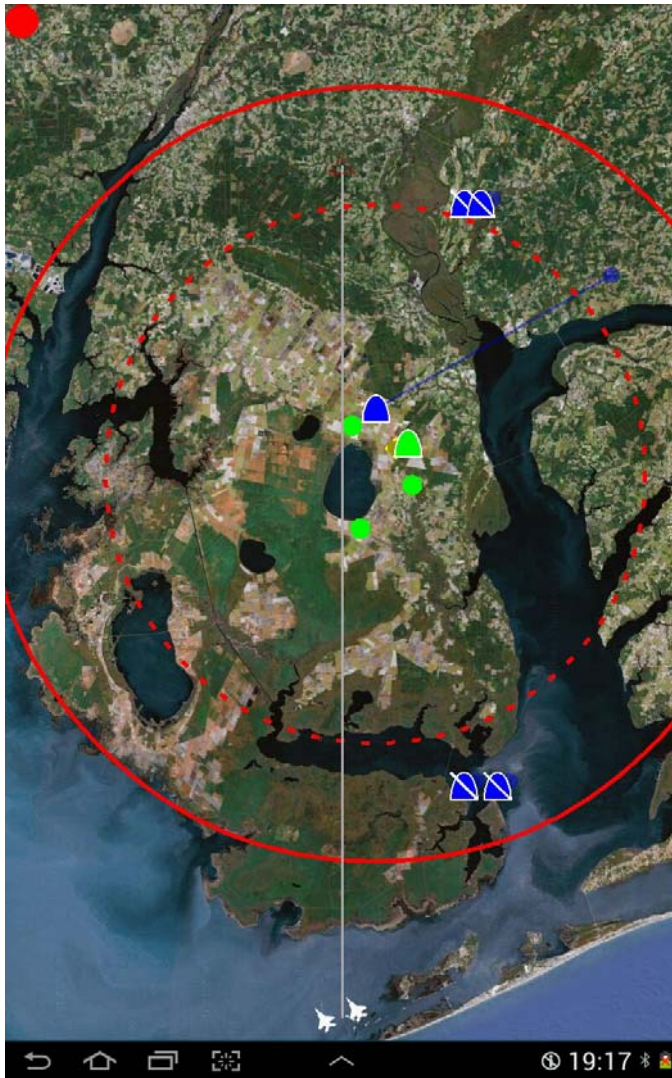
Figure 32.   a. Imagery ready for last IS   b. Imagery indicates possible threat (mid-left side)
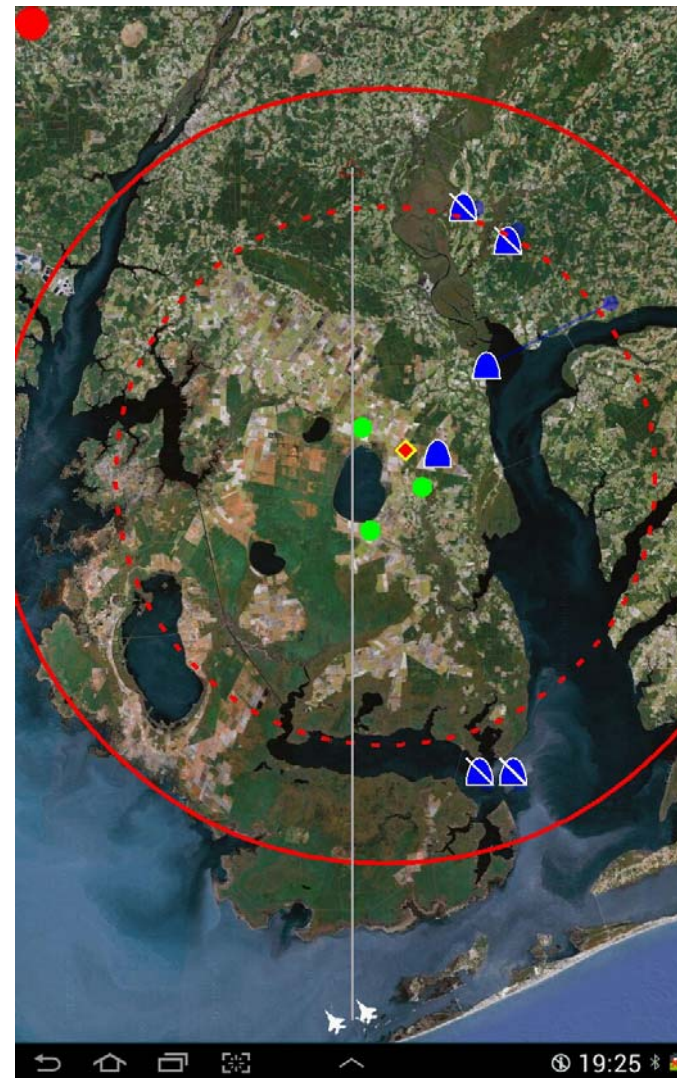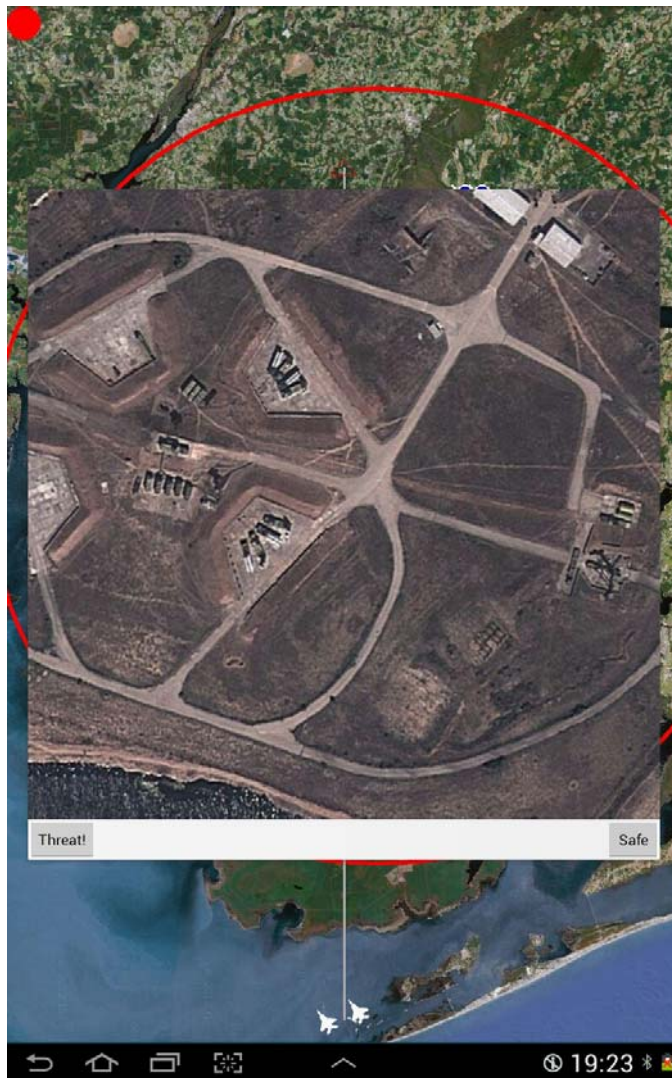
Figure 33.   a. Zoomed image shows two SA-XX TELs   