Dynamic Human-Computer Collaboration in Real-time Unmanned Vehicle Scheduling

by

Andrew S. Clare

S.B. Aerospace Engineering with Information Technology
Massachusetts Institute of Technology, Cambridge, MA, 2008

Submitted to the Department of Aeronautics and Astronautics
in partial fulfillment of the requirements for the degree of

Master of Science in Aeronautics and Astronautics
at the
MASSACHUSETTS INSTITUTE OT TECHNOLOGY

June 2010

Author ....
Department of Aeronautics and Astronautics
May 14, 2010

Certified by ....
Mary L. Cummings
Associate Professor of Aeronautics and Astronautics
Thesis Supervisor

Accepted by ....
Eytan H. Modiano
Associate Professor of Aeronautics and Astronautics
Chair, Committee on Graduate Students

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
JUN 2010		**00-00-2010 to 00-00-2010**

4. TITLE AND SUBTITLE
Dynamic Human-Computer Collaboration in Real-time Unmanned Vehicle Scheduling

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Massachusetts Institute of Technology,Department of Aeronautics and Astronautics,Cambridge,MA,02139

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
see report

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**Same as Report (SAR)**	**127**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

Dynamic Human-Computer Collaboration in Real-time Unmanned Vehicle Scheduling

by

Andrew S. Clare

Submitted to the Department of Aeronautics and Astronautics
on May 14th, 2010 in partial fulfillment of the requirements for the degree of
Master of Science in Aeronautics and Astronautics.

Abstract

Advances in autonomy have made it possible to invert the operator-to-vehicle ratio so that a single operator can control multiple heterogeneous Unmanned Vehicles (UVs). This autonomy will reduce the need for the operator to manually control each vehicle, enabling the operator to focus on higher-level goal setting and decision-making. Computer optimization algorithms that can be used in UV path-planning and task allocation usually have an a priori coded objective function that only takes into account pre-determined variables with set weightings. Due to the complex, time-critical, and dynamic nature of command and control missions, brittleness due to a static objective function could cause higher workload as the operator manages the automation. Increased workload during critical decision-making could lead to lower system performance which, in turn, could result in a mission or life-critical failure.

This research proposes a method of collaborative multiple UV control that enables operators to dynamically modify the weightings within the objective function of an automated planner during a mission. After a review of function allocation literature, an appropriate taxonomy was used to evaluate the likely impact of human interaction with a dynamic objective function. This analysis revealed a potential reduction in the number of cognitive steps required to evaluate and select a plan, by aligning the objectives of the operator with the automated planner.

A multiple UV simulation testbed was modified to provide two types of dynamic objective functions. The operator could either choose one quantity or choose any combination of equally weighted quantities for the automated planner to use in evaluating mission plans. To compare the performance and workload of operators using these dynamic objective functions against operators using a static objective function, an experiment was conducted where 30 participants performed UV missions in a synthetic environment. Two scenarios were designed, one in which the Rules of Engagement (ROEs) remained the same throughout the scenario and one in which the ROEs changed.

The experimental results showed that operators rated their performance and confidence highest when using the dynamic objective function with multiple objectives. Allowing the operator to choose multiple objectives resulted in fewer modifications to the objective function, enhanced situational awareness (SA), and increased spare mental capacity. Limiting the operator to choosing a single objective for the automated planner led to superior performance for individual mission goals such as finding new targets, while also causing some violations of ROEs, such as destroying a target without permission. Although there were no significant differences in system performance or workload between the dynamic and static objective functions, operators had superior performance and higher SA during the mission with changing ROEs. While these results suggest that a dynamic objective function could be beneficial, further research is required to explore the impact of dynamic objective functions and changing mission goals on human performance and workload in multiple UV control.

Thesis Supervisor: Mary L. Cummings
Title: Associate Professor of Aeronautics and Astronautics

Acknowledgments

I owe thanks to numerous people for the successful completion of this thesis.

First, to my advisor Missy Cummings, for taking me on as an undergraduate researcher 4 years ago, believing in my potential as a graduate student, and for all of the guidance and feedback that you provided on my research. You have had a tremendous impact on me, both professionally and personally, and I look forward to further research under your leadership in the future.

To Yves Boussemart, Ryan Castonia, and Kristopher Thornburg for providing feedback on this thesis. Your comments and advice were outstanding and truly helped me throughout the writing process.

To the National Defense Science and Engineering Graduate Fellowship, through the Air Force Office of Scientific Research, for funding my graduate education, and the Office of Naval Research for supporting the development of the testbed used for this experiment.

To Christin, Dan, Scott, Ian, and Pierre, for all of your help in the development of the OPS-USERS Interface. I learned so much from all of you, had a great time (even when the system crashed countless times), and I couldn't have asked for better teammates on this project.

