Assessing the Effects of Multi-Modal Communications on Mental Workload During the Supervision of Multiple Unmanned Aerial Vehicles

Sharon Claxton Bommer
Wright State University

Follow this and additional works at: https://corescholar.libraries.wright.edu/etd_all
Part of the Operations Research, Systems Engineering and Industrial Engineering Commons

Repository Citation
Bommer, Sharon Claxton, "Assessing the Effects of Multi-Modal Communications on Mental Workload During the Supervision of Multiple Unmanned Aerial Vehicles" (2013). Browse all Theses and Dissertations. 1349.
https://corescholar.libraries.wright.edu/etd_all/1349

This Thesis is brought to you for free and open access by the Theses and Dissertations at CORE Scholar. It has been accepted for inclusion in Browse all Theses and Dissertations by an authorized administrator of CORE Scholar. For more information, please contact corescholar@www.libraries.wright.edu, library-corescholar@wright.edu.
ASSESSING THE EFFECTS OF MULTI-MODAL COMMUNICATIONS ON MENTAL WORKLOAD DURING THE SUPERVISION OF MULTIPLE UNMANNED AERIAL VEHICLES

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Engineering

By

Sharon Claxton Bommer
B. S., Tennessee State University, 1996
M.B.A., University of Cincinnati, 2008

2013
Wright State University
WRIGHT STATE UNIVERSITY
GRADUATE SCHOOL

Date: 12/19/2013

I HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER MY SUPERVISION BY Sharon Claxton Bommer ENTITLED Assessing the Effects of Multi-Modal Communications on Mental Workload During the Supervision of Multiple Unmanned Aerial Vehicles BE ACCEPTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF Master of Science in Engineering.

______________________________
Mary Fendley, Ph.D.
Thesis Director

______________________________
Thomas N. Hangartner, PhD, FAAPM
Department Chair

Committee on Final Examination

______________________________
Mary Fendley, Ph.D.

______________________________
Frank Ciarallo, Ph.D.

______________________________
Subhashini Ganapathy, Ph.D.

______________________________
R. William Ayres, Ph.D.
Interim Dean, Graduate School

"Distribution A: Approved for public release; distribution unlimited. 88ABW Cleared 12/03/2013; 88ABW-2013-5056."
ABSTRACT

Bommer, Sharon Claxton. M.S. Engineering, Department of Biomedical, Industrial, and Human Factors Engineering, Wright State University, 2013. Assessing the Effects of Multi-Modal Communications on Mental Workload during Supervision of Multiple Unmanned Aerial Vehicles.

Human supervisory control (HSC) systems that require intermittent human input of commands to create actions within the system are helpful in reducing the human’s mental workload (MWL). HSC systems are commonly employed in the piloting and control of unmanned aerial vehicles (UAVs). As the ratio of UAVs to operators increases there is a clear need to understand the effort required to complete task requirements. Specifically, this research explored MWL relative to multi-modal communications using subjective and physiological measures. To accomplish this, a cognitive model was developed to understand the relationship between tasks and MWL. MWL was manipulated by measuring three in-flight variables: increased communication tasks, task interruptions, and a combination of these two variables. Results showed that task interruptions have the greatest influence on MWL suggesting that reducing the tasks complexity rather than number of UAVs, would be more effective in reducing operator workload.
# Table of Contents

1.0 INTRODUCTION ................................................................. 1  
1.1 Overview and Problem Description ........................................ 1  
1.2 Research Objectives .......................................................... 2  
1.3 Research Hypothesis ......................................................... 4  

2.0 RELATED RESEARCH .......................................................... 6  
2.1 Supervisory Control ............................................................ 6  
2.2 Hierarchical Task Analysis ................................................... 8  
2.3 Modes of Communication .................................................... 9  
2.4 Comms Usage Diagram ....................................................... 10  
2.5 Multiple Resource Theory (MRT) .......................................... 11  
2.6 Information Processing ..................................................... 13  
2.7 Mental Workload Assessment .............................................. 14  
2.7.1 Physiological Measures .................................................. 17  
2.7.2 Subjective Workload Techniques ...................................... 18  
2.8 Crew Awareness Rating Scale .............................................. 19  
2.9 NASA-TLX ........................................................................ 22  

3.0 RESEARCH COMPONENTS .................................................... 24  
3.1 Cognitive Model ................................................................. 24  
3.2 Vigilant Spirit Control Station .............................................. 25  
3.3 Eye Tracking ....................................................................... 26  
3.4 Mobile Test Environment (MTE) .......................................... 28  
3.5 Research Framework ........................................................... 31  

4.0 RESEARCH METHODOLOGY ............................................... 34  
4.1 Phase 1: Preliminary Design Assessment ............................... 34  

"Distribution A: Approved for public release; distribution unlimited. 88ABW Cleared 12/03/2013; 88ABW-2013-5056."
4.2 Phase 2: Experimental Design ................................................................. 36
  4.2.1 Test Scenario ...................................................................................... 36
  4.2.2 Test Objective ..................................................................................... 38
  4.2.3 Components of the Experiment .......................................................... 39
  4.2.4 Operating Procedure .......................................................................... 40
  4.2.5 Results .................................................................................................. 41
  4.2.6 Discussions ........................................................................................ 47

5.0 CONCLUSION ............................................................................................. 50

6.0 FUTURE WORK ........................................................................................ 53

7.0 APPENDIX ................................................................................................ 54
  7.1 WORKLOAD ASSESSMENT TOOLS ..................................................... 54
  7.2 CREW AWARENESS RATING SCALE ................................................ 55
  7.3 NASA-TLX SCALE ............................................................................... 57
  7.3 PHASE 2-DATA COLLECTION ............................................................ 58
    7.3.1 CARS Data Set for n=2 for the Increased Communication with Task Interruption Trials ........................................................................ 58
    7.3.2 CARS Data Set for n=5 for the Baseline Trials ................................ 58
    7.3.3 CARS Data Set for n=5 for the Increased Communication Trials ............. 59
    7.3.4 CARS Data Set for n=5 for the Task Interruptions Trials ...................... 59
    7.3.5 CARS Data Set for n=5 for the Increased Communications with Task Interruptions Trials ................................................................. 60
    7.3.6 NASA-TLX Data Set for n=2 ............................................................ 60
    7.3.7 NASA-TLX Data Set for n=5 ............................................................ 61
    7.3.8 Pupil Diameter Data for the Baseline Trials ........................................ 62
    7.3.9 Pupil Diameter Data for the Increased Communication Trials ............. 64
    7.3.10 Pupil Diameter Data for the Task Interruption Trials ....................... 66

"Distribution A: Approved for public release; distribution unlimited. 88ABW Cleared 12/03/2013; 88ABW-2013-5056."
LIST OF FIGURES

Figure 1.0   UAV Manning for this Experimental Scenario ..........................3
Figure 2.0   Yerkes-Dodson Curve ................................................................16
Figure 3.0   CARS Rating Scale ..................................................................19
Figure 4.0   Hybrid Cognitive Model .........................................................25
Figure 5.0   Image of VSCS Screens .............................................................26
Figure 6.0   VSCS Station with Eye Tracking ...............................................28
Figure 7.0   Vigilant Spirit Trailer ...............................................................29
Figure 8.0   Test Administrator Stations ....................................................30
Figure 9.0   Simulation Station .................................................................31
Figure 10.0  Research Framework .............................................................32
Figure 11.0  Simulation Test Intervals .........................................................38
Figure 12.0  CARS Data Collection from the 5-Vehicle Scenarios ..............41
Figure 13.0  NASA-TLEX Data Collection from the 5-Vehicle Scenarios ....42

"Distribution A: Approved for public release; distribution unlimited. 88ABW Cleared 12/03/2013; 88ABW-2013-5056."
LIST OF FIGURES (continued)

Figure 14.0   Comparisons of the 2-Vehicle and 5-Vehicle Scenarios with Increased Communications and Task Interruptions………………………………………………………………….43

Figure 15.0   Mean Pupil Data Collection……………………………………………………44

Figure 16.0   Mean Fixation Duration Data Collection…………………………………….44

Figure 17.0   CARS Mean MWL Rating for each 5-Vehicle Scenario…………………….45

Figure 18.0   NASA-TLX Mean Rating for each 5-Vehicle Scenario……………………45

Figure 19.0   Mean Fixation Duration for each 5-Vehicle Scenario……………………46

Figure 20.0   Mean Pupil Diameters for Each 5-Vehicle Scenario…………………….47

Figure 21.0   Mean Total MRT Resource Rating for each 5-Vehicle Scenario……….51
LIST OF TABLES

Table 1.0  Phase 2 – Experiment Factors for the CARS Tool…………………………39
Table 2.0  Phase 2 - Experiment Factors for the NASA-TLX Tool.......................39
ACRONYMS

AFB   Air Force Base
AFRL  Air Force Research Laboratory
ATC   Air Traffic Control
BLOS  Beyond Line of Site
CARS  Crew Awareness Rating Scale
Comms Communications
HSC   Human Supervisory Control
LOS   Line of Site
MTE   Mobil Test Environment
MWL   Mental Workload
NASA-TLX National Aeronautics and Space Administration-Task Load Index
ROZ   Restricted Operating Zone
SA    Situation Awareness
SME   Subject Matter Expert
UAV   Unmanned Aerial Vehicle
USAF  United States Air Force
VSCS  Vigilant Spirit Control Station
WSU   Wright State University

"Distribution A: Approved for public release; distribution unlimited. 88ABW Cleared 12/03/2013; 88ABW-2013-5056."
ACKNOWLEDGEMENTS

This research was part of a study with the Air Force Research Laboratory (AFRL), 711th Human Performance Wing, Human Effectiveness Directorate, Supervisory Control and Cognition Branch located at Wright-Patterson AFB. So, with sincere appreciation, I thank them for their leadership and collaboration on this project. Also, it is imperative that I thank my advisor and thesis director, Dr. Mary Fendley, for her direction, assistance and support over the duration of this research experiment.
1.0 INTRODUCTION

1.1 Overview and Problem Description

Human supervisory control (HSC) occurs when a human operator monitors a system and intermittently interacts with a computer interface to transform operator commands to detailed control actions on the system (Sheridan, 2012; Cummings & Guerlain, 2007). This method is used by pilots or supervisory controllers in the operations of unmanned aerial vehicles (UAVs). Most UAVs require multiple pilots to control a single vehicle; so, the United States Air Force is looking to change how UAVs are handled in efforts to increase the UAV-to-pilot ratio (Pond, Webster, Machuca, Colombi, Miller, & Gibb, 2012; Ruff, Calhoun, Draper, Fontejon, & Guilfoos, 2004). Therefore, there is a need to understand the effort that must be applied and constraints experienced to complete the pilot’s task requirements during the operation of multiple UAVs. One critical aspect of the pilot’s task requirements is the ability to manage multiple modes of communication. The pilot and the UAV are not co-located; therefore, the pilot controls the vehicle through radio and satellite communications.