b. User selects "Threat," icon changes to red/yellow
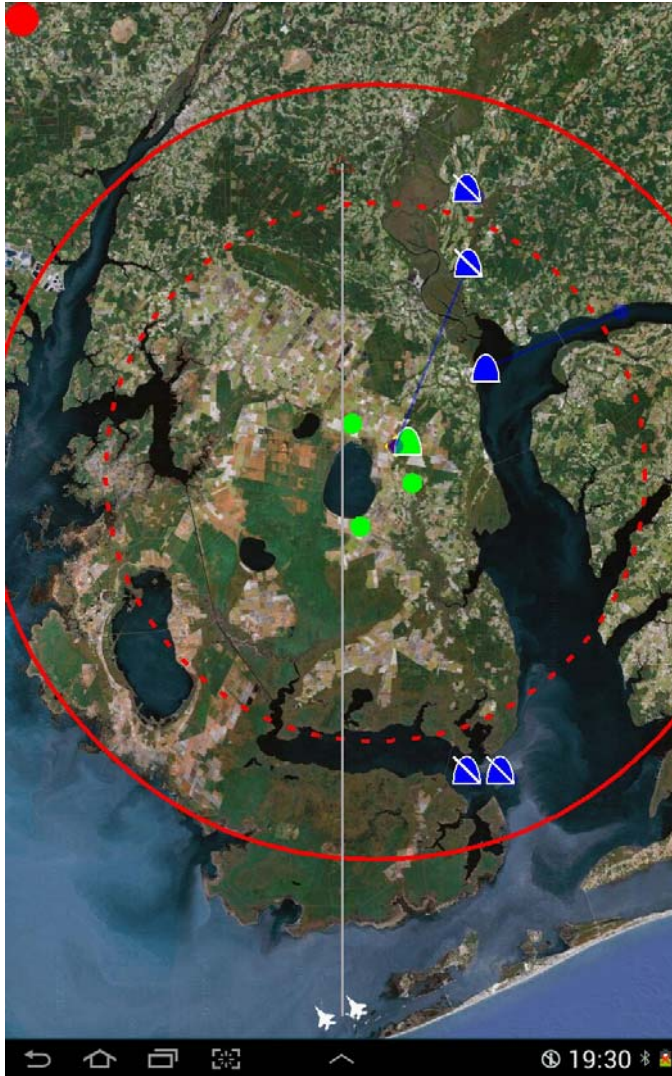
Figure 34.   a. Attack UAV assigned to strike   b. First attack UAV lost (shot down—grayed out)
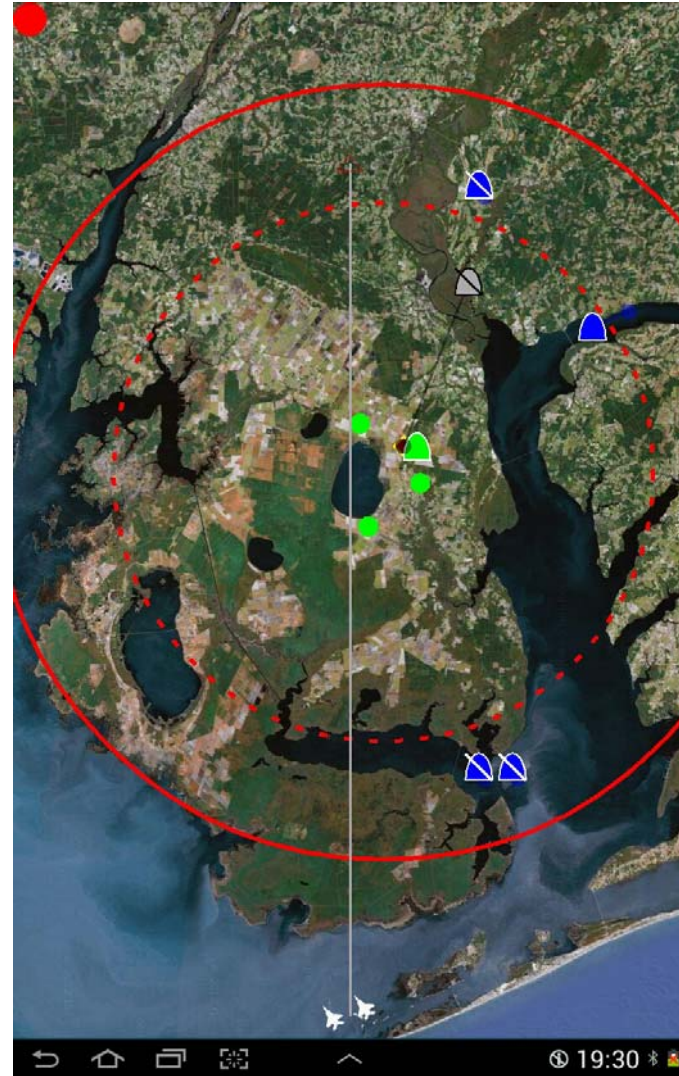
Figure 35.   a. Second attack UAV assigned   b. Second attack UAV lost