To Olivier Toupet and Karl Kulling at Aurora Flight Sciences and Andy Whitten, Cameron Frasier, and Jon How at the Aerospace Controls Laboratory, for all of your help in the development of the OPS-USERS system. I couldn't have done this project without all of your expertise and support.

To my fellow HALiens past and present: Christin, Dan, Dave, Armen, Geoff, Anna, Birsen, Luca, Yves, Ryan, Jason, and many others, thank you for the constant support, especially during the Quals, as well as plenty of good times on and off the volleyball court.

To my brothers in Sigma Phi Epsilon and other friends at MIT through both my undergraduate and graduate years. All of you have had such an impact on me as I've developed over the past six years. I will always remember all of the memories that we had, while I look forward to more good times in the future.

To my Mom, Dad, and brother for your unwavering love and support through all of my years in school. You were the first to believe in me and your encouragement and advice have always lifted my spirits and pushed me to pursue my dreams.

And to Bri, for being my rock through the ups and downs of school. Whenever I got distracted from my work, you kept me motivated and always believed in me. Our time together has been wonderful and I am forever grateful for your love and support.

Table of Contents

List of Figures

List of Tables

List of Acronyms

ANOVA	Analysis of Variance
HACT	Human-Automation Collaboration Taxonomy
LOA	Levels of Automation
LOC	Levels of Collaboration
MABA-MABA	“Men are better at – Machines are better at”
MIT	Massachusetts Institute of Technology
OPS-USERS	Onboard Planning System for UxVs Supporting Expeditionary Reconnaissance and Surveillance
ROE	Rule of Engagement
SA	Situational Awareness
SCT	Schedule Comparison Tool
SRK	Skill, Rule, and Knowledge-based
TLAM	Tomahawk Land Attack Missiles
UAV	Unmanned Aerial Vehicle
USV	Unmanned Surface Vehicle
UV	Unmanned Vehicle
WUAV	Weaponized Unmanned Aerial Vehicle

1 Introduction

1.1 Motivation

In the past decade, the use of Unmanned Vehicles (UVs) has increased dramatically for scientific, military, and civilian purposes. UVs have been successfully used in dangerous and remote environments, with Underwater Unmanned Vehicles exploring the deepest trenches of the ocean (e.g., [1]) and NASA's rovers traversing the surface of Mars [2]. Unmanned Aerial Vehicles (UAVs) have enabled the military to conduct long duration missions over hostile territory without placing a pilot in harm's way. Unmanned Ground Vehicles have been utilized by soldiers and civilian bomb squads to investigate and defuse explosive devices (e.g., [3]). Scientists have studied global warming by surveying the polar ice caps (e.g., [4]) with UAVs, while civilian agencies have employed UAVs for border patrol [5] and forest firefighting [6].

While these UVs contain advanced technology, they typically require multiple human operators, often many more than a comparable manned vehicle would require. This barrier to further progress in the use of UVs can be overcome through an increase in the autonomous capabilities of UVs [7]. Many advanced UVs can execute basic operational and navigational tasks autonomously and can collaborate with other UVs to complete higher level tasks, such as surveying a designated area [8, 9]. The United States Department of Defense already envisions inverting the operator-to-vehicle ratio in future scenarios where a single operator controls multiple UAVs simultaneously [10]. This concept has been extended to single operator control of multiple heterogeneous (air, sea, land) UVs [11], as illustrated in Figure 1.

In this concept of operations, a single operator will supervise multiple vehicles, providing high level direction to achieve mission goals, and will need to comprehend a large amount of information while under time-pressure to make effective decisions in a dynamic environment.