There are two types of communication loops for controlling UAVs: Line-of-Site (LOS) and Beyond Line-of-Site (BLOS). Line-of-Site Vehicles use radio frequencies and
antennas to communicate with the vehicles; however, the Beyond Line-of-Site Vehicles use satellites. The Ground Control Station sends information through a circuit to a communication line or step-site (i.e. satellite ground communication up/down link). Next, from the step site a signal is sent to the satellite in the sky above the step-site. Afterwards, the satellite does a crosslink between satellites or a downlink to the UAV.

There are various modes of communication used to support the communication loop: radio, internet relay chat and telephones. This research will focus on communication tasks during the transit phase of operations. This phase of operations has demanding communication requirements; however, in addition to these demanding communication tasks, the UAV operator has a host of other tasks to include vehicle routing, which involves creating emergency and operational inputs, sensor manipulation to evaluate weather, and vehicle system checks. Consequently, this experiment will explore the mental workload relative to multi-modal communications while operating multiple UAVs during the transit phase of operations.

1.2 Research Objectives

The objective of this experiment is to assess mental workload (MWL) relative to multi-modal communications in operating multiple UAVs during the transit phase of operations. Mental workload reflects how difficult it is for the brain to accomplish task demands. This is an important measurement because it provides awareness as to where increased task
Demands within user operations may lead to poor or unacceptable performance (Cain, 2007). Mental workload and situation awareness have three major categories of measures based upon the nature of data collected: performance, subjective ratings and physiological measures (Vidulch & Tsang, 2012). Therefore, this research will use subjective ratings and physiological measures to evaluate mental workload in this experimental design.

The following (Figure 1) illustrates the manning of the UAV operation for this experimental scenario:

![UAV Manning for this Experimental Scenario](http://www.howstuffworks.com/predator.htm)

**Figure 1: UAV Manning for this Experimental Scenario**
**Operator 1** is responsible for launch and recovery tasks which include controlling the aircraft during takeoffs and landing using a LOS data link. At a pre-coordinated location, this Operator hands off the vehicle to Operator 2.

**Operator 2** is responsible for transit tasks, which includes monitoring the vehicle and navigating it to Operator 3. Once Operator 3 completes its mission tasks, it returns the vehicle to Operator 2, and Operator 2 hands off to Operator 1 for landing.

**Operator 3** is responsible for mission tasks, which means executing the mission using BLOS data link. Once the mission is complete, Operator 3 returns the vehicle to Operator 2 and performs a handoff.

### 1.3 Research Hypothesis

Based upon previous research (Dixon and Wickens, 2003; Pomransky and Wojciechowski, 2007; Schneider and McGrogan, 2011), the hypothesis of this experiment is that monitoring two UAVs by one operator is the ideal UAV-to-pilot ratio for optimal mental workload, as compared to five UAVs. In order to test this hypothesis, human-in-the-loop testing will be conducted with real operators supervising two or five UAVs simultaneously. The choice of two or five UAVs was dictated by the developers of the simulation. The simulations will be used to attempt validating the assumption that five UAVs are not manageable with one supervisor. Also, subjective measures will be used to measure mental workload with physiological measures applied to validate the results.
According to Dixon and Wickens (2003), the operation of two UAVs falls within the capabilities of a single operator but with degradation of performance. Pomransky and Wojciechowski (2007) conducted a study similar to the pilot’s role in this experiment. Their study evaluated operator tasks for manually making adjustments, monitoring video feedback and systems status of vehicles. Their results showed that one operator supervising two UAVs would meet the optimal threshold for overall workload but two (i.e. cognitive and visual workloads) of the four components that make up the overall workload measurement were above the optimal threshold, which could lead to performance errors. And, operation of three UAVs was clearly an overload condition. Another point of reference for optimal number of UAVs supervised by one operator was simulated at the Air Force Institute of Technology (AFIT) in 2011. This study analyzed workload performed for multi-aircraft control by using a single pilot to control multiple aircraft along with a sensor operator and mission intelligence coordinator for each vehicle (Schneider & McGrogan, 2011). Their model predicts that the pilot experiences low workload when operating one or two UAVs; however, there is a conflict with multi-tasking overlap when operating three or more vehicles. The aforementioned case studies provide the basis for my test hypothesis relative to manageable number of UAVs by one supervisor. Therefore, the test hypothesis is defined as five UAVs supervised by one pilot is a mental overload situation.
2.0 RELATED RESEARCH

2.1 Supervisory Control

Supervisory control is control by a human operator of an automated or semi-automated system; the operator controls such a system via a computer or a bank of computers that are controlling the system (Mitchell & Miller, 1986). In a broader sense, supervisory control entails interacting with a computer to transform data or to generate control actions (Sheridan, 2012). In human-computer systems with a human supervisor, the computer typically takes over the routine and simple decision making tasks. However, the human supervisor is in control of the higher level tasks. The concept of supervisory control developed as a part of research on how people on earth might control vehicles on the moon, and has since extended to other domains such as aircraft and search and rescue.

With supervisory control, function allocation plays an important role, because function allocation provides a rational means of determining which system–level functions should be carried out by human supervisors and which by machines. The human supervisor takes on five different functions in supervisory control: planning, teaching, monitoring, intervening and learning (Sheridan, 2012). The planning function requires the operator to
identify how to do a computer-based task. Secondly, the teaching function requires instructing the computer on what was planned. Next, the monitoring function necessitates checking the system’s actions to detect failures. Intervening is needed once the plan has been successfully completed or if the system needs assistance by specifying a new goal state or reprogramming. Lastly, the learning function is vital to acquiring information from the experience for continuous improvement.

The following outlines some of Sheridan’s suggestions on advantages of the supervisory control method (Moray, 1986):

1. Improves performance,
2. Reduces operator workload,
3. Improves task planning with the ability to view predictor displays,
4. Assists with monitoring, detecting and diagnosing system issues,
5. Makes system control easier by display and control aids.

This method is used by pilots or supervisory controllers in the operations of unmanned aerial vehicles (UAVs), which are used for military missions to include intelligence, surveillance, reconnaissance (ISR) and air combat. State and local governments also use UAVs for “such task as border patrol, search and rescue, forest fire monitoring, disaster response, and air traffic control” (Rowe & Liggett, 2009). Since most UAVs require multiple pilots to control a single vehicle, there is a need to understand the pilot’s task requirements in order to increase the UAV-to-pilot ratio. Therefore, if pilots are
overwhelmed with manual controls, there will be a management of fewer vehicles because they must dedicate their attention to lower level skill-based cognitive tasks. In order to understand these task requirements, there are many models and methodologies that can be applied. A **hierarchical task analysis** was chosen to analyze the operator’s functions and tasks, and will be explained in the next section.

### 2.2 Hierarchical Task Analysis

The Hierarchical Task Analysis (HTA) is a step-by-step methodology to understanding a task or system, and the task or system can be analyzed to any required level of detail. The HTA works by decomposing systems into a hierarchy of goals, sub-ordinate goals, operations and plans; it focuses on what the operator is required to do in terms of actions and/or cognitive processes to achieve a system goal (Salmon, Jenkins, Stanton, & Walker, 2010). The HTA specifies the overall goal of a particular system, the sub-goals to be undertaken to achieve this goal, the operations required to achieve each of the sub-goals specified and the plans, which are used to ensure that the goals are achieved (Salmon, Jenkins, Stanton, & Walker, 2010). It is a simplistic process that involves collecting data relative to the task or system being analyzed through observation, surveys, interviews, walkthroughs, and user trials to name a few. This data is then used to breakdown and describe the goals and sub-goals involved. The process of developing the HTA enables the analyst to deeply understand the task or system. The output of a HTA is extremely useful and forms the input for numerous human factors analyses, such as error analysis, interface design and evaluation and allocation of function analysis (Stanton, Salmon,
Walker, Baber, & Jenkins, 2010). Therefore, a HTA is used to understand and examine the interaction between the pilots and their UAV system tasks in order to support examining mental workload associated with communication tasks while operating multiple UAVs.

2.3 Modes of Communication

The UAV system has a reliance on robust and reliable communication both for the payload and the aircraft system; there is a need to ensure that safety and effectiveness are not compromised, because “communication plays a much more important role in the overall operation of a UAV than it does for manned aircraft because the men-in-the-loop are on the ground” (Clot, 1999). A primary concern of the separation between aircraft and operator is that the operator is withdrawn from a range of sensory cues available to the pilot of a manned vehicle. Instead of receiving sensory input directly from the environment, the UAV pilot receives only that information provided by onboard systems via remote communication (McCarley & Wickens, 2007). Therefore, some of the issues involved are loss of sensory cues for flight control, delays in the control and communication loop, and trouble with scanning the visual environment around the UAV.