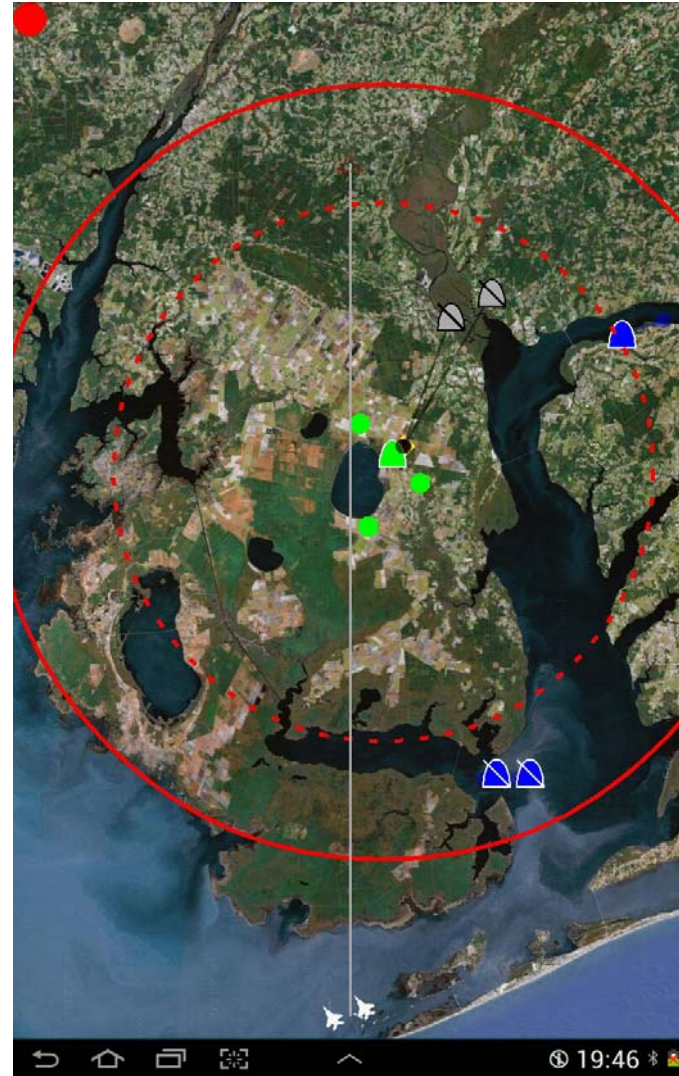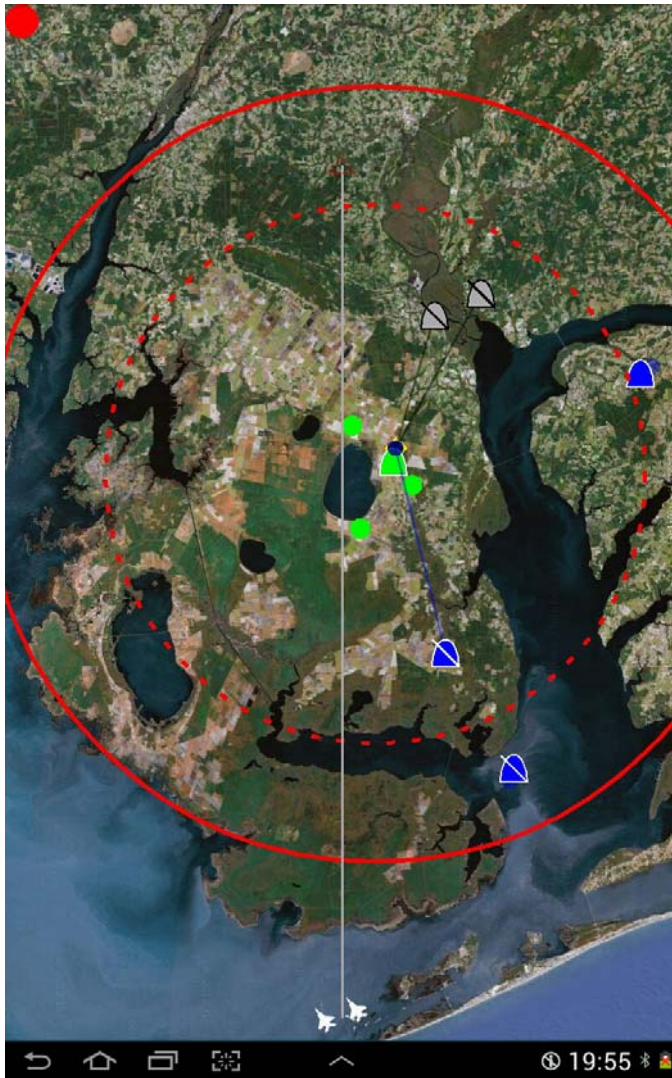
Figure 36.   a. Third attempt at targeting   b. First successful support to impact
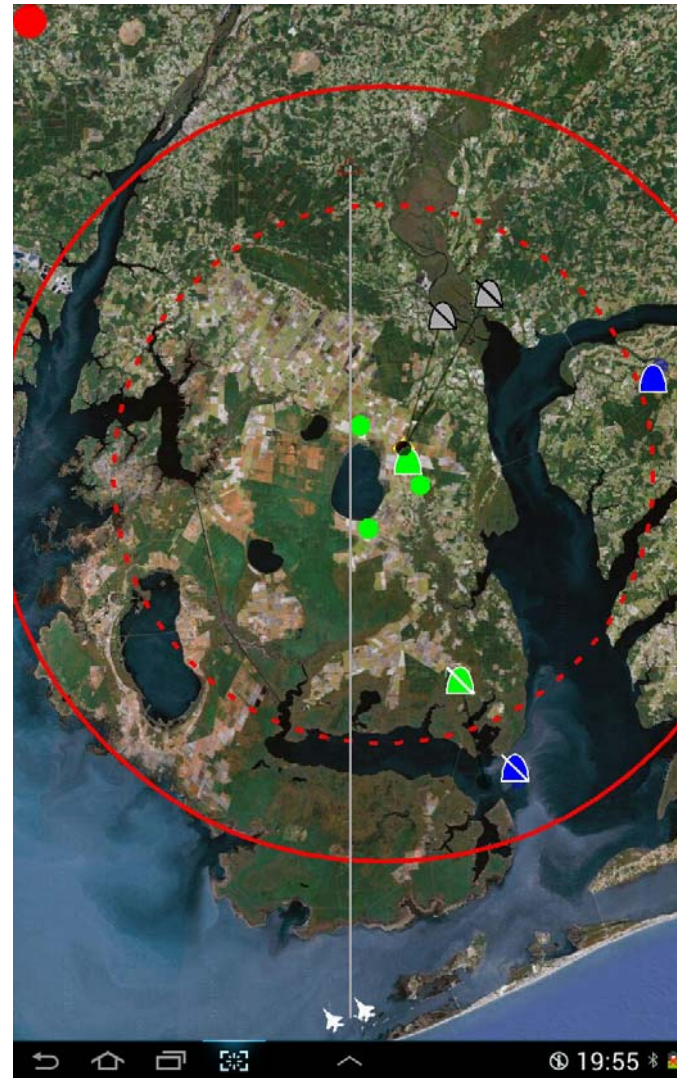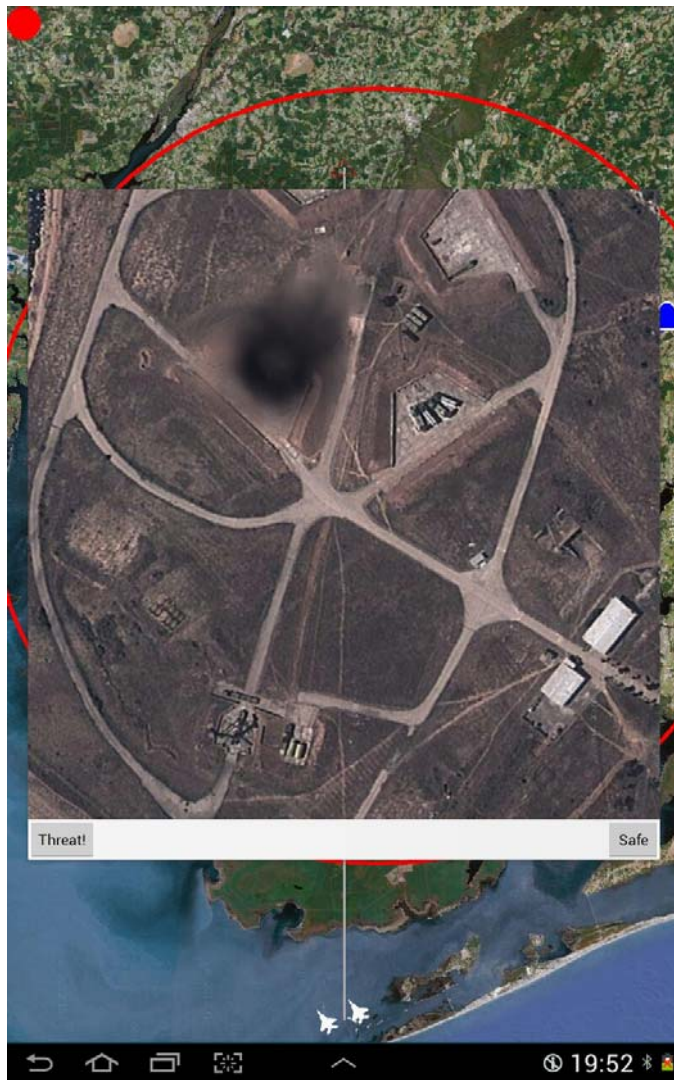