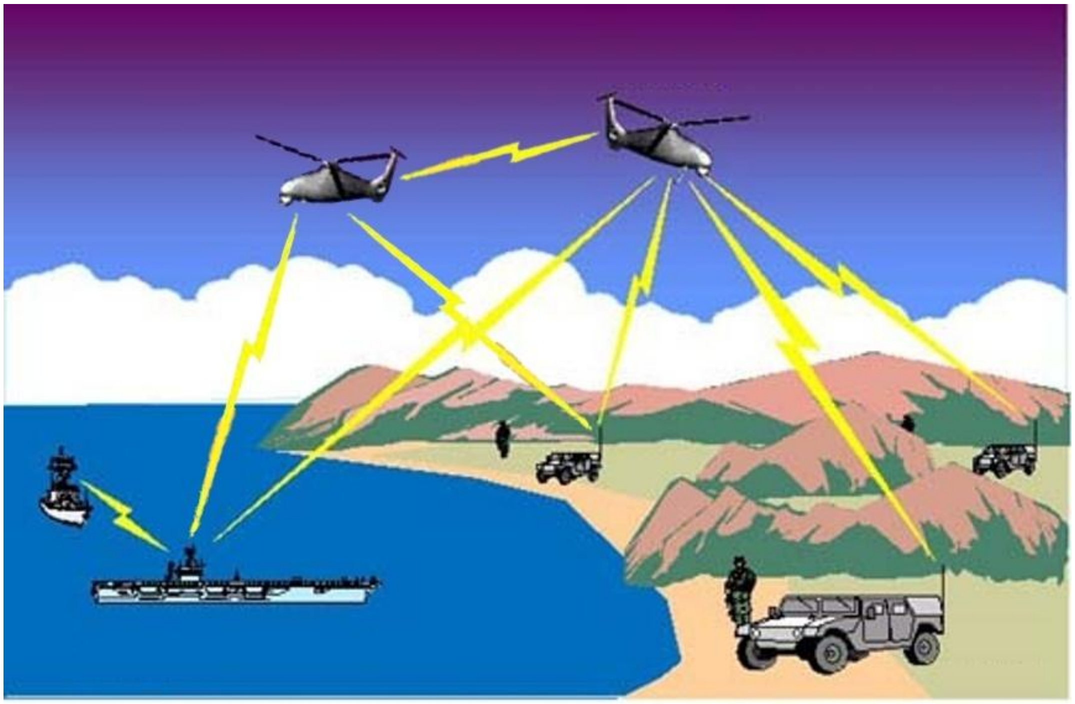


Figure 1. Coordinated Operations with Heterogeneous Unmanned Vehicles [12].

This large amount of data provides a challenge for system designers, as it may cause cognitive saturation, which has been shown to correlate with poor performance [13, 14]. The capacity of a single operator to control multiple UVs has been demonstrated in multiple studies [15, 16]. Operators will be assisted by automated planners, which can be faster and more accurate than humans at path planning [17] and task allocation [18] in a multivariate, dynamic, time-pressured environment.

Outside of the world of UV control, path planning with the assistance of automated planners has become routine, with the proliferation of Global Positioning Systems on mobile devices and in automobile navigation systems, as well as advances in online route planners such as MapQuest© and Google Maps©. While extensive research has been conducted in the computer science field to develop better algorithms for planning, comparatively little research has occurred on the methods by which human users utilize these tools, especially when working in dynamic, time-critical situations with high uncertainty in information [19].

Human management of the automated planner is crucial, as automated planners do not always generate accurate solutions, especially in the presence of unknown variables and possibly inaccurate prior information. Though fast and able to handle complex computation far better than humans, computer optimization algorithms are notoriously "brittle" in that they can only take into account those quantifiable variables identified in the design stages that were deemed to be critical [20, 21]. In a command and control situation such as supervising multiple UVs, where events are often unanticipated, automated planners are unable to account for and respond to unforeseen problems [22, 23]. Additionally, operators can become confused when working with automation, unaware of how the "black box" automated planner came to its solution. Various methods of human-computer collaboration have been investigated to address the inherent brittleness and opacity of computer algorithms [19, 21, 24, 25]. To truly assist human supervisors of multiple UVs, however, automated planners must be capable of *dynamic* mission replanning. As vehicles move, new tasks emerge, and mission needs shift, the way that the automated planner works will need to change to assist in real-time decision making. This will require greater flexibility and transparency in the computer algorithms designed for supporting multi-UV missions.

This thesis will investigate the impact of human-computer collaboration in the context of dynamic objective function manipulation for multiple UV control. Computer optimization algorithms, such as those used in most automated path planning and task allocation problems, typically have an a priori coded objective function that only takes into account pre-determined variables with set weightings. In this work, human operators will be given the ability to modify the weightings of these optimization variables during a mission. One significant concern in this concept of operations where one operator supervises multiple UVs is the potential high workload for the operator, and possible negative performance consequences. This work will investigate the operator workload and both human and system performance implications of providing this additional level of human-computer collaboration.

1.2 Problem Statement

To effectively supervise multiple UVs simultaneously, operators will need the support of significant embedded collaborative autonomy. This autonomy will reduce the need for the operator to manually control each vehicle, enabling the operator to focus on higher-level goal setting and decision-making. Automated planners can conduct path planning and scheduling faster and possibly more efficiently than humans. Due to the complexity and dynamic nature of command and control missions, however, the brittleness of automated planners could cause overall lower system performance or higher workload as the operator manages the automation. This thesis seeks to determine how best to divide responsibility for mission replanning in a dynamic environment between the human and automation, with the ability to designate degrees of collaboration. Additionally, this thesis seeks to evaluate whether there is a difference in system performance when a human operator controlling multiple, heterogeneous UVs collaborates with an automated planner that has a static objective function or a dynamic objective function that can be modified during the mission.

1.3 Research Objectives

To address this goal, the following research objectives were posed:

- Objective 1: **Determine the motivating principles for dynamic objective function manipulation in human-computer collaborative multi-UV control.** In order to achieve this objective, current research in human-computer collaboration for scheduling, resource allocation, and path planning was reviewed, as described in

Chapter 2. Also, a theoretical model of dynamic objective function manipulation was developed, as outlined in Chapter 3.