The UAV system is a complex structure; it has many elements to include one or more air vehicles, ground control stations (GCS), a suite of communications (including chat, radios, phones, satellite link, etc.), and operations and maintenance crews. Therefore, the UAV
flight is mainly accomplished by teams of human operators using multimodal communications for crew coordination. In the computer-human interface, multimodal means interfacing using more than one sensory-motor modality (touch, sight, etc.) (Taylor, Neel, & Bouwhuis, 2000). In this system, there are typically three modes of communication utilized to support the UAV communication loop: radio, internet relay chat (IRC) and telephone. The radio provides secure voice communications. The IRC provides channels to send text messages in real time, and the telephone is used as a back-up system if radio or chat is not operative. Therefore, the ability to manage multiple modes of communication is a critical aspect of the pilot’s tasks, because “frequent, unpredictable, and oftentimes overwhelming, volumes of communication events can produce unmanageable levels of system induced workload” (Pond, et al., 2012). And, the risks of missed communications could be substantial in the command and control system, causing the UAV to collide or wreck with a manned vehicle. So, in efforts to understand the communication activity associated with the operation of UAVs, a Comms Usage Diagram (CUD) was developed.

2.4 Comms Usage Diagram

The Comms Usage Diagram (CUD) describes collaborative activity between users that are not geographically co-located (Stanton, Salmon, Walker, Baber, & Jenkins, 2010). This method is generic in nature but can be applied in many domains to include, naval warfare, fire services, air traffic control, military, rail and aviation. The output of the CUD method provides a description of the task under analysis along with a description of the
communications between operators performing the tasks, the technology used for the communications, and the advantages or disadvantages associated with the technology mode can be incorporated. The CUD draws attention to communication flaws. A number of different methods can be used to collect data for the CUD: interviews, walk-through analysis, questionnaires, and/or observational studies. The initial data collection phase of the CUD can be labor intensive, because a HTA is required in the process in order to define the task under analysis. So, in order to assess the mental workload associated with UAV communication tasks, the CUD will play a major role in understanding this complex command and control system and the multiple resource theory was applied to assess the ability of the UAV operator’s ability to perform.

2.5 Multiple Resource Theory (MRT)

Multiple Resource Theory (MRT) is a predictive model of an operator’s ability to perform while multi-tasking in a high workload environment or a complex system (Wickens, 2002). A complex system refers to a task environment that has a number of simultaneous activities which are time-shared (Liu & Wickens, 1988). An example of a complex system is a pilot flying an aircraft. Flying requires tracking multiple sources of information and managing several systems and controls, and it doesn’t take much to be inundated to the point that critical information is overlooked or important actions are neglected or performed incorrectly (Hopkins, 2013). So, the primary value of the MRT model is predicting relative differences in multitasking among varying conditions (Wickens, 2008). It consists of four different dimensions and three components.
MRT proposes that there are four dichotomous dimensions of information processing that explains the difference in time-sharing performance. The four dimensions are processing stages, perceptual modalities, visual channels and processing codes, which are outlined as follows:

1. Processing Stages - This dimension suggests that perceptual and cognitive demand tasks (e.g. working memory) use different resources from those essential for the selection and execution of responses. Individuals are able to effectively time-share between activities that require perceptual and cognitive demand, because these resources appear to be the same.

2. Perceptual Modalities – This dimension indicates that time-sharing is better in cross-modal conditions versus intramodal conditions, because attention can be distributed better between the eye and ear rather than two visual or two auditory channels since visual perception uses different resources than auditory perception.

3. Visual Channels – This dimension reveals that within the visual channel there are two different types of resources: focal and ambient vision. Focal vision primarily supports object recognition (e.g. reading text); whereas, ambient vision is responsible for perception of orientation and movement.

4. Processing Codes – This dimension implies that there is a contrast between the resources used for spatial tasks and verbal activity. Separating the spatial and verbal resources should account for a high degree of efficiency.
In summary, MRT consists of three components: demand, resource overlap and allocation. People are capable of multi-tasking until task demands exceed available resources. There are several resources that the human mind can devote to task demands either individually or collectively: visual, auditory, cognitive, motor and speech. Therefore, the resources concept is based upon the assumption that human operators have a limited capacity for processing resources that may be allocated to task performance. Two task demands require more resources than one, so time-sharing can lead to where one or both have fewer resources available. Also, MRT predicts that tasks will interfere when competing for common processing resources, therefore leading to poor performance (Liu & Wickens, 1988), which could lead to a compromise in system safety and effectiveness.

2.6 Information Processing

Information processing is very important to human performance, because in situations where humans interact with systems, “the operator must perceive information, must transform into different forms, must take actions on the basis of the perceived and transformed information, and must process the feedback from that action, assessing its effect on the environment” (Wickens & Carswell, 2006). According to the Wickens and Holland (2000) framework for information-processing, there are three stages to information processing: sensory memory, working memory and long-term memory. Sensory memory is affiliated with interpreting data based upon past experiences that was sensed or perceived. However, working memory (i.e. short term memory) is a brain system that deals with immediate conscious processing of data; it is of limited capacity and heavily
demanding to attention of information that was not sensed or perceived but generally an internal thought. And, long-term memory is cognitive storage of data; it keeps extensive amount of information but it is not always able to be fully recalled. So, in the UAV system model, the operator primarily utilizes their working memory, which has limited capacity. Therefore, it is important that the operator is not over saturated with tasks that can cause a mental overload situation, because overloading the operator could lead to performance degradation. With the UAV system model, there is interaction between an operator and an assigned task; this is referred to as mental workload or simply ‘workload’. This is an important measurement because it provides awareness as to where unacceptable performance may result from an increase in task demands.

### 2.7 Mental Workload Assessment

Mental workload plays a significant role in many complex command and control systems, because it reflects how difficult it is for the brain to accomplish task demands. In other words, it indicates when there are limits on the brain’s inability to deliver. So, through the examination of tasks confronted by pilots and other crew members, Gartner and Murphy (1979) define workload in terms of task demand, effort and performance. Task demands refer to what the operator is required to complete and the operating constraints experienced during the time for completion. Effort is dependent upon how demanding is the operator’s task. Performance is based on the actual accomplishments of the operator. In other words, workload attempts to measure task performance relative to an operator’s capability.
is an important measurement because it provides awareness as to where increasing task demands could result in poor performance (Cain, 2007). So, workload measurements can be used in the assessment of various capacities to include system desirability and system optimization. These measurements can help to assess the desirability of various systems if performance is used to distinguish between choices.

Workload of a single operator can be measured through a utilization metric. “The utilization metric is defined as the percentage of time the operator is busy, \( p = \frac{\text{operator busy}}{\text{total operation time}} \)” (Cummings & Guerlain, 2007). Busy in this metric refers to the time the operator is actively engaged in either a retargeting decision or responding to communication messages. According to a study by Cummings and Guerlain (2007), the best performance for operator utilization is found between 40% and 70%. Their study showed that utilization rates below 40% yield a performance that declined under low workload. However, on the other hand, their analysis confirmed that operators performed significantly worse at utilization rates greater than 70%. The Cummings and Guerlain investigation was based upon supervising multiple tactical tomahawk missiles; the supervision of these missiles is similar to managing UAVs. “Tomahawk missiles could be stationed in loiter patterns to await further orders, much like what is envisioned for both surveillance and combat UAVs” (Cummings & Guerlain, 2007). There are similarities for the demand to re-plan missions and rearrange resources in real time under pressure. Therefore, their experimental results are equally applicable to unmanned vehicles. In addition, the Cummings and Guerlain (2007) experimental results coincide with the
Yerkes-Dodson Law. This theory was originally developed by Robert M. Yerkes and John D. Dodson in 1908 to measure performance and arousal levels, which refers to stress, anxiety, and motivation levels (Figure 2). This law indicates that improved performance can result from arousal, as long as the arousal levels are not too high, which can lead to detriment in information processing and decision making as well as performance.

![Yerkes-Dodson Curve](http://psychtopics.files.wordpress.com/2011/08/yerkes-dodson.jpg)

Figure 2: Yerkes-Dodson Curve

Therefore, increased operator performance, reduction of operator stress, and increased system safety could potentially be achieved by optimizing mental workload (Ahistrom & Friedman-Berg, 2005).
Although mental workload cannot exactly be observed, it can be deduced from the operator’s behavior and measured indirectly through its effects on system performance, psychological or physiological processes drawn from subjective judgments.

2.7.1 Physiological Measures

As cited in (Ahistrom & Friedman-Berg, 2005), “findings indicate that blink rate, blink duration and saccade duration all decreased while pupil diameter, the number of saccades and the frequency of long fixations all increased with increased workload (Iqbal, Adamczyk, Zheng, & Bailey 2004, 2005; Iqbal, Zheng, & Bailey, 2004; Lin, Zhang, & Watson, 2003; Rognin, Grimaud, Hoffman, & Zeghal, 2004; Stein, 1992; Van Orden, 2000; Van Orden et al., 2000; Veltman & Gaillard, 1998; Zeghal, Grimaud, Hoffman, & Rognin, 2002)”. Two psychological measures have been chosen as the index to measure mental workload in this experiment: fixations and pupil diameter. Fixations are moments when the eyes are relatively stationary. So, the fixation duration can be interpreted as increased mental workload can increase the fixation duration. Just as, pupil diameter is an indication of cognitive effort; large pupil diameters indicate increased mental workload (Poole & Ball, 2005). These measures will be employed to support the subjective workload techniques for this experimental design.
2.7.2 Subjective Workload Techniques

A variety of subjective workload techniques have been developed to assess mental workload. One approach for a proposed assessment methodology to choose a workload measurement technique is outlined as follows (O'Donnel & Eggemeir, 1986):

1. Sensitivity: method must be capable to discriminate between significant changes in workload inflicted by a task or group of tasks.
2. Diagnosticity: method must be capable to indicate the amount of workload variation on different operator capacities.
3. Intrusiveness: method should not interfere with the performance of the operator’s tasks, causing degradation in task performance.
4. Implementation Requirements: elements dealing with measurement equipment and procedures must be considered for ease of implementing a technique.
5. Operator Acceptance: method should have face validity and be acceptable to the operator.