Figure 37.   a. BHA of one good hit, one TEL remaining   b. Final strike asset assigned to threat
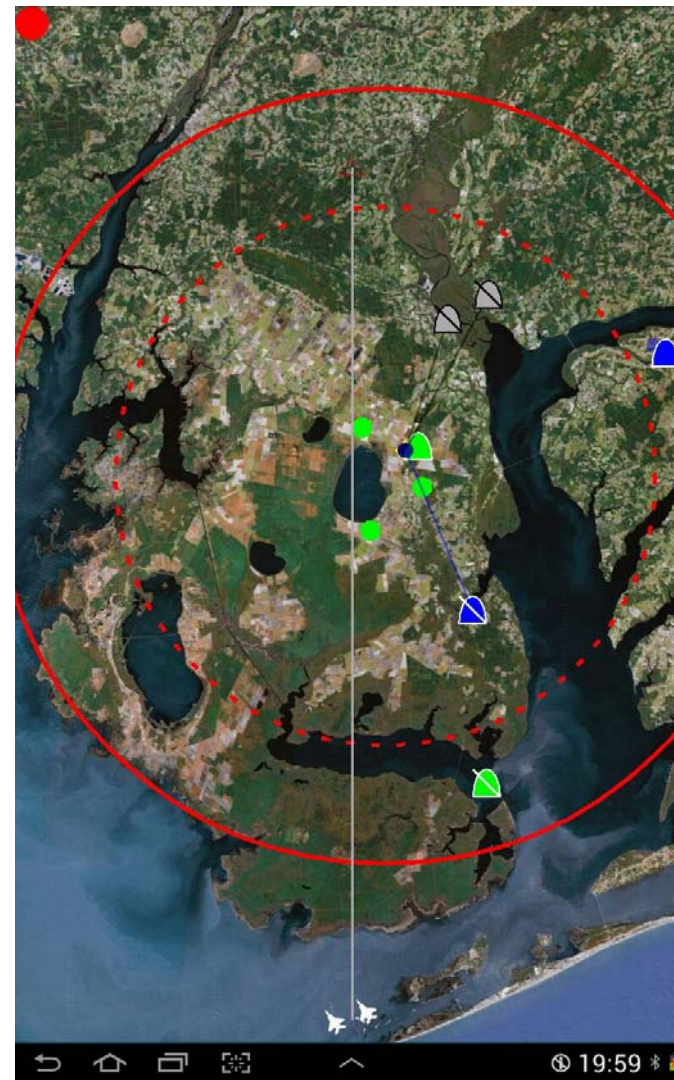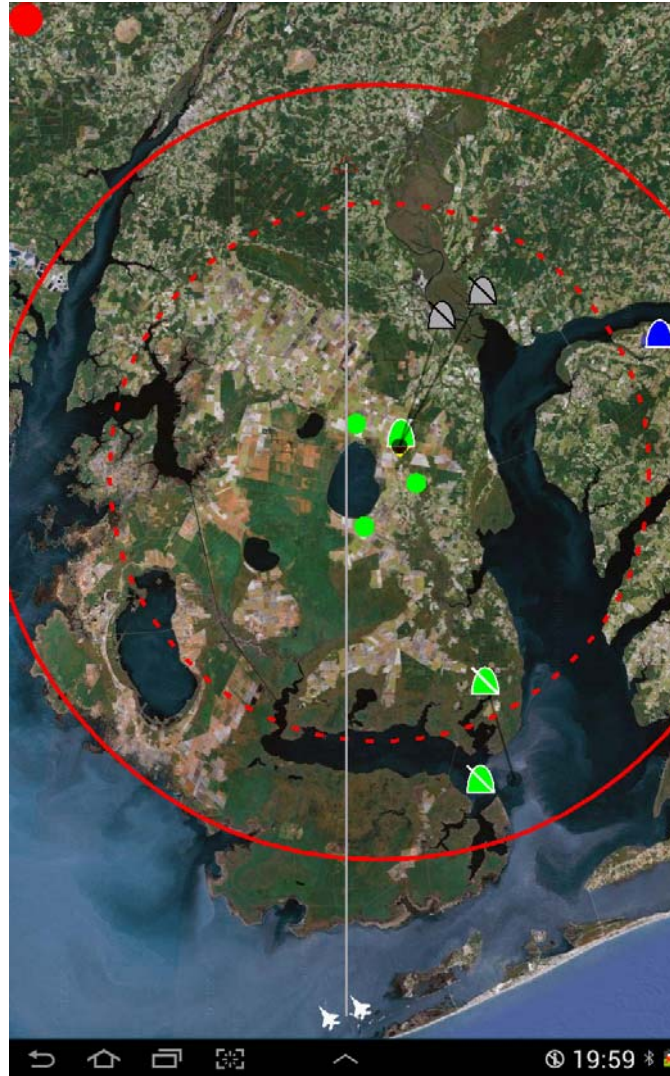
Figure 38.   Final support to impact, attack UAV returning to orbit, recon UAV ready with imagery