- Objective 2: **Develop a tool to enable operators to dynamically modify the objective function of an automated planner**. From the motivating principles described in Objective 1, as well as mission-specific information, a dynamic objective function tool was designed, described in Chapter 4. This tool was integrated into the Onboard Planning System for UxVs Supporting Expeditionary Reconnaissance and Surveillance (OPS-USERS), a previously developed multi-UV mission simulation testbed for evaluating the impact of embedded autonomy distributed across networked UVs [18, 26].
- Objective 3: **Evaluate the effectiveness of real-time human manipulation of objective function in multi-UV scheduling algorithms**. To address this objective, human performance experimentation (Chapters 4 and 5) was conducted to analyze how well the dynamic objective function tool is able to support single operator multi-UV control compared to an automated planner with a static objective function.

1.4 Thesis Organization

This thesis is organized into the following chapters:

- Chapter 1, Introduction, describes the motivation and research objectives of this thesis.
- Chapter 2, Background, provides a summary of a previous experiment that motivated this thesis, discusses current human-computer collaboration research, and frames the context of the research objectives introduced in Chapter 1.
- Chapter 3, Human-Automation Role Allocation, provides an overview of function allocation literature to discuss methods for dividing responsibility for mission replanning

in a dynamic environment between the human and automation. A theoretical model of function allocation is applied to the chosen simulation testbed. This model is extended to incorporate dynamic objective function manipulation and to describe the potential benefits of a dynamic objective function tool.

- Chapter 4, Human Performance Experimentation, describes the human-performance experiment used to test the hypotheses of this research. Details include a discussion of the interfaces designed to enable manipulation of the objective function of an automated planner for multiple UV control, objectives of the experiment, participants, procedures, and experimental design.
- Chapter 5, Results, presents the statistical results of the experiment from Chapter 4.
- Chapter 6, Discussion, compares the results of the human performance experiment with the hypotheses.
- Chapter 7, Conclusions, summarizes the motivation and objectives of this research, how well the objectives were met, and the key contributions. Suggestions for future work are also provided.

2 Background

This chapter discusses previous research relevant to human supervisory control of multiple UVs with the support of an automated planning algorithm. Previous experimental work on human-automation collaboration for scheduling, path planning, and task allocation is described to detail both the benefits and drawbacks of collaboration with automated planners. Through this initial research, three gaps in previous methods of human collaboration with automated planners were revealed: methods for dealing with dynamic and uncertain environments, decision-making support under time-pressure on the order of seconds, and methods for operators to align the objective function of the automated planner with their desires. These gaps can be addressed through the development of a dynamic objective function method for collaborative human-automation control of multiple UVs.

2.1 Motivating Experiment

In a previous experiment, human operators used a simulation environment to supervise multiple UVs with the assistance of a decentralized automated planner with a static objective function [18]. This system was utilized to examine the impact of increasing automation replanning rates on operator performance and workload [13]. The operator was prompted to replan at various intervals, but could also choose to replan whenever he or she desired. When replanning, the operator could accept, reject, or attempt to modify automation-generated plans manually.

Results showed that the rate of replanning by the human operator had a significant impact on workload and performance. Specifically, rapid replanning caused high operator workload, which resulted in poorer overall system performance [13]. Workload was characterized through a utilization metric, which measured percent busy time. Results from the experiment also showed that operators with the ability to collaborate effectively with the automated planner, labeled “Consenters” in the study, had significantly higher performance and lower workload [27].

Surveys conducted after each trial revealed that approximately 35% of the participants were frustrated by the automated planner. Participants wrote or stated that they did not always understand what the automated planner was doing. A few participants specifically wrote that they desired the ability to modify the way that the automated planner worked. For example, participants wrote “automation [is] not very smart, [and] doesn't have same priorities I do,” and “the algorithm does its own thing most of the time…there was a clash between what I wanted to have the UAVs do and what the [algorithm] decided” [13]. Operators were unable to express their desires to the automated planner, which was too brittle for the dynamic environment and mission. This thesis seeks to address this shortcoming by developing a method for dynamic objective function manipulation, which should enable operators to more effectively collaborate with an automated planner for multi-UV control.

2.2 Human-Automation Collaboration Empirical Research

This section outlines experiments which have explored the ability of humans to collaborate closely with an automated planner for a path-planning, scheduling, or resource allocation problem. These experiments show previous attempts to develop systems that address the communication gap between humans and the automated systems. Human-automation collaboration can be beneficial due to the uncertainty inherent in supervisory control systems, such as weather, target movement, changing priorities, etc. Numerous previous experiments have shown the benefits of human-guided algorithms for search, such as in vehicle-routing problems [28-30] or trade space exploration for large scale design optimization [31]. However, the inability of the human to understand the method by which the automation developed its solution, or whether a solution is optimal, especially in time-pressured situations, can lead to automation bias [32]. This automation bias can cause complacency, degradation in skills and performance, and potential loss of Situational Awareness (SA) [15].