Upon careful review and use of the above stated Workload Assessment Methodology as a guide, nine workload techniques (Appendix 7.1) were evaluated for this study and the **Crew Awareness Rating Scale (CARS)** was chosen.
2.8 Crew Awareness Rating Scale

As previously stated, mental workload reflects how difficult it is for the brain to accomplish task demands. This is an important human performance measurement because it provides awareness as to where an increase in task demands may result in unacceptable performance. In order to measure mental workload, the Crew Awareness Rating Scale (CARS) was used. CARS is a subjective and multidimensional situation awareness (SA) and workload assessment tool. This subjective measurement tool uses two sets of four questions as it relates to mission performance to assess various dimensions (Figure 3): perception, comprehension, projection and intention.

\[
\begin{array}{|c|c|}
\hline
\text{Situation Awareness} & \text{Workload} \\
\hline
1: Perception - Content & 2: Perception - Process \\
3: Comprehension - Content & 4: Comprehension - Process \\
5: Projection - Content & 6: Projection - Process \\
7: Intention - Content & 8: Intention - Process \\
\hline
\end{array}
\]

The eight CARS rating scales (reproduced from Situational Awareness and the Crew Awareness Rating Scale by B. McGuinness, 1999)

According to McGuinness (1999), the dimensions are defined as follows:
1. Perception – acquiring and understanding information of external sources from readouts, displays, etc.
2. Comprehension – understanding the meaning of the sum of information
3. Projection – comprehending how the situation will likely change over time
4. Intention – recognizing the available courses of action in the current situation

The CARS method allows the experimenter to select only those dimensions of interest for inclusion in the survey (McGuinness B., 1999). Therefore, only the workload dimensions of the CARS model were used for the purpose of assessing mental workload in this experiment. So, after each simulation run, the participant was given a CARS survey (Appendix 7.2) to complete, which gave a quantitative measure to each MWL dimension. Each dimension was rated on a scale from 1 – 4: easy (1), manageable (2), difficult (3) and unmanageable (4).

This subjective measurement tool is based on the three-level model of situation awareness by Endsley, 1995 that has been and could be used for piloting aircraft, power generation, petro-chemical, nuclear, command and control, and medicine (Stanton, Chambers, & Piggott, 2001). Subjective measures, which are based upon the judgment of the operator, have been widely used to assess mental workload, because these measurement tools have proven to be valid and reliable. They have the advantage of posing the fewest implementation problems as only paper and pencil are required for responses (O'Donnel & Eggemeir, 1986). CARS can be applied in a number of applications to include civil or
military cockpit evaluation trials, field assessments of SA in army commanders or the evaluation of control room operators (McGuinness B., 1999). However, its primary domain of application is for Military use. In a case study by McGuinness and Ebbage (2002), the CARS method was used to assess digitized versus radio communications in an Army battlespace simulation model. With this model two exercises were performed by subjects: one in which all communications between the headquarters and its battlefield units took place using the standard radio net and a second in which most of the communications were either supplemented or replaced with digital technology, such as data-link, text messaging and automatic location reporting. In the experiment, CARS was used to measure the mental content and mental processing of the situational awareness involved in the simulation. The CARS metrics gave data and insight into the impact of digitizing their communication methods. Similarly, the goal for using the CARS model in this research experiment is to gain data and insight relative to mental workload of the communication tasks on one operator supervising multiple UAVs.

Another study, that utilized the CARS methodology, was during an Army field experiment (Matthews & Beal, 2002). Cadets were expected to complete a series of infantry missions and the purpose of the exercise was to develop leadership skills in the cadets. The intent of the mission was to compare the function of the positions between the Platoon and Squad Leaders during the field exercise. The Mission Awareness Rating Scale (MARS), which is based on CARS, was used to analyze situation awareness and workload as a function of the leadership position. This case demonstrates another aspect as to how the CARS tool

"Distribution A: Approved for public release; distribution unlimited. 88ABW Cleared 12/03/2013; 88ABW-2013-5056."
can be applied. Although the CARS tool has been validated, there are limited studies published using the CARS method; therefore, to validate the CARS results, this research employed the National Aeronautics and Space Administration-Task Load Index (NASA-TLX) as a support method.

2.9 NASA-TLX

NASA-TLX is a subjective MWL assessment tool; it is the most commonly applied MWL assessment tool. It has been used in numerous domains including civil and military aviation, driving, nuclear, power plant control room operation and air traffic control and cited in previous studies, like (Bruce, 2003; Endsley & Rodgers, 1998; Bertram, et al., 1992; Cao, Chintamani, Pandya, & Ellis, 2009). Its utilization has extended far beyond its original application of aviation and crew complement. It is “being used as a benchmark against which the efficacies of other measures, theories or models are judged” (Hart, 2006). It is taught in university courses and used in diverse situations such as aircraft certifications, operator rooms and nuclear power plant control rooms.

The NASA-TLX uses a multi-dimensional scale to measure operator task performance, which consists of six subscales (Appendix 7.3):

1. Mental demand – measures whether the perceptual activity (e.g. thinking, looking, searching etc.) was easy or demanding
2. Physical demand – measures whether the physical activity (e.g. controlling, activating, etc.) was strenuous or relaxed
3. Temporal demand – measures how much pressure was felt with the pace of activity (e.g. slow, fast, hectic, etc.)
4. Effort – measures how difficult work was applied to accomplish the task
5. Performance – measures how satisfied the subject was with their execution of accomplish the tasks
6. Frustration level – measures how the subject felt during the task (e.g. irritated, relaxed, stressed, etc.)

The six dimensions are rated on a scale from 0 to 100, where 0 represents the lowest task demand and 100 is highest task demand with the exception of the performance dimension, 0 indicates high demand and 100 is low demand.

This method deploys a weighting procedure that presents 15 pairwise combinations to the participants and has them to select the scale from each pair that has the most effect on the workload during the task under analysis (Stanton, Salmon, Walker, Baber, & Jenkins, 2010). The most common modification to the NASA-TLX has been to eliminate the weighting process all together (Hart, 2006). This method has been referred to as the Raw NASA-TLX (RTLX); the sub-scale ratings are averaged to create an estimate of the overall workload. RTLX will be used in this research experiment as the subjective MWL validation tool.
3.0 RESEARCH COMPONENTS

3.1 Cognitive Model

A cognitive model was developed in order to better understand the human-system interaction that impacts the operator’s workload. This model (Figure 4) assists with analyzing the system’s process as it relates to communication tasks while supervising multiple UAVs in order to measure human performance. It is a hybrid model composed of three validated human factors methods: Hierarchical Task Analysis (HTA), Comms Usage Diagram (CUD) and Crew Awareness Rating Scale (CARS). First, the HTA was used to gain a deeper understanding of the system’s overall task requirements. Subsequently, a CUD was developed to illustrate the communications and implied relationships between teams/crews. This provided an understanding of the communication activity that takes place in the coordination of the UAV mission. Also, the CUD was used to establish the assessment tool in order to measure mental workload. Lastly, the CARS tool was employed to measure mental workload of the communication tasks.
3.2 Vigilant Spirit Control Station

The testbed used in this experiment is a Vigilant Spirit Control Station (VSCS). VSCS (Figure 5) allows researchers to explore different supervisory control interface scenarios to evaluate the supervision of multiple UAVs by a single operator.
According to Rowe & Liggett (2009), it is a simulated environment used to model:

- various vehicle payloads, sensors, and human factors testing tools; dynamic mission planning interfaces for interacting with vehicle supervision and control; a robust and flexible software architecture that allows for multiple configurations to accommodate diverse missions across a multitude of vehicle platforms; and finally the interoperability and communication across these vehicle platforms and the associated Ground Control Stations.

In summary, VSCS provides analysts a research testbed to evaluate advancement in multi-UAV supervisory control.

### 3.3 Eye Tracking

In order to gather quantitative data to support the workload subjective measures, eye tracking was used. Eye tracking is the process of measuring eye activity: measuring where an individual is looking (i.e. gaze direction) at any given time or the sequence in which the
eyes are shifting from one location to another. An eye tracker is a device used for measuring eye positions and eye movement. Eye trackers are used to evaluate the human-machine interaction. The eye tracking data is collected using either a remote or head-mounted ‘eye-tracker’ connected to a computer. The eye tracking data can be used to collect information relative to gaze direction, blink frequency and pupil diameter. This information can be employed to validate design options and explore how the user gathers information relative to visual processing. Since eye movements are generally thought to be involuntary, eye tracking provides objective data of user’s visual interaction with a system (Bruneau, Sasse, & McCarthy, 2002).

This research will use a remote, off-body eye tracker, the Smart Eye Pro. The Smart Eye Pro uses multiple cameras (Figure 6) to determine where the participant is looking on a set of screens. This system is non-intrusive; therefore, there is no contact with the human body. It collects gaze data at a sampling rate of 60 Hz.
3.4 Mobile Test Environment (MTE)

The MTE (Figure 7) is a truck-and-trailer mobile test and evaluation station for unmanned systems.
Figure 7: Vigilant Spirit Trailer

It is 35 feet in length with a 550-square foot mobile office. The MTE is a climate control environment with standalone power. It has analysis and operator capability for research and development. Its hardware and software capabilities are as follows (WSRI VS-MTE Trailer, 2012):

- 11 rack-mounted CPUs
- 72 Terabyte RAID
- 6 dual-monitor workstations
- 2 quad-monitor workstations
- 6-Camera Smart Eye Trackers
- Real IR Software
- FLAMES /Viper
- Morae software
- MAPPS
- Vigilant Spirit Control Stations (hosting for AFRL)
The WSU research team played the simulation staff, which required playing the role of Operator 3, Operator 1 and other players. This took place at the workstation in Figure 8.

Figure 8: Test Administrator Stations
The participants acted as Operator 2 in the VSCS station in Figure 9.

![Simulation Stations](image)

**Figure 9: Simulation Stations**

### 3.5 Research Framework

This experimental research follows a two-phase approach (Figure 10) with phase 2 building on the previous. The following outlines the process:
Phase 1 is an exploratory phase in which observations and interviews with the expert and novice participants took place. This phase was used to assemble a knowledge base of information for understanding the problem. A secondary objective was to gain feedback on MWL and determine if there may be a potential difference between novice and expert users.