Figure 39.   a. Alternate technique: third and fourth attempt together ("Double Tapping") b. Successful support to impacts

Figure 40. a. BHA showing two good hits   b. Successfully cleared threat—threat rings removed

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX E. DATA TABLE

Table 7.   Data summary—Biographical and Experiment setup

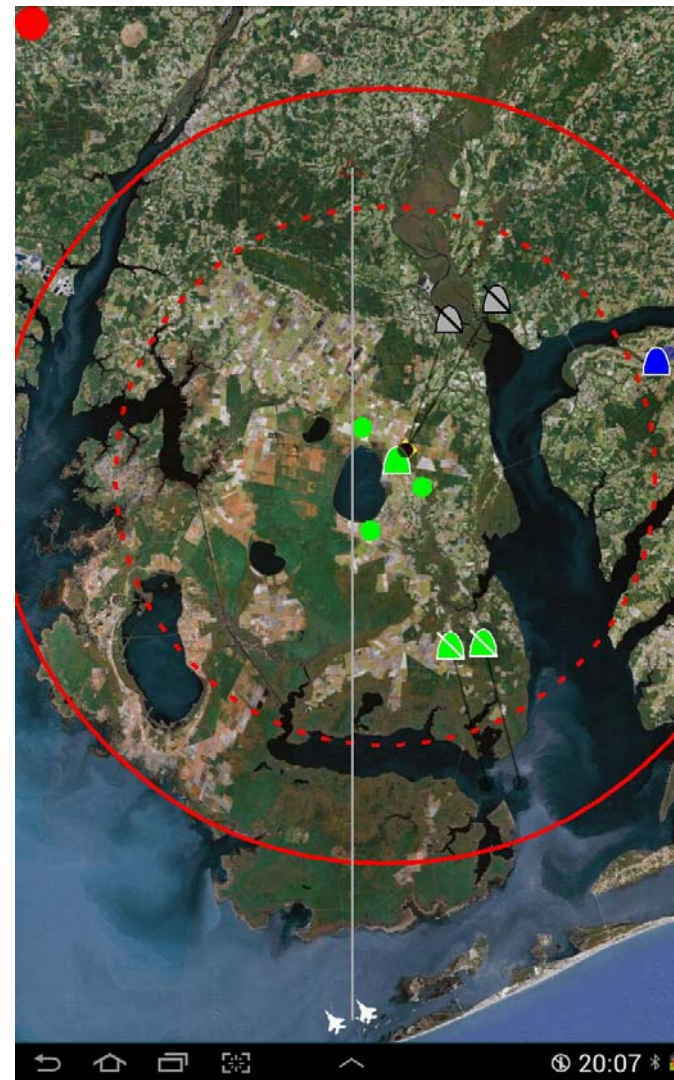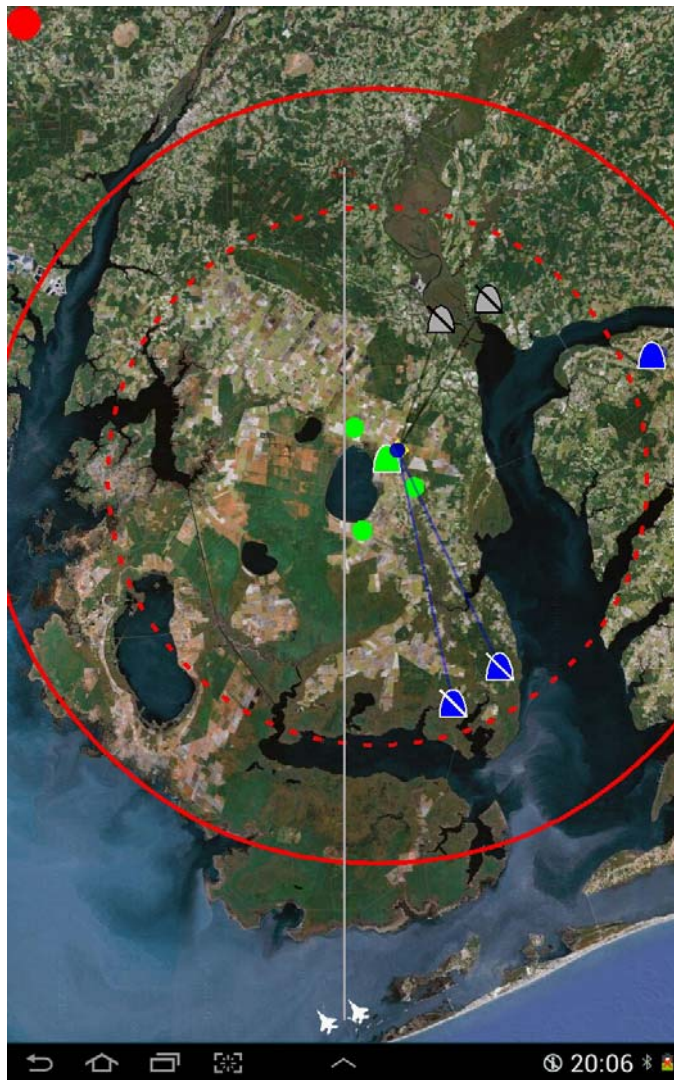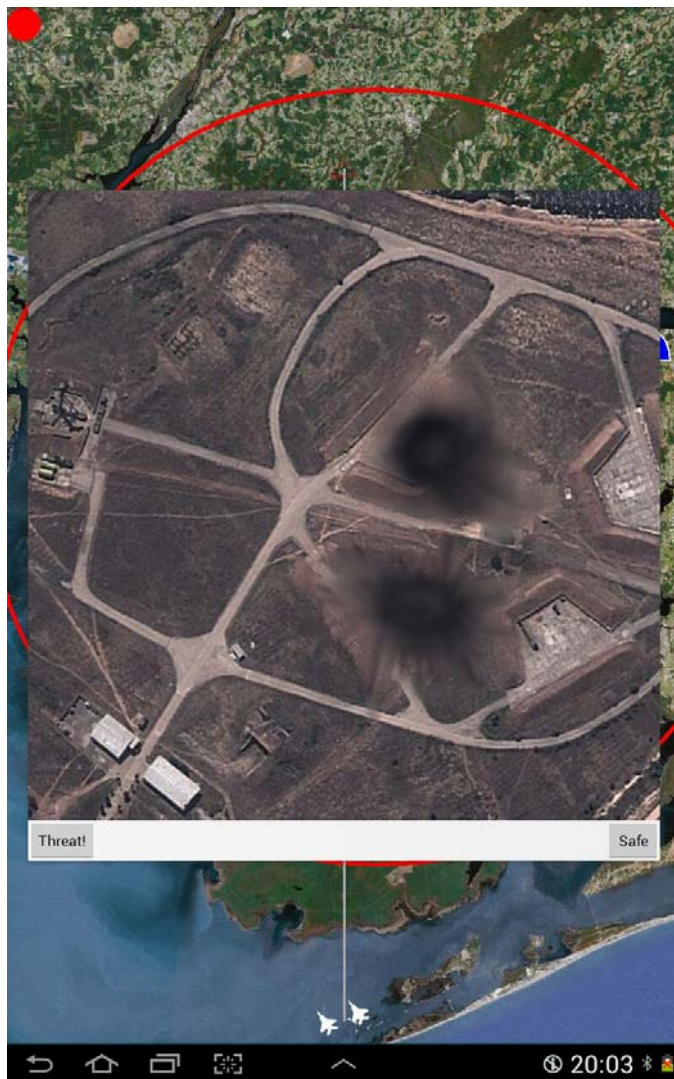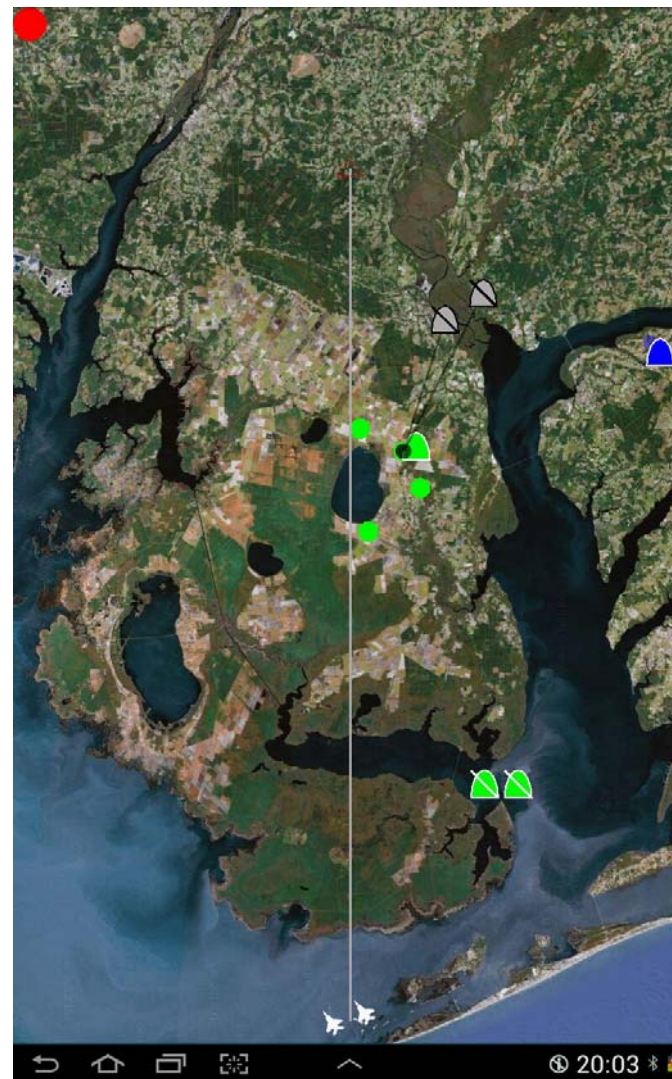| Partic- ipant #	Rank (O-_)	Age (Years)	Primary Series	Flight Hours					Days Since Last		F-18 Simulator Hours			Sleep Hours	Health (E/G/A)	Handed (L/R)	Own/Operate		Alert Test		Sim Type	Sim # (TOFT)	Bio xmit (R/B)	Task Order
				Total	F-18	Last 30	Last 7	Today	Form	Sim	Total	Last 30	Last 7				Games	Touch Screen	A (s)	B (s)				
1	3	32	F	1900	1550	10	0	0	14	14	300	2	1	8	E	R	N	N	17	28	F	132	B	1
2	3	28	F	1100	900	10	0	0	14	14	200	5	0	9	E	R	N	Y	14	24	F	132	R	1
3	4	34	F	2200	1800	8	2	0	7	7	400	4	2	8	E	R	N	Y	19	38	F	132	B	2
4	3	29	F	1250	950	1.1	0	0	90	30	150	2	0	7	E	L	N	N	16	31	F	132	R	2
5	3	34	F	800	600	10	0	0	14	14	150	2	0	8	E	R	N	Y	22	34	F	132	R	1
6	3	29	F	1000	700	3	0	0	14	14	200	12	0	7.5	E	L	Y	Y	15	26	F	132	B	2
7	4	34	F	2500	1500	5	1	0	90	21	100	1	0	7	E	R	N	Y	12	22	F	132	R	1
8	3	27	F	600	450	1.4	0	0	17	24	80	0	0	8	E	R	Y	Y	14	26	F	132	B	1
9	3	28	F	750	550	10	2	0	10	2	300	10	3	7.5	E	R	N	Y	17	44	F	132	R	1
10	3	27	F	1140	820	3	0	0	21	14	60	1	0	7.5	E	R	N	N	16	26	F	132	B	1
11	4	35	F	3200	2100	2	0	0	45	1	250	2	1	6	G	R	N	Y	19	30	F	132	R	2
12	4	37	D	2000	1700	15	0	0	8	10	500	2	0	8	E	L	N	Y	14	30	F	132	B	2