Many researchers have found success in addressing challenging scheduling problems using mixed-initiative systems, where a human guides a computer algorithm in a collaborative process to solve a problem. The "initiative" in such systems is shared in that both the human and computer can independently contribute to the formulation and analysis of solutions [33]. For example, a mixed-initiative tool to solve an over-constrained scheduling problem could provide operators with the ability to relax constraints for a sensitivity analysis. This is essentially a “what-if” tool to compare the results of changes made to the schedule [34]. Scott, Lesh, and Klau showed that in experiments with humans utilizing mixed-initiative systems for vehicle routing, operator intervention can lead to better results, but there is variation in the way that operators interact with the system and in their success in working with the automation [29]. Howe et al. developed a mixed initiative scheduler for the U.S. Air Force satellite control network, implementing a satisficing algorithm, which recommends plans despite the fact that a solution that satisfies all constraints does not exist [35]. The user can choose the “best” plan despite constraint violations and modify the plan to address mistakes and allow for emergency high priority requests. The authors argued that it was difficult to express the complete objective function of a human through an a priori coded objective function because of the likely non-linear evaluations made by the human and the unavailability of all information necessary for the algorithm to make a decision [35].

Hanson et al. found that human operators paired with an algorithm for scheduling multiple UAVs desired a greater understanding of why the algorithm made certain recommendations [36]. The authors also observed that operators tend to think less in terms of numerical optimization when planning UAV routes, but in abstract terms about the overall goals or tactical objectives that they want to accomplish. The authors argue that developing a method to communicate these goals to the optimization algorithm would help the user develop increased trust in the automation and result in solutions that match the desires of the operator. Miller, et al. attempted to address this challenge through the development of the Playbook™ human-automation integration architecture, which identified a set of common tasks performed by semi-autonomous UVs, grouped them into "plays," and provided the operator with a set of play templates to utilize [37]. This system limited the human operators' interactions with the automation to selecting pre-made plays instead of directly communicating their desires to the automated planner. Although this method worked successfully in an experimental setting, it may be too limiting for the highly complex, dynamic, and uncertain environments found in command and control missions.

Much of this previous research focused on methods for humans to work with automation to solve a problem, such as changing the inputs to the algorithm. Comparatively little research has investigated methods by which the human operator could, in real-time, change the way that the automation actually *works* in order to aid in accomplishing mission objectives. Techniques for guiding optimization algorithms, for changing the constraints, and for modifying solutions developed by an algorithm were all described in detail. There was a constant assumption, however, that the automation was static and unchanging throughout the period in which the human was interacting with the automation. Despite enhanced collaboration, operator SA was low and operators complained about the lack of transparency in how the automation generated plans [13, 19, 35, 36]. For example, Marquez concluded that an improvement to her collaborative lunar path planning aid would be adding additional flexibility, stating that users should "have the ability to change the cost function (variables or relationships) and observe how the [solution] itself changes based on the cost function modifications" [19]. Thus, developing a method for human operators to modify the objective function of the automated planner in real-time could provide the transparency necessary to maintain operator SA, while enabling operators to communicate their desires to the automation.

More recent research has increased the focus on the concept of providing the human operator with the ability to modify the way the automated planner works for collaborative decision-making. Bruni and Cummings developed a series of studies on human interactions with an automated planner for mission planning with Tomahawk Land Attack Missiles (TLAM) [38, 39]. In their experiment, human planners paired missiles, which could come from different launchers, with preplanned missions or targets. This was a highly complex optimization problem, where operators needed to consider many pieces of information. One of the interfaces tested in the experiment featured a customizable heuristic search algorithm, where the human operator could choose and rank criteria that would adjust the weights of variables in the objective function. The authors emphasized that while heuristic algorithms are fast and will generally find a solution if one exists, the algorithms provide no guarantee of finding the "best" solution, as the algorithm can become stuck in local optima. The interface also allowed the human operator to manually adjust the solution after utilizing the heuristic search algorithm to develop an initial solution. Results showed that there was no statistical difference in performance between this method of collaborative human-automation planning as compared to a more manual method of planning. In terms of the number of information processing steps required to generate a solution, which relates directly to operator workload [40], the collaborative interface utilizing the customizable search algorithm required significantly fewer steps than the manual interface. Although lower workload was achieved, the mission was not time-critical on the order of seconds (despite the fact that subjects were timed) and was not performed in a real-time, dynamic environment.