Phase 2 is an actual test run administered by the WSU team at the Mobil Training Environment (MTE). This phase used all novice users. The objective of this phase was to assess and manipulate MWL by measuring three in-flight variables: increased communication tasks, task interruptions and a combination of these two tasks simultaneously.
The following section describes the research methodology, to include detail on each of the two phases.
4.0 RESEARCH METHODOLOGY

4.1 Phase 1: Preliminary Design Assessment

This model investigates using a single UAV pilot to control multiple aircraft during the transit phase of operations: moving to, from, and between target areas. In addition, this model seeks to create a multi-vehicle transit operation in order to maximize UAV time-over-target when 24/7 coverage is a requirement or if transit makes up a large percentage of the overall sortie duration. A sortie is “a sudden issuing of troops from a defensive position against the enemy or one mission or attack by a single plane” (Merriam-Webster, Incorporated, 2013).

Potential end users (i.e. Subject Matter Experts, SMEs) were brought in to conduct simulated walkthroughs for the Phase 1 assessment. The following describes some background on the scenario and the types of tasks that Operator 2 had to execute using VSCS:

1. Gaining/losing handover procedure from/to the Operator 1
2. In-flight vehicle supervision during the transit phase of operations
   a. Intermittent system status checks
   b. Transit environment status checks (e.g. weather, traffic, airspace changes)
   c. Communications with different entities
d. Off-nominal events (e.g. system emergencies, mission changes)

3. Losing/gaining handover procedure to/from Operator 3

The VSCS interface was used in order to simulate real-time operations. During the simulated scenario many events were triggered to include:

1. Restricted Operating Zone (ROZ) popup
2. Icing alert
3. Traffic avoidance
4. Radio calls
5. Lost datalink alert
6. Radio frequency changes

This scenario provided a “realistic” environment for operations. It consisted of operating in civilian airspace with civilian air traffic controllers (ATC) and in combat airspace for Operator 3. The scenario has a 24-hour schedule in which the aircraft flies 16 hour sorties from takeoff to landing, with four hour transit times each way. The takeoff and landing take a half hour each; and, handoffs are almost instantaneous.

The Phase 1 observations and interviews support there is not a significant difference between mental workload experienced by the novice and expert users for the Operator 2 functions, so this affords the opportunity to use novice users for the Phase 2 design as I
design and test scenarios to manipulate MWL. Also, based upon the expert user’s feedback, more of the “real environment” should be captured to develop a richer simulation for Phase 2.

4.2 Phase 2: Experimental Design

4.2.1 Test Scenario

The rationale for this experimental design is based upon feedback from interviews with the Phase 1 SMEs. Many suggestions were given on how to improve the Phase 1 preliminary design, which is also referred to hereafter as baseline design, in order to capture more of the real environment. Therefore, this experimental design focused on three variables:

1. Increase of chat activity
2. More task interruptions
3. Two simultaneous emergencies

Chat Activity

The Phase 2 design increased auto chat messaging frequency by 200% from the baseline scenario. So, in the Phase 2 scenario, three auto chat messages were sent to the Operator 2 in the 16-minute scenario, in which they had to read and determine the appropriate course of action.
Also, this experimental scenario incorporated chat communication as an option during all clearances. Therefore, Operator 2 was not limited to radio only during the civilian airspace. This would allow for mapping and measuring of the increase use of chat mental resources as it relates to workload.

**Task Interruptions**

During the interviews with the Phase 1 SMEs, a suggestion was made to increase the task interruptions. Therefore, this design incorporated more quick action items. An example of a quick action item is weather. During this experimental scenario, two quick action items of weather are incorporated into the 16-minute test interval.

**Simultaneous Emergencies**

This experimental design applied an emergency to two vehicles at the same time.

During the Phase 2 test, discrete tasks from the baseline scenario were used to evaluate how unpredicted events can affect the pilot’s MWL. The simulation was approximately one-hour and twenty minutes to include the following test intervals (Figure 11):
There were five test intervals (Figure 8) running approximately sixteen minutes each. Each test interval included high communication tasks along with the normal operating tasks for Operator 2. In order to eliminate the potential of the nuisance variable of training effect, the test intervals were randomized among the participants during the simulation.

4.2.2 Test Objective

The Phase 2 experiment will:

1. Objective 1: validate whether there is a significant difference between the 2-vehicle and 5-vehicle scenario relative to MWL using a different set of subjects

2. Objective 2: evaluate total MWL for four different scenarios while monitoring 5-vehicles using the CARS method

3. Objective 3: validate the CARS method for measuring MWL using two additional MWL tools: NASA-TLX and physiological measures of pupil diameter and fixations

"Distribution A: Approved for public release; distribution unlimited. 88ABW Cleared 12/03/2013; 88ABW-2013-5056."
4. Objective 4: determine if there is a significant difference in MWL among the baseline scenario with increased communications, task interruptions and a combination of both while monitoring 5 vehicles.

4.2.3 Components of the Experiment

Part 1 Level of Factors: This is a one factor experiment with five variables and four levels, which is outlined as follows (Table 1):

<table>
<thead>
<tr>
<th>Factors</th>
<th>Variables</th>
<th>Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 Vehicles</td>
<td>Baseline MWL</td>
<td>easy</td>
</tr>
<tr>
<td></td>
<td>Increase Comms MWL</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Increased Tasks MWL</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Increased Comms &amp; Tasks MWL</td>
<td></td>
</tr>
</tbody>
</table>

Part 2 Level of Factors: This is a one factor experiment with four variables and twenty levels, which is outlined as follows (Table 2):

<table>
<thead>
<tr>
<th>Factors</th>
<th>Variables</th>
<th>Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 Vehicles</td>
<td>Baseline MWL</td>
<td>5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100</td>
</tr>
<tr>
<td></td>
<td>Increase Comms MWL</td>
<td>5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100</td>
</tr>
<tr>
<td></td>
<td>Increased Tasks MWL</td>
<td>5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100</td>
</tr>
<tr>
<td></td>
<td>Increased Comms and Tasks MWL</td>
<td>5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100</td>
</tr>
</tbody>
</table>

"Distribution A: Approved for public release; distribution unlimited. 88ABW Cleared 12/03/2013; 88ABW-2013-5056."
4.2.4 Operating Procedure

This was a controlled test that took place in the WSU MTE. Four novice participants were used for the test. These are the steps that they followed:

1. Each participant was given a consent form to review and sign for participating in the experiment.
2. Each participant was given a pre-test questionnaire in order to obtain background information.
3. Each participant had a profile in the SmartEye eye tracking system developed and the system was calibrated.
4. Each participant used the VSCS testbed for the flight scenario.
5. Each participant was asked to act as a UAV operator and remotely operate multiple aircraft. The operator’s job was to monitor each UAV’s progress and re-plan aspects of the mission in response to unexpected events.
6. Communications was a major portion of the experiment. So, the scenario provided a realistic and robust test of the anticipated communications issues, which encompasses radios and internet relay chat.
7. All participants received a training session on the computer interface. Immediately following the training, the simulation took place.
8. Each participant was given five different scenarios requiring 16 minutes each during the simulation.
9. After each scenario, the CARS (Appendix 7.2) post-questionnaire and NASA-TLX workload assessment (Appendix 7.3) tools were administered.

"Distribution A: Approved for public release; distribution unlimited. 88ABW Cleared 12/03/2013; 88ABW-2013-5056."
4.2.5 Results

At the end of each test run, data was collected using the CARS and NASA-TLX tools and the SMARTEye eye tracking system. The CARS and NASA-TLX data were analyzed using the JMP software for statistical analysis and the SMART Eye data was analyzed using MAPPs software to further examine MWL.

The data was collected from four novice users; there were a total of 80 CARS responses (Appendix 7.3.1, 7.3.2, 7.3.3, 7.3.4 and 7.3.5) and 120 NASA-TLX responses (Appendix 7.3.6 and 7.3.7). Figures 12 and 13 summarize the CARS and NASA-TLX data collection.

![CARS Mean MWL Rating for Each 5-CAP Scenario](image)

**Figure 12: CARS Data Collection from the 5-Vehicle Scenarios**

"Distribution A: Approved for public release; distribution unlimited. 88ABW Cleared 12/03/2013; 88ABW-2013-5056."
Figure 13: NASA-TLX Data Collection from the 5-Vehicle Scenarios

The following graph (Figure 14) summarizes the responses for the 2-vehicle and 5-vehicle scenarios with increased communications and task interruptions using the CARS scale. The 2-vehicle scenario data was collected in order to perform a statistical comparison with the 5-vehicle scenario data. This provided an opportunity to verify with a different set of subjects whether there is a significant difference between the 2-vehicle and 5-vehicle scenario relative to MWL.
The SMARTEye tracking system was used to collect data relative to pupil diameter (Appendix 7.3.8, 7.3.9, 7.3.10 and 7.3.11) and fixation duration for each participant (Appendix 7.3.12, 7.3.13, 7.3.14 and 7.3.15). The following summarizes the pupil diameter (Figure 15) and fixation (Figure 16) data collection:
Figure 15: Mean Pupil Data Collection

Figure 16: Mean Fixation Duration Data Collection

"Distribution A: Approved for public release; distribution unlimited. 88ABW Cleared 12/03/2013; 88ABW-2013-5056."
Figures 17 and 18 illustrate the mean MWL experienced during operations of each 5-Vehicle Scenario using the CARS and NASA-TLX subjective rating scales.

![Graph](image1.png)

**Figure 17: CARS Mean MWL Rating for each 5-Vehicle Scenario**

![Graph](image2.png)

**Figure 18: NASA-TLX MWL Mean Rating for each 5-Vehicle Scenario**
The CARS and NASA-TLX scales both show the same trend. The increased communication scenario showed a slightly higher load over the baseline. And, the participants experienced the greatest load during the task interruptions trial. However, there was a reduction of workload experienced during the increased communication with task interruption trial in comparison to the task interruption trial.