13	2	26	F	750	500	3	0	0	30	2	100	3	1	7	G	R	Y	Y	25	38	F	132	R	1
14	3	30	C	1200	750	0	0	0	60	2	250	5	1.5	8	E	L	N	Y	13	25	C	24	B	2
15	4	33	C	1900	1700	15	5	0	1	7	300	5	2	7	E	R	N	Y	10	21	C	24	R	1
16	3	27	C	1050	760	10	6	0	1	2	200	5	2	7	G	R	N	Y	19	27	C	24	B	1
17	3	31	F	1400	1100	8	2	0	30	90	200	0	0	6	G	R	N	Y	14	51	F	133	R	1
18	3	33	F	1600	1400	20	5	0	1	1	500	15	4	7.5	E	L	N	Y	11	41	F	133	B	1
19	3	35	F	1250	1000	10	3	0	30	30	100	0	0	8	E	R	Y	Y	13	26	F	133	R	1
20	3	31	F	1850	1600	6	0	0	14	60	200	0	0	7	G	R	Y	Y	20	47	F	133	B	2
21	3	31	E	1500	1250	15	5	0	7	7	300	5	0	6	G	R	Y	Y	13	24	F	133	B	2
22	3	31	E	2000	1400	5.5	0	0	14	40	200	0	0	7.5	E	R	N	Y	14	39	F	133	B	2
23	2	26	C	525	260	11	3	0	1	21	250	1.5	0	8	E	R	Y	Y	15	24	C	24	B	1
24	3	27	C	650	450	12	4.5	0	1	2	60	2	2	6	E	R	N	Y	8	26	C	24	R	2
25	4	36	C	2500	2350	15	2	0	1	6	150	1.5	0	6	E	R	N	Y	16	43	C	24	B	1
26	3	31	C	1100	800	10	2	0	1	7	100	3	0	7	E	L	N	Y	17	37	C	24	R	1
27	2	26	C	400	200	15	3	0	2	7	60	2	1	7	E	R	Y	Y	19	30	C	24	B	1
28	3	30	F	1500	1250	2.5	0.8	0	120	120	500	0	0	8	E	R	N	Y	12	20	F	133	R	2
29	4	33	E	2100	1800	3	0	0	14	21	300	2	0	7	E	R	N	Y	18	31	F	133	B	2
30	4	36	C	2000	1800	20	8	0	2	3	250	1	0	8	E	R	N	Y	12	22	F	133	R	2
31	4	39	E	2700	2500	15	2	0	17	47	200	0	0	8	A	R	Y	Y	21	38	F	133	B	2
32	3	31	F	2000	1600	10	5	0	90	16	100	5	0	6	G	R	Y	Y	13	25	F	133	B	1
33	3	27	F	750	500	20	5	1	1	0	30	2	1	6.5	G	R	Y	N	12	24	F	133	B	2
34	3	27	F	1000	400	15	5	0	4	14		1	0	7	E	R	N	Y	15	35	F	133	B	1
Mean	3.2	30.9		1475	1147	9.4	2.1	0.03	23	20	213	3.0	0.6	7.3					15	31				
SD	0.6	3.6		693	615	5.8	2.3	0.17	31	26	128	3.4	1.0	0.8					4	8				