Finally, Forest et al. conducted an experiment during which operators created a schedule for multiple UAVs with the assistance of a human-guided algorithm [25]. The subjects were presented with different interfaces to pre-plan a mission based on pre-existing targets with given values and risks. Certain interfaces had sliding bars that enabled the operator to modify the weights on the five factors that the objective function used to calculate scores for the plans: total target value, risk, percentage of available missiles used (utilization), distance, and mission time. Although the operator could utilize any of these factors to evaluate plans, the mission instructions encouraged operators to maximize target value while minimizing mission time.

Results showed that, based purely on mission time and target value, the "best" plans were created in an interface where the human operator did not have the ability to modify the objective function of the automated planner [25, 41]. The authors concluded that it was likely that operators chose plans based on a number of additional factors, including risk or distance metrics. Discussions with participants after the experiment confirmed that they determined their own risk tolerances and included metrics beyond just time and target value in their selection of plans. These results show that while automation is excellent at optimizing a solution for specific goals, automation may be too brittle to take into account all factors that could influence the success of a complex command and control mission in an uncertain environment.

This experiment highlighted the difficulty of human-automation collaboration when humans have different internal objective functions from the automation. In subjective ratings, participants gave the highest rating to the interface where they had the most control of the objective function [41]. They found it intuitive to adjust the weights and had higher trust in the automation's solution. It should be noted that these results were obtained for a pre-planning scenario, where algorithm searches took 20-30 seconds, and the entire planning process could take up to 15 minutes. While these experiments show that dynamic objective functions can result in improved collaboration between humans and automation, only six participants were involved in the study.

2.3 Summary

In summary, previous research has shown that humans and automation can collaborate to achieve superior results in resource allocation and path planning problems, with potentially lower workload. These results have also demonstrated the need for better methods for human operators to express their internal objectives and desires to automated planners.

Three key gaps have been identified in the experimental research reviewed here. First, most of the previous experiments in human-automation collaboration occurred in fairly static environments with high certainty. Typically, the experiments involved mission pre-planning, where targets were known in advance and information was certain and did not change during the decision-making process. Realistic command and control missions involve highly dynamic and uncertain environments, and collaborative control methods need to be developed that can operate in these environments.

A second gap in the previous literature is the lack of experiments that required users to make decisions under time-pressure. Many of the collaborative systems were developed for pre-planning scenarios, when operators have minutes, hours, or days to make decisions. The algorithms in some of the experiments required seconds, if not minutes, to generate solutions. To account for highly dynamic environments, collaborative control will be necessary during mission replanning. The time scale for decision making will be reduced dramatically, to mere seconds, and previous research indicates that under this type of time-pressure, operators will often change their strategies, including those concerning the use of automation [42, 43]. While these adjustments in strategies for managing the automation may be beneficial, research is needed in human-automation collaborative control in time-pressured environments to understand the strategies of operators under these conditions.

A third gap is the lack of methods for operators to express their desires to the automated planner to ensure alignment of the objective functions of the human and automation. A number of the participants in the experiments reviewed here complained of a mismatch between their own goals and the plans generated by the automated planner. Few attempts have been made to enable operators to change the way the automation works to generate and evaluate plans.

This thesis seeks to address these gaps by investigating the use of objective function weight adjustments as a potential method for enhancing human-automation collaboration in multi-UV control in a highly dynamic, real-time command and control environment. In the following chapter, function allocation literature is reviewed in order to select an appropriate taxonomy to apply in order to evaluate the potential impact of human manipulation of a dynamic objective function. Based on this analysis, dynamic objective functions will be implemented in an existing multiple UV simulation testbed, and a human performance experiment will be used to evaluate the performance and workload implications of the dynamic objective function.

3 Human-Automation Role Allocation

In this chapter, a review of function allocation literature highlights various taxonomies for dividing responsibility between the human operator and automation. These taxonomies are evaluated in order to select an appropriate method for modeling a collaborative human-automation system. In order to evaluate the impact of a dynamic objective function, an existing multiple UV simulation testbed is chosen for human performance experiments. The system is described and then analyzed using the selected taxonomy. The taxonomy is extended to include the proposed method for manipulating the objective function of an automated planner. Finally, the theoretical impact of utilizing a dynamic objective function on human operator workload and system performance is explored.

3.1 Function Allocation Taxonomies

Human-computer collaboration for controlling multiple UVs raises the issue of the determining the appropriate roles of the human operator and automated planner. In the scope of this thesis, an example would be determining the impact of providing the human operator with the role of manipulating the automated planner for collaborative UV control. The field of function allocation has traditionally focused on the question of whether a human *or* computer is better suited to perform a task.

One method of comparing the capabilities of humans and computers is through Rasmussen's Skill, Rule, and Knowledge-based (SRK) taxonomy of cognitive control [44, 45]. Typically, automation is utilized to reduce human workload, for example, by automating skill-based tasks such as controlling the altitude of an airplane or manufacturing a component on an assembly line. As computers have grown more powerful, automation has become more useful in tasks that are cognitively demanding for humans, such as controlling unstable aircraft.