As previously mentioned in section 2.7.1, fixation duration can be interpreted as there is an increase in cognitive effort or mental workload as the fixation duration increases. Therefore, the results of the mean fixation duration (Figure 19) for each 5-vehicle scenario indicate that task interruptions placed the greatest load on the participants.
As previously mentioned in section 2.7.1, increased pupil diameter is an indicator of increased cognitive effort or mental workload. Although the participants rated the task interruption trial with the highest load, their physiological measure of pupil dilatation does not correlate with their subjective ratings. The mean pupil diameter results (Figure 20) indicate that the task interruptions trial experienced the lowest load.

4.2.6 Discussions

In summary, for each objective the results support the following:

1. Objective 1: This experiment could not determine a significant difference relative to MWL between the operation of 2-vehicles and 5-vehicles for Operator 2, \( t(5.268737) = 0.727607, p = 0.7511 \) as shown in Appendix 7.4.1. The total MWL mean for Operator 2 indicates that managing 2 or 5 vehicles simultaneously with
the experimental scenario is manageable. However, the 95% confidence interval for mean MWL when managing 5 vehicles includes values above 2, indicating that this experiment was not able to establish that the average workload was below a 2 rating (Appendix 7.4.1); one out of the four participants experienced a total MWL rating slightly above the manageable range.

2. Objective 2: The CARS rating scale was used to measure MWL for each 5-vehicle scenario. The results showed that Operator 2 experienced the greatest load during the operation of the task interruption trial (Figure 17).

3. Objective 3: The NASA-TLX was applied to support the findings of the CARS scale. The NASA-TLX results (Figure 18) showed the same trend as the CARS scale for MWL over each scenario. Also, the mean fixation durations (Figure 19) for the participants showed the same trend as the CARS and NASA-TLX scales. However, the mean pupil diameter (Figure 20) showed a quite different trend for MWL of the Task Interruption and Increased Comms and Task Interruption trials, in comparison to the other three measures of CARS, NASA-TLX and mean fixations. This could be caused by the sensitive nature of the pupil diameter measure. This measure is less diagnostic in environments that require substantial eye and head movement; and, when overload arises, it can become unresponsive to variations or even reverse it responses (Cain, 2007).

4. Objective 4: The CARS data was analyzed using the JMP statistical software to do a means comparison of the MWL ratings (Appendix 7.4.2). The results indicate that there is not a significant difference between the mean MWL ratings for the four various scenarios. In addition, the NASA-TLX data was analyzed in the same

"Distribution A: Approved for public release; distribution unlimited. 88ABW Cleared 12/03/2013; 88ABW-2013-5056."
manner (Appendix 7.4.3) and the results also indicated that a significant difference between the mean MWL ratings for the four various scenarios could not be determined. In addition, both data sets (i.e. CARS and NASA-TLX) showed a peak in MWL above the other scenarios during the task interruption trial. Therefore, the results of this experiment support that the greatest workload was experienced during task interruptions.
5.0 CONCLUSION

As previously stated, the Multiple Resource Theory (MRT) can be used to predict relative differences in multitasking among varying conditions. It can assist with identifying diagnostic information about workload sources. So, in order to map the MRT resources for the communication tasks in the Phase 2 experimental scenarios, the cognitive model from Section 3.1 was applied. First, the elements of the cognitive model were referenced. Next, the communication tasks were crossed referenced to the HTA, which were then matched to the CUD source. The CUD specified the computer interface used for the communication tasks. Afterwards, each communication task was mapped with the MRT ratings scale. The ratings for each mental resource were averaged in order to evaluate the differences among the varying conditions. Therefore, the communication tasks for each scenario were mapped using the Multiple Resource Theory (MRT) rating scale (Appendix 7.4.4). The following graph summarizes the findings (Figure 21):
The MRT mapping supports that the Task Interruptions scenario experiences the greatest workload of the four scenarios. The task interruption scenario had the highest demand of the other scenarios for the cognitive and visual tasks. However, when the task interruptions are combined with increased chat communication tasks, the users experienced a lower demand for auditory tasks, which reduces the total mental workload experienced.

A plausible cause for the reduction in load during the increased comms and task interruption scenario is diversion. The operator could be diverted from a more strenuous task when a comms task interrupts which reduces the overall mental load as indicated by
the fixations measures. During the increased comms and tasks interruptions trial, the user’s fixation duration (Figure 19) is less than the task interruption trial.

In summary, this thesis developed and applied a cognitive model for understanding the human-system interaction that impacts an operator’s workload, and followed a two-phased approach to analyzing the effects of multimodal communications on MWL during the supervision of two and five UAVs. In addition, the objectives of each of the phases were met. Phase 1 was used to gain a knowledge base of the problem and experience through literature reviews, interviews and observations. In addition, the Phase 1 feedback indicated similar workload expectations between novice or expert users. So, in Phase 2, novice users were able to be used in a new experimental design in efforts to manipulate MWL using three in-flight variables. Also, Phase 1 and 2 analyses conveyed that the operation of five vehicles by one operator could cause an overload situation for mental resources, which can lead to an unsafe operating condition. In Phase 2, subjective and physiological measures were used to evaluate MWL. The results indicated that the test variable of task interruptions had the greatest influence on human effectiveness relative to MWL while managing five UAVs. This suggests that reducing the complexity of tasks rather than number of UAVs controlled would be more effective in reducing the operator’s MWL.
6.0 FUTURE WORK

Four areas for future work have been identified in order to test the effectiveness of one pilot managing multiple vehicles during transit operations:

1. Based upon feedback from the SMEs, more of the “real environment” should be captured in order to develop a richer simulation for a more accurate measure of MWL.

2. This study was limited to monitoring only two or five vehicles: two vehicles are manageable and five vehicles can experience difficulty with tasking. Therefore, future studies should explore monitoring three or four vehicles in order to confirm the optimal UAV-to-pilot ratio for MWL measures or evaluate reducing task complexity.

3. Further research should be performed to examine possible performance errors during the scenarios and map these back to the task demands and resources in MRT.

4. The participants for the Phase 2 study were limited to four; however, I recommend that a larger subject pool is utilized in order to further validate significant differences among the variables.
7.0 APPENDIX

7.1 WORKLOAD ASSESSMENT TOOLS

Workload Assessment Tools

Assumptions:
1. Pen and paper are the required tools for each method.
2. Training time for each method is considered low.

<table>
<thead>
<tr>
<th>Method</th>
<th>Type of Method</th>
<th>Domain</th>
<th>Validation Studies</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. NASA-TLX</td>
<td>Multi-dimensional subjective rating tool</td>
<td>Generic</td>
<td>Yes</td>
<td>1. Quick and easy to use, requiring little training or cost</td>
<td>1. More complex to analyse than uni-dimensional tools</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2. Consistently performs better than SWAT</td>
<td>2. TLX weighting procedure is laborious and adds more time to the procedure</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3. TLX scales are generic allowing the technique to be applied in any domain</td>
<td>3. Caters for individual workload only</td>
</tr>
<tr>
<td>2. MCH - Modified Cooper Harper Scales</td>
<td>Uni-dimensional subjective rating tool</td>
<td>Generic</td>
<td>Yes</td>
<td>1. Quick and easy to use, requiring little training or cost</td>
<td>1. Unsophisticated measure</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2. Widely used in a number of domains</td>
<td>2. Limited to manual control tasks</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3. Data obtained is easier to analyze than multi-dimensional data</td>
<td>3. Not as sensitive as the TLX or SWAT</td>
</tr>
<tr>
<td>3. SWAT - Subjective Workload Assessment Technique</td>
<td>Multi-dimensional subjective rating tool</td>
<td>Generic (Aviation)</td>
<td>Yes</td>
<td>1. Quick and easy to use, requiring little training or cost</td>
<td>1. More complex to analyse than uni-dimensional tools</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2. Multi-dimensional</td>
<td>2. Data is collected post-trial. There are a number of problems with this, such as a correlation with performance</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3. SWAT sub-scales are generic, allowing the technique to be applied in any domain</td>
<td>3. The output does not offer a rating of participant MWL as such, only a rating of which tasks or devices imposed greater MWL than others</td>
</tr>
<tr>
<td>4. SWORD - Subjective Workload Dose</td>
<td>Subjective paired comparison technique</td>
<td>Generic (Aviation)</td>
<td>Yes</td>
<td>1. Quick and easy to use, requiring little training or cost</td>
<td>1. More complex to analyse than uni-dimensional tools</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2. Very effective when comparing the MWL imposed by two or more interfaces</td>
<td>2. Data is collected post-trial. There are a number of problems with this, such as a correlation with performance</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3. MWL ratings may correlate with task performance dimensional tools</td>
<td>3. The output does not offer a rating of participant MWL as such, only a rating of which tasks or devices imposed greater MWL than others</td>
</tr>
<tr>
<td>5. DRAWS - Defence Research Agency Workload Scales</td>
<td>Multi-dimensional subjective rating tool</td>
<td>Generic (Aviation)</td>
<td>No</td>
<td>1. Quick and easy to use, requiring little training or cost</td>
<td>1. More complex to analyse than uni-dimensional tools</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2. Limited use and validation</td>
<td>2. Intrusive to primary task performance</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3. More complex to analyse than uni-dimensional tools</td>
<td></td>
</tr>
<tr>
<td>6. MACE - Malvern Capacity Estimate</td>
<td>Uni-dimensional subjective rating tool</td>
<td>Air Traffic Control (ATC)</td>
<td>No</td>
<td>1. Quick and easy to use, requiring little training or cost</td>
<td>1. Data is collected post-trial. There are a number of problems with this, such as a correlation with performance</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2. Limited evidence of use or reliability and validity</td>
<td>2. Technique has only been used in simulators and would be very intrusive if applied on-line during task performance in the ‘real-world’</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3. Technique has only been used in simulators and would be very intrusive if applied on-line during task performance in the ‘real-world’</td>
<td></td>
</tr>
<tr>
<td>7. Workload Profile Technique</td>
<td>Multi-dimensional subjective rating tool</td>
<td>Generic</td>
<td>Yes</td>
<td>1. Quick and easy to use, requiring little training cost</td>
<td>1. More complex to analyse than uni-dimensional tools</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2. Based upon sound theoretical underpinning (Multiprocess source theory)</td>
<td>2. Data is collected post-trial. There are a number of problems with this, such as a correlation with performance</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3. More complex than other MWL techniques. Dimensions used in the technique may not be fully understood by participants with limited experience of psychology and human factors</td>
<td></td>
</tr>
<tr>
<td>8. Bedford Scale</td>
<td>Multi-dimensional subjective rating tool</td>
<td>Generic</td>
<td>Yes</td>
<td>1. Quick and easy to use, requiring little training or cost</td>
<td>1. More complex to analyse than uni-dimensional tools</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2. Data is collected post-trial. There are a number of problems with this, such as a correlation with performance</td>
<td>2. Limited evidence of use and validation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3. MWL ratings may correlate with task performance dimensional tools</td>
<td>3. Little evidence of actual use and validation of the technique</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2. Less intrusive than on-line techniques</td>
<td>2. Limited evidence of use and validation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3. Quick, easy to use requiring little training</td>
<td>3. Possible correlation with performance</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4. Low cost means of assessing SA and workload</td>
<td></td>
</tr>
</tbody>
</table>

Workload Assessment Tools (primary data source from Stanton, Salmon, Walker, Baber and Jenkins, 2010)
7.2 CREW AWARENESS RATING SCALE
Multi Role Control Station Simulation
Subjective Workload Evaluation

RE: Utilizing the CARS process ratings scale to assess mental workload relative to communication tasks

Please rate the following questions by marking your response:

**Perception:** How easy, in terms of mental workload required, would you say it was to monitor all relevant communication events and information?