= F-18C Simulator Used

(Best E/F)
(Second Best E/F)
Best C
Second Best C

Second Worst
Worst

95

Table 8.   Data summary—Primary task and Traditional secondary task performance

Partic-ipant #	Primary Task (Following Distance)												Traditional Task Complete (Y/N)
	Raw Data						Filtered Data						
	Base μ	Trad μ	Fut μ	Base SD	Trad SD	Fut SD	Base μ	Trad μ	Fut μ	Base SD	Trad SD	Fut SD	
1	384	415	803	198	206	558	255	415	803	85	206	558	Y
2	328	550	414	179	482	564	173	308	246	67	124	112	N
3	324	325	334	199	140	158	178	249	256	75	94	81	Y
4	304	338	272	188	152	155	163	285	160	65	111	56	Y
5	307	311	320	166	169	157	202	192	290	83	73	133	N
6	318	303	323	175	131	160	170	160	225	56	52	61	Y
7	308	357	425	113	121	114	237	295	372	41	68	85	Y
8	273	272	277	139	138	134	149	191	186	37	62	52	Y
9	302	341	313	160	128	157	169	253	250	65	63	89	Y
10	325	292	296	167	133	129	149	183	205	55	51	63	N
11	268	265	262	185	128	131	145	211	199	73	96	63	N
12	229	181	257	149	127	88	152	181	218	85	88	91	Y
13	259	319	275	153	165	124	183	230	230	100	105	88	Y
14	110	143	130	23	69	165	111	121	129	23	43	70	Y
15	112	337	668	150	209	477	159	337	657	150	209	474	Y
16	336	316	418	401	120	322	336	316	420	401	120	325	Y
17	276	322	403	137	118	133	192	274	371	75	81	109	Y
18	224	579	342	131	564	140	164	340	312	74	106	107	N
19	161	291	246	106	150	135	146	220	246	73	103	71	N
20	173	289	310	122	157	135	142	222	240	67	85	101	Y
21	156	238	298	119	142	138	121	147	210	54	65	70	Y
22	268	270	300	772	125	135	140	198	227	42	85	81	Y
23	228	270	368	98	113	109	228	270	370	98	113	109	Y
24	268	560	755	130	550	716	268	472	762	130	337	559	N
25	128	125	179	25	28	82	128	125	180	25	28	81	Y
26	122	261	156	27	122	47	122	261	154	27	122	43	Y
27	102	119	166	37	37	76	102	119	167	37	37	77	Y
28	218	307	371	102	119	128	176	258	312	52	97	73	Y
29	172	308	344	117	186	167	114	226	175	46	103	88	Y
30	300	1064	768	117	487	367	300	1064	749	117	367	472	N
31	146	291	262	74	162	186	137	173	131	53	100	66	Y
32	181	276	304	138	161	184	136	171	163	57	70	88	Y
33	209	303	272	206	145	171	131	210	187	47	80	46	Y
34	176	304	262	137	150	108	110	240	220	46	120	72	N
Mean	235	331	350	157	180	198	170	262	295	76	108	139	
SD	78	165	165	128	132	156	55	162	181	64	73	148	

(Best E/F)
(Second Best E/F)
Best C
Second Best C

Second Worst
Worst

96

Table 9. Data Summary—Futuristic secondary task peformance

Partic-ipant #	SAWCI Data						Notes:
	Time to: (sec.)						
	1st Interact	Indentify	Damage	Destroy	(1st Dest. View)	Cleared	
1	4.3	140.3	213.5	224.8		252.1	user double-tapped threat.
2	3.9	106.4	176.2	216.0		231.3	
3	40.7	167.4	224.8	287.7	290.6		user viewed destroyed threat @ 4:51 but did not mark clear before end of session
4	3.0	80.7	119.0	131.1		143.6	user double-tapped threat.
5	-27.0	242.2	273.8	277.1	288.2		user double-tapped threat. User misidentified as still a threat @ 4:51, but then correctly after session
6	-19.4	110.4	179.8	185.8		235.2	user double-tapped threat.
7	-24.2	80.9	131.1	138.3		150.3	user double-tapped threat.
8	-26.8	78.1	112.8	156.0		252.6	user reported threat clear @ 3:41, but hadn't actualy set on system
9	-14.2	95.2	167.2	202.2		232.6	
10	11.4	113.4	177.3	190.6		203.2	user double-tapped threat.
11	-27.4	198.0	289.3				
12	-19.8	145.1	241.5	251.5			user double-tapped threat. last interaction after session end
13	-24.0	106.3	161.9	166.6		240.5	user double-tapped threat.
14	-24.5	102.2	160.8	226.0		266.9	
15	-25.1	149.8	208.9	283.5	315.1		user viewed destroyed threat @ 5:15 elapsed but did not mark clear before end of session
16	-3.1	111.4	146.4				User fired on wrong investigation point first- then misidentified correct threat as destroyed after only
17	-8.5	125.8	167.8	186.0		205.7	user double-tapped threat.
18	-42.6	118.8	186.1	235.3		276.3	
19	20.9	163.8	254.8				user had just re-idetified as still a threat @ end of session
20	-19.2	139.9	170.3	192.4	239.2		user double-tapped threat. User misidentified as still a threat @ 3:59
21	-14.6	167.0	230.3	291.7			user did not task ISR UAV to verify threat clear
22	-16.8	156.2	212.5	268.1	301.8		user viewed destroyed threat @ 5:02 but did not mark clear before end of session
23	-19.2	126.7	none				user had second UAV destroyed @ end of session. Stated confusion on CDR's intent w/ loss of UAVs
24	-8.8	137.6	254.1	260.2	300.8		user double-tapped threat. User misidentified as still a threat @ 5:01, but then correctly after session
25	-44.4	141.9	215.1	248.1	327.4		user double-tapped threat. User misidentified as still a threat @ 5:27
26	-8.4	159.1	230.3	308.2			
27	-11.6	191.3	294.4				user had just re-idetified as still a threat @ end of session
28	-30.5	138.4	187.2	204.2		311.1	
29	-17.7	116.9	183.1	184.4	220.7		user double-tapped threat. User misidentified as still a threat @ 3:41
30	-49.9	220.0	288.0				last user interaction was opening the view of the damaged target
31	-55.4	121.1	172.9	176.6		193.3	user double-tapped threat.
32	-20.2	110.8	174.6	183.5		220.2	user double-tapped threat.
33	-19.1	139.5	207.7	249.9		262.1	
34	-5.6	115.0	166.4	167.2		235.5	user double-tapped threat.
Mean	-14.4	135.8	199.4	217.6	285.5	230.2	
SD	20.4	37.8	47.8	49.0	36.9	42.5	

Control:best possible times		66.9	108.4	117.3		123.8	With double-tap
Control:best possible times		66.9	108.4	131.6		139.3	Without double-tap (sequential)
Control: "reasonable speed		81.4	131.8	136.2		144.8	Undistracted, double-tapping