Computers have also been shown to have the ability to plan optimal paths when the environment is known with moderate certainty [46]. Humans, however, have the ability to conduct knowledge-based reasoning [44] because of their superior improvisation, flexibility, and inductive reasoning skills as compared to computers. Computers are typically unable to perform this higher level reasoning because they simply follow a set of predetermined rules, known as rule-based behavior [19]. Although the SRK taxonomy is descriptively useful for classifying tasks into broad categories and enumerating the generalized strengths of humans and computers, it lacks a prescriptive methodology for allocating functions.

One of the first formal treatments of function allocation is known as Fitts List [47]. An example of a Fitts list is shown in Table 1. Fitts and his colleagues aimed to identify those functions or tasks that were performed better by machines or humans. For many years, this paper was regarded as the seminal work in the field of function allocation, despite the fact that the authors noted that their method was highly limiting.

Table 1. Example Fitts List

Attribute	Machine	Human
Speed	Superior	Comparatively slow
Power Output	Superior in level in consistency	Comparatively weak
Consistency	Ideal for consistent, repetitive action	Unreliable, learning & fatigue a factor
Information Capacity	Multi-channel	Primarily single channel
Memory	Ideal for literal reproduction, access restricted and formal	Better for principles & strategies, access versatile & innovative
Reasoning Computation	Deductive, tedious to program, fast & accurate, poor error correction	Inductive, easier to program, slow, accurate, good error correction
Sensing	Good at quantitative assessment, poor at pattern recognition	Wide ranges, multi-function, judgment
Perceiving	Copes with variation poorly, susceptible to noise	Copes with variation better, susceptible to noise

Price argued that Fitts list remains a valuable heuristic aid to design, despite the generalizations and the assumption that a task will be performed solely by humans or machines

[48]. Price, however, asserted that function allocation by formula alone cannot be achieved and that we must rely on expert judgment as the final means of making allocation decisions, based on past experience and empirical tests. He also advocated for an iterative design process instead of the typical one-time step of allocating functions that occurs early in the design of technical systems. Price introduced a decision matrix for function allocation, as shown in Figure 2. This decision matrix rejects the assumption that the choice between human and machine is zero-sum. The six regions shown in Price's decision matrix are: 1) there is no difference in the relative capabilities of human & machine, 2) human performance is clearly superior than machine performance, 3) machine performance is clearly superior to human performance, 4) machine performance is so poor that the functions should be allocated to humans, 5) human performance is so poor that the functions should be allocated to machine, and 6) unacceptable performance by both human and machine. By adding the concept that humans and machines may have comparable or even complementary skills, Price brought the function allocation world closer to the concept of human-automation collaboration.

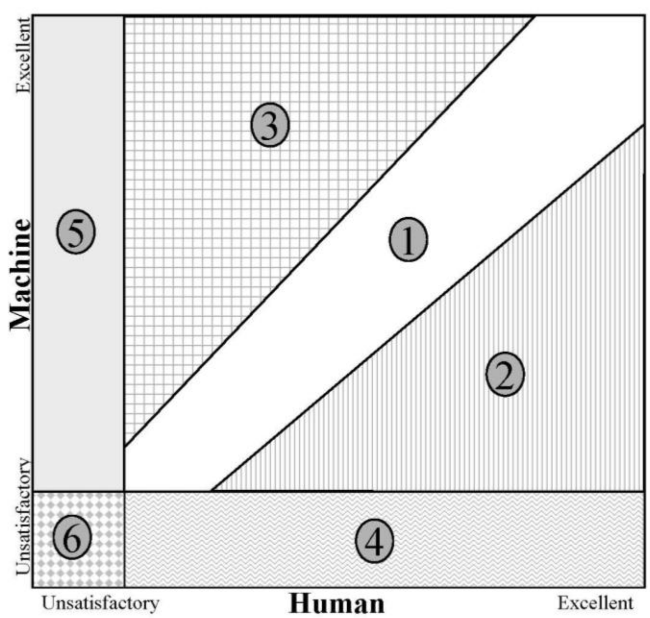


Figure 2. Decision Matrix for Function Allocation [48]

A further attempt to describe the interactions between humans and computers is the Levels of Automation (LOA) scale [49, 50]. Shown in Table 2, the LOA scale describes a human-computer system that ranges from fully manual to fully automatic. At lower LOAs, the human is very active and involved in decision-making and control, while at higher LOAs, the human is taken more and more out of the decision-making loop. While this scale addresses the allocation of decision-making and action selection authority, it is limited in its ability to fully describe the many methods of collaboration between humans and automation.