- [ ] Easy … monitoring all relevant communication events and information was straightforward
- [ ] Manageable … monitoring all relevant communication events and information was acceptable
- [ ] Difficult … monitoring all relevant communication events and information was challenging
- [ ] Unmanageable … monitoring all relevant communication events and information was unacceptable

**Comprehension:** How easy, in terms of mental workload, would you say it was to interpret all the available information?

- [ ] Easy … interpreting all the available information was straightforward
- [ ] Manageable … interpreting all the available information was acceptable
- [ ] Difficult … interpreting all the available information was challenging
- [ ] Unmanageable … interpreting all the available information was unacceptable

**Projection:** How easy, in terms of mental workload, would you say it was to anticipate/predict the next step in and from the communication process?

- [ ] Easy … predicting the next step in and from the communication process was straightforward
- [ ] Manageable … predicting the next step in and from the communication process was acceptable
- [ ] Difficult … predicting the next step in and from the communication process was challenging
- [ ] Unmanageable … predicting the next step in and from the communication process was unacceptable

**Intention:** How easy, in terms of mental workload, would you say it was to consider your options and decide upon the best course of action to take after communicating with or reviewing all communication sources (e.g. LRE, MCE, ATC, chat, comm alerts, etc.)?

- [ ] Easy … deciding upon the best course of action was straightforward
- [ ] Manageable … deciding upon the best course of action was acceptable
- [ ] Difficult … deciding upon the best course of action was challenging
- [ ] Unmanageable … deciding upon the best course of action was unacceptable

Please see back for additional questions.
Subjective Situational Awareness Evaluation

RE: Utilizing the CARS content ratings scale to assess situational awareness relative to the various modes of communications

Please rate the following questions by marking your response:

Perception: Please rate your ability to identify critical communication cues during the mission.

__ Easy: able to identify all cues
__ Manageable: could identify most cues
__ Difficult: many cues hard to identify
__ Unmanageable: had substantial problems identifying most cues

Comprehension: Please rate how well you understood the communication cues relative to what was going on during the mission.

__ Easy: fully understood the situation as it unfolded at any time
__ Manageable: understood most aspects of the situation at any time
__ Difficult: had difficulty understanding much of the situation at any time
__ Unmanageable: the situation did not make sense to me at any time

Projection: Please rate how well you could predict the next step in the mission using all the communication sources.

__ Easy: could predict with accuracy what was about to occur at any time for each vehicle
__ Manageable: could make accurate predictions most of the time for each vehicle
__ Difficult: misunderstood the situation most of the time for each vehicle
__ Unmanageable: unable to predict what was about to occur with each vehicle

Intention: Please rate how well you were able to best achieve your goals during the mission with the sources of communication available.

__ Easy: very aware of how to achieve your goals at all times using the communication sources
__ Manageable: fairly aware most of the time how to achieve your goals using the communication sources
__ Difficult: somewhat unaware of how to achieve some goals using the communication sources
__ Unmanageable: very unaware of how to achieve the mission goals using the communication sources

Thank you for your participation.
7.3 NASA-TLX SCALE

**NASA Task Load Index**

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

<table>
<thead>
<tr>
<th>Name</th>
<th>Task</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mental Demand</td>
<td>How mentally demanding was the task?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Very Low</td>
<td>Very High</td>
</tr>
<tr>
<td>Physical Demand</td>
<td>How physically demanding was the task?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Very Low</td>
<td>Very High</td>
</tr>
<tr>
<td>Temporal Demand</td>
<td>How hurried or rushed was the pace of the task?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Very Low</td>
<td>Very High</td>
</tr>
<tr>
<td>Performance</td>
<td>How successful were you in accomplishing what you were asked to do?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Perfect</td>
<td>Failure</td>
</tr>
<tr>
<td>Effort</td>
<td>How hard did you have to work to accomplish your level of performance?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Very Low</td>
<td>Very High</td>
</tr>
<tr>
<td>Frustration</td>
<td>How insecure, discouraged, irritated, stressed, and annoyed were you?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Very Low</td>
<td>Very High</td>
</tr>
</tbody>
</table>

Source: [http://humansystems.arc.nasa.gov/groups/TLX/paperpencil.html](http://humansystems.arc.nasa.gov/groups/TLX/paperpencil.html)
7.3 PHASE 2-DATA COLLECTION

7.3.1 CARS Data Set for n=2 for the Increased Communication with Task Interruption Trials

CARS MWL Ratings for 2-CAP Scenario with Increased Comms and Tasks Interruptions

Dimensions

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Perception</th>
<th>Comprehension</th>
<th>Projection</th>
<th>Intention</th>
</tr>
</thead>
<tbody>
<tr>
<td>Novice 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Novice 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Novice 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Novice 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

7.3.2 CARS Data Set for n=5 for the Baseline Trials

CARS MWL Ratings for 5-CAP Baseline Scenario

Dimensions

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Perception</th>
<th>Comprehension</th>
<th>Projection</th>
<th>Intention</th>
</tr>
</thead>
<tbody>
<tr>
<td>Novice 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Novice 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Novice 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Novice 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

"Distribution A: Approved for public release; distribution unlimited. 88ABW Cleared 12/03/2013; 88ABW-2013-5056."
7.3.3 CARS Data Set for n=5 for the Increased Communication Trials

CARS MWL Ratings for 5-CAP Scenario with Increased Comms

7.3.4 CARS Data Set for n=5 for the Task Interruptions Trials

CARS MWL Ratings for 5-CAP Scenario with Task Interruptions

"Distribution A: Approved for public release; distribution unlimited. 88ABW Cleared 12/03/2013; 88ABW-2013-5056."
7.3.5 CARS Data Set for n=5 for the Increased Communications with Task Interruptions Trials

CARS MWL Ratings for 5-CAP Scenario with Increased Comms and Task Interruptions

7.3.6 NASA-TLX Data Set for n=2

Mean NASA-TLX Ratings for the 2-Vehicle Scenario with Increased Comms and Tasks

"Distribution A: Approved for public release; distribution unlimited. 88ABW Cleared 12/03/2013; 88ABW-2013-5056."
7.3.7 NASA-TLX Data Set for n=5

Mean NASA-TLX Ratings for each 5-Vehicle Scenario

<table>
<thead>
<tr>
<th>Subjective Rating</th>
<th>Baseline</th>
<th>Increased Comms</th>
<th>Task Interruptions</th>
<th>Increased Comms &amp; Tasks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mental Demand</td>
<td>28.75</td>
<td>33.75</td>
<td>48.75</td>
<td>31.25</td>
</tr>
<tr>
<td>Physical Demand</td>
<td>6.25</td>
<td>6.25</td>
<td>6.25</td>
<td>6.25</td>
</tr>
<tr>
<td>Temporal Demand</td>
<td>25</td>
<td>30</td>
<td>45</td>
<td>30</td>
</tr>
<tr>
<td>Effort</td>
<td>21.25</td>
<td>20</td>
<td>28.75</td>
<td>23.75</td>
</tr>
<tr>
<td>Performance</td>
<td>25</td>
<td>32.5</td>
<td>32.5</td>
<td>37.5</td>
</tr>
<tr>
<td>Frustration Level</td>
<td>17.5</td>
<td>28.75</td>
<td>33.75</td>
<td>32.5</td>
</tr>
<tr>
<td>Total MWL</td>
<td>20.6241667</td>
<td>25.2075000</td>
<td>32.4991667</td>
<td>26.8741667</td>
</tr>
</tbody>
</table>

"Distribution A: Approved for public release; distribution unlimited. 88ABW Cleared 12/03/2013; 88ABW-2013-5056."
7.3.8 Pupil Diameter Data for the Baseline Trials

![Participant 1 Baseline Trial](image1)

![Participant 2 Baseline Trial](image2)

"Distribution A: Approved for public release; distribution unlimited. 88ABW Cleared 12/03/2013; 88ABW-2013-5056."
Participant 3_Baseline Trial

Participant 4_Baseline Trial

"Distribution A: Approved for public release; distribution unlimited. 88ABW Cleared 12/03/2013; 88ABW-2013-5056."
7.3.9 Pupil Diameter Data for the Increased Communication Trials

![Participant 1_Increased Communication Trial](image1)

![Participant 2_Increased Communication Trial](image2)

"Distribution A: Approved for public release; distribution unlimited. 88ABW Cleared 12/03/2013; 88ABW-2013-5056."
Participant 3_Increased Communication Trial

Participant 4_Increased Communication Trial

"Distribution A: Approved for public release; distribution unlimited. 88ABW Cleared 12/03/2013; 88ABW-2013-5056."
7.3.10 Pupil Diameter Data for the Task Interruption Trials