(Best E/F)
(Second Best E/F)
Best C
Second Best C

Second Worst
Worst

THIS PAGE INTENTIONALLY LEFT BLANK
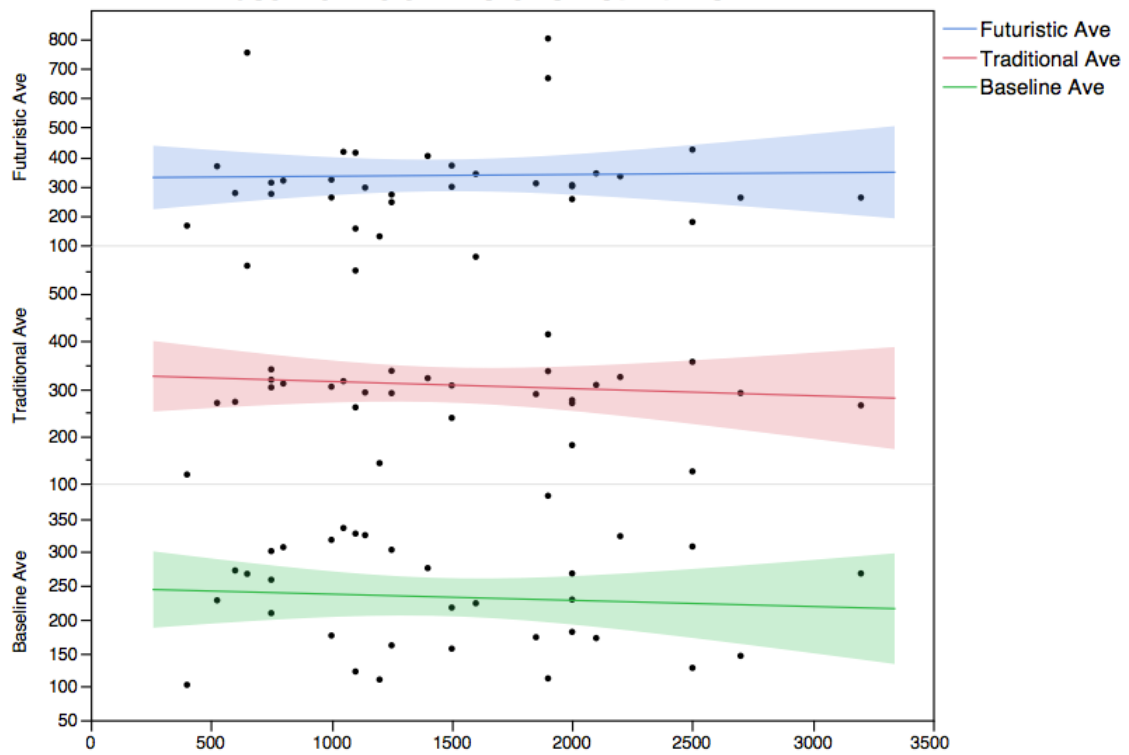
APPENDIX F. RAW TIME VS. DISTANCE PLOT

Figure 41. Total flight hours vs raw mean following distance

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

Anderson, R.E. (1992). Social impacts of computing: Codes of professional ethics. *Social Science Computing Review, 2*, 453–469.

Burts, L., Martins, K., Quimby, A., & Shattuck, L. (1998). Driving and cellular telephones: an analysis of dual task performance in the real world. *Proceedings of the Applied Behavioral Sciences Symposium 22–23*, 156–160.

Endsley M.R. (1995). Model of Situational Awareness Toward a Theory of Situation Awareness in Dynamic Systems [Image from Article]. *Human Factors,* 37(1), 32–64. Retrieved from http://www.nap.edu/openbook.php?record_id=6173&page=177

Endsley, Mica R. (1999). *Situation Awareness in Aviation Systems. Handbook of Aviation Human Factors.* (pp. 257–276). Mahwah NJ: Lawrence Erlbaum Associates.

General Electric I-Sim. (2006). PatrolSim GE-ISIM driving simulator [Photograph]. Retrieved from http://www.distraction.gov/download/research-pdf/Comparison-of-CellPhone-Driver-Drunk-Driver.pdf

Hitchcock, Lloyd. (1999). *Pilot Performance. Handbook of Aviation Human Factors.* (pp. 311–326). Mahwah, NJ: Lawrence Erlbaum Associates.

Massachusetts Institute of Technology (MIT). (2011). Research Environment for Supervisory Control of Heterogeneous Unmanned Vehicles [Software Screen Capture]. Retrieved from http://www.thinkmind.org/index.php?view=article&articleid=cognitive_2011_6_10_40057

Mulligan, N. W. (2002). Attention and perceptual implicit memory: effects of selective versus divided attention and number of visual objects. *Psychological research, 66*(3)*,* 157–65. doi:10.1007/s00426–002–0089–2

Rantanen, E.M., & Goldberg, J.H. (1999). The effect of mental workload on the visual field size and shape. *Ergonomics, 42*, 816–834.

Ratwani, R., & Trafton, J. G. (2010). An eye movement analysis of the effect of interruption modality on primary task resumption. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 52*(3)*,* 370–380. doi:10.1177/0018720810374195

Rodas, O.R., Veronda, M., & Szatkowski C. (2011). Developing a Decision Tool to Evaluate Unmanned System's Command and Control Technologies in Network Centric Operations Environments. *Cognitive 2011: The Third International Conference on Advanced Cognitive Technologies and Applications.*

Steelman, K. S., McCarley, J. S., & Wickens, C. D. (2011). Modeling the control of attention in visual workspaces. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 53*(2)*,* 142–153. doi:10.1177/0018720811404026

Strayer, D. L., Drews, F.A, & Crouch, D. J. (2006). A comparison of the cell phone driver and the drunk driver. *Human Factors, 48(2),* 381–91. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16884056

Thomas, L.C., & Wickens, C.D. (2001). Visual displays and cognitive tunneling: frames of reference effects on spatial judgments and change detection. *Proceedings of the 45th Annual Meeting of the Human Factors and Ergonomics Society,* 1–5.

Tombaugh, T.N.T.N. (2004). "Trail Making test A and B: Normative Data Stratified by Age and Education." *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists, 19*(2*),* 203–214. http://dx.doi.org/10.1016/S0887–6177(03)00039–8

Troyer, A K., & Craik, F. I. (2000). The effect of divided attention on memory for items and their context. *Canadian Journal of Experimental Psychology = Revue canadienne de psychologie expérimentale, 54*(3)*,* 161–71. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11021036

U.S. Navy Archive. (2012). F/A-18 Hornet Tactical Operational Flight Trainer (TOFT) [Photograph]. Retrieved from http://www.navair.navy.mil/index.cfm?fuseaction=home.NAVAIRNewsStory&id=4988

Wickens, C. D. (2008). Multiple resources and mental workload. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 50(3),* 449–455. doi:10.1518/001872008X288394

Wickens, C.D. (2003). Control of Multiple-UAVs: A Workload Analysis. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 47*(3)*,* 479–487. doi: 10.1518/001872005774860005

Wickens, C.D. (2005). Attentional Tunneling and Task Management, Presented at the 13th International Symposium on Aviation Psychology, Dayton OH, 2005. http://130.203.133.150/viewdoc/download?doi=10.1.1.135.283&rep=rep1&type=pdf

Wickins, C.D. (2002). Multiple Resources and Performance Prediction. *Theoretical Issues in Ergonomics Science, 3*(2), 159–177.

THIS PAGE INTENTIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1.   Defense Technical Information Center
        Ft. Belvoir, Virginia

2.   Dudley Knox Library
        Naval Postgraduate School
        Monterey, California

3.   LCDR Eric McMullen
        MOVES Institute
        Monterey, California

4.   Naval Crew Resource Management Program Office
        Naval Aviation Schools Command
        Pensacola, Florida