Table 2. Levels of Automation [50]

Automation Level	Automation Description
1	The computer offers no assistance: human must take all decision and actions.
2	The computer offers a complete set of decision/action alternatives, or
3	narrows the selection down to a few, or
4	suggests one alternative, and
5	executes that suggestion if the human approves, or
6	allows the human a restricted time to veto before automatic execution, or
7	executes automatically, then necessarily informs humans, and
8	informs the human only if asked, or
9	informs the human only if it, the computer, decides to.
10	The computer decides everything and acts autonomously, ignoring the human.

Sheridan himself argued that the LOA scale, along with Fitts List, both of which focus on "Men are better at – Machines are better at" (MABA-MABA), is too narrow, writing that "the public, and unfortunately too many political and industrial decision-makers, have been slow to realize that function allocation does not necessarily mean allocation of a whole task to either human or machine, exclusive of the other" [51]. Others agree with Sheridan that the traditional scales of function allocation are too narrow, by assigning a task specifically to human or machine, and that flexibility in the allocation of functions is necessary [52-55].

This concept of changing the role of the human and computer during operation has been explored in the body of research on adjustable autonomy and adaptive automation. Both domains focus on adjusting *how* automated a system is, for example, changing from a completely automated system, LOA 10, to management-by-consent, LOA 5 [50]. These adjustments can be made during a mission, either with the human operator instigating the change through adjustable autonomy [56], or with the computer automatically deciding to adjust the level of automation through adaptive automation [37, 57]. The purpose of these adjustments is usually to prevent the operator from becoming either too overloaded with tasks or too bored due to a lack of stimulating tasks.

Both adaptive automation and adjustable autonomy, however, are subtly different from the concept of an automated planner with a dynamic objective function that can be adjusted by a human. Neither the human operator nor the computer would be controlling whether the vehicles are more or less autonomous. Instead, the operator would be directly manipulating the method by which the automated planner optimizes the task allocation, scheduling, and path planning of the various UVs, which remain at the same level of automation. The purpose of these manipulations would not be to maintain an ideal workload for the operator, but to directly impact the plans generated and selected by the human-automation team, which would influence the overall system performance.

In an attempt to take into account greater collaboration between humans and computers than the previously mentioned LOA system, newer models of function allocation have been developed. Riley [58] described an automation taxonomy that can be used in a framework to represent human-machine systems. The taxonomy includes two factors that define the automation levels: the level of intelligence and level of autonomy. At the highest levels of automation and intelligence, the human and machine act as partners to command the system. Kaber, Onal, and Endsley explored the idea of "human-centered levels of automation" in contrast to technology-centered function allocation [59]. They reviewed numerous LOA taxonomies and, as opposed to automating as much as possible and leaving the "left-over" functions for the human operator, they advocated for intermediate LOAs that keep the human operator's SA at higher levels. They argued that potentially higher system performance could be obtained through human-automation collaboration, but they caution that the resulting loss of operator SA at higher LOAs can lead to poorer performance during automation failure.

Many of these researchers have stressed the challenge of developing a framework for designing systems that deal with the uncertainty inherent in dynamic environments [58, 60]. Constraints or preferences are typically not coded completely into the optimization algorithm's objective function, making the collaborative aspect even more important. Specifically, Kirkpatrick, Dilkina, and Havens write, "domains with unmodellable [sic] aspects will benefit from systems that allow the operator to add specific constraints and call for a revised solution" [60]. Cummings and Bruni argue that it is rarely clear what characterizes an "optimal" solution in uncertain scenarios, and that the definition of optimal is a constantly changing concept, particularly in command and controls settings [24]. This theory is depicted in Figure 3, as with increasing uncertainty in the world, additional human interaction is necessary to maintain satisfactory performance. Also, they argue that computer-generated solutions are often suboptimal because in optimization problems with many variables and constraints, the algorithm may make erroneous assumptions, may become trapped in a local minima, and can only take into account those quantifiable variables that were deemed critical in early design stages [61].

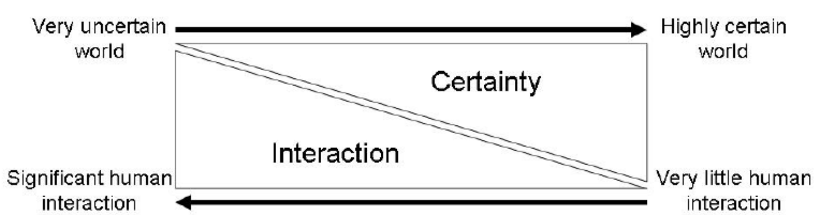


Figure 3. Human-automation interaction as a function of certainty

The Human-Automation Collaboration Taxonomy (HACT) was developed to provide system designers with a model that can be used to analyze collaborative human-computer decision making systems [24, 62]. HACT extends the Parasuraman [63] information processing model by adding to the decision-making component, as shown in Figure 4. HACT adds an iterative data analysis stage combined with an evaluation step where operators can request more information or analysis. Once feasible solutions