Participant 1_Task Interruption Trial

Participant 2_Task Interruption Trial

"Distribution A: Approved for public release; distribution unlimited. 88ABW Cleared 12/03/2013; 88ABW-2013-5056."
Participant 3_Task Interruption
Trial

Participant 4_Task Interruption
Trial

"Distribution A: Approved for public release; distribution unlimited. 88ABW Cleared 12/03/2013; 88ABW-2013-5056."
7.3.11 Pupil Diameter Data for the Increased Communication with Task Interruptions Trials

![Participant 1]  
Participant 1_Increased Comms with Task Interruptions Trial

![Participant 2]  
Participant 2_Increased Comms with Task Interruptions

"Distribution A: Approved for public release; distribution unlimited. 88ABW Cleared 12/03/2013; 88ABW-2013-5056."
Participant 3 _Increased Comms with Task Interruptions Trial

Participant 4 _Increased Comms with Task Interruptions

"Distribution A: Approved for public release; distribution unlimited. 88ABW Cleared 12/03/2013; 88ABW-2013-5056."
7.3.12 Fixation Data for the Baseline Trials

Participant 1 _ Baseline Trial

![Graph showing fixation duration for Participant 1](image)

Participant 2 _ Baseline Trial

![Graph showing fixation duration for Participant 2](image)

"Distribution A: Approved for public release; distribution unlimited. 88ABW Cleared 12/03/2013; 88ABW-2013-5056."
"Distribution A: Approved for public release; distribution unlimited. 88ABW Cleared 12/03/2013; 88ABW-2013-5056."
7.3.13 Fixation Data for the Increased Communication Trials

**Participant 1** _Increased Communications Trial_

![Graph showing fixation data for Participant 1.](image)

**Participant 2** _Increased Communication Trial_

![Graph showing fixation data for Participant 2.](image)
Participant 3_ Increased Communications Trial

Participant 4_ Increased Communications Trial

"Distribution A: Approved for public release; distribution unlimited. 88ABW Cleared 12/03/2013; 88ABW-2013-5056."
7.3.14 Fixation Data for the Task Interruption Trials

**Participant 1_ Task Interruption Trial**

**Participant 2_ Task Interruption Trial**

"Distribution A: Approved for public release; distribution unlimited. 88ABW Cleared 12/03/2013; 88ABW-2013-5056."
Participant 3_Task Interruption
Trial

Participant 4_Task Interruption
Trial

"Distribution A: Approved for public release; distribution unlimited. 88ABW Cleared 12/03/2013; 88ABW-2013-5056."
7.3.15 Fixation Data for the Increased Communications with Task Interruptions Trials
Participant 3_Increased Comms with Task Interruptions Trial

Participant 4_Increased Comms with Task Interruptions trial

"Distribution A: Approved for public release; distribution unlimited. 88ABW Cleared 12/03/2013; 88ABW-2013-5056."
7.4 PHASE 2 - ANALYSIS

7.4.1 Total Mental Workload Comparison between 2- vs. 5-Vehicle Experimental Design for the Increased Communications with Task Interruptions Scenario
7.4.2 CARS Means Comparison of Total MWL by Scenario

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F Ratio</th>
<th>Prob &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>6</td>
<td>0.875</td>
<td>0.145</td>
<td>0.5283</td>
<td>.</td>
</tr>
<tr>
<td>Error</td>
<td>9</td>
<td>2.484</td>
<td>0.276</td>
<td></td>
<td>Prob &gt; F</td>
</tr>
<tr>
<td>C. Total</td>
<td>15</td>
<td>3.393</td>
<td></td>
<td>0.7747</td>
<td>.</td>
</tr>
</tbody>
</table>

**Task Interruptions**

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F Ratio</th>
<th>Prob &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.5000</td>
<td>1</td>
<td>0.250</td>
<td>0.091</td>
<td>0.3501</td>
<td>.</td>
</tr>
<tr>
<td>1.4000</td>
<td>1</td>
<td>0.062</td>
<td>0.022</td>
<td>0.6455</td>
<td>.</td>
</tr>
</tbody>
</table>

**Comms**

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F Ratio</th>
<th>Prob &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.4000</td>
<td>1</td>
<td>0.146</td>
<td>0.146</td>
<td>0.4695</td>
<td>.</td>
</tr>
<tr>
<td>1.4000</td>
<td>3</td>
<td>0.218</td>
<td>0.080</td>
<td>0.6866</td>
<td>.</td>
</tr>
</tbody>
</table>

"Distribution A: Approved for public release; distribution unlimited. 88ABW Cleared 12/03/2013; 88ABW-2013-5056."
"Distribution A: Approved for public release; distribution unlimited. 88ABW Cleared 12/03/2013; 88ABW-2013-5056."
Graph Builder

Residual Total MWL vs. Predicted Total MWL

"Distribution A: Approved for public release; distribution unlimited. 88ABW Cleared 12/03/2013; 88ABW-2013-5056."
### 7.4.3 NASA-TLX Means Comparison of Total MWL by Scenario

#### Response Total MWL

<table>
<thead>
<tr>
<th>Summary of Fit</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>R Square</td>
<td>0.339902</td>
</tr>
<tr>
<td>R Square Adj</td>
<td>-0.10016</td>
</tr>
<tr>
<td>Root Mean Square Error</td>
<td>18.84181</td>
</tr>
<tr>
<td>Mean of Response</td>
<td>26.30125</td>
</tr>
<tr>
<td>Observations (or Sum Wgts)</td>
<td>16</td>
</tr>
</tbody>
</table>

#### Analysis of Variance

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Ratio</th>
<th>Prob &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>5</td>
<td>1645.2538</td>
<td>274.209</td>
<td>0.7724</td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td>9</td>
<td>3195.1231</td>
<td>355.014</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. Total</td>
<td>15</td>
<td>4840.3769</td>
<td></td>
<td>0.6107</td>
<td></td>
</tr>
</tbody>
</table>

#### Parameter Estimates

<table>
<thead>
<tr>
<th>Source</th>
<th>Nparm</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>F Ratio</th>
<th>Prob &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task Interruptions</td>
<td>1</td>
<td>1</td>
<td>103.3722</td>
<td>0.5135</td>
<td>0.4906</td>
</tr>
<tr>
<td>Comms</td>
<td>1</td>
<td>1</td>
<td>1.0847</td>
<td>0.0031</td>
<td>0.9571</td>
</tr>
<tr>
<td>Task Interruptions*Comms</td>
<td>1</td>
<td>1</td>
<td>104.2033</td>
<td>0.2935</td>
<td>0.6011</td>
</tr>
<tr>
<td>Subject</td>
<td>3</td>
<td>3</td>
<td>1356.5936</td>
<td>1.2737</td>
<td>0.3400</td>
</tr>
</tbody>
</table>

"Distribution A: Approved for public release; distribution unlimited. 88ABW Cleared 12/03/2013; 88ABW-2013-5056."
<table>
<thead>
<tr>
<th>Level</th>
<th>Sq Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>29.596625</td>
</tr>
<tr>
<td>0</td>
<td>22.915875</td>
</tr>
</tbody>
</table>

Levels not connected by the same letter are significantly different.

### Comms - Task Interruptions Comms

<table>
<thead>
<tr>
<th>Level</th>
<th>Sq Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>25.561625</td>
</tr>
<tr>
<td>1</td>
<td>26.040875</td>
</tr>
</tbody>
</table>

Levels not connected by the same letter are significantly different.

### Subject

<table>
<thead>
<tr>
<th>Level</th>
<th>Sq Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>32.499000</td>
</tr>
<tr>
<td>1.1</td>
<td>26.874250</td>
</tr>
<tr>
<td>0.1</td>
<td>25.207500</td>
</tr>
<tr>
<td>0.0</td>
<td>20.624250</td>
</tr>
</tbody>
</table>

Levels not connected by the same letter are significantly different.

"Distribution A: Approved for public release; distribution unlimited. 88ABW Cleared 12/03/2013; 88ABW-2013-5056."
"Distribution A: Approved for public release; distribution unlimited. 88ABW Cleared 12/03/2013; 88ABW-2013-5056."
"Distribution A: Approved for public release; distribution unlimited. 88ABW Cleared 12/03/2013; 88ABW-2013-5056."
7.4.4 MRT Resource Mapping

**MRT Mapping for Auditory Resources**

- Baseline: ~4 MRT Rating
- Increased Comms: ~4 MRT Rating
- Task Interruptions: ~3 MRT Rating
- Increased Comms + Tasks: ~3 MRT Rating

**MRT Mapping for Cognitive Resources**

- Baseline: ~1 MRT Rating
- Increased Comms: ~2 MRT Rating
- Task Interruptions: ~2 MRT Rating
- Increased Comms + Tasks: ~2 MRT Rating

"Distribution A: Approved for public release; distribution unlimited. 88ABW Cleared 12/03/2013; 88ABW-2013-5056."
MRT Mapping for Visual Resources

MRT Mapping for Speech Resources

"Distribution A: Approved for public release; distribution unlimited. 88ABW Cleared 12/03/2013; 88ABW-2013-5056."
MRT Mapping for Motor Resources

<table>
<thead>
<tr>
<th>Scenario</th>
<th>MRT Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>4.5±0.5</td>
</tr>
<tr>
<td>Increased Comms</td>
<td>4.5±0.5</td>
</tr>
<tr>
<td>Task Interruptions</td>
<td>5.0±0.5</td>
</tr>
<tr>
<td>Increased Comms + Tasks</td>
<td>4.5±0.5</td>
</tr>
</tbody>
</table>
8.0 BIBLIOGRAPHY


"Distribution A: Approved for public release; distribution unlimited. 88ABW Cleared 12/03/2013; 88ABW-2013-5056."


"Distribution A: Approved for public release; distribution unlimited. 88ABW Cleared 12/03/2013; 88ABW-2013-5056."

"Distribution A: Approved for public release; distribution unlimited. 88ABW Cleared 12/03/2013; 88ABW-2013-5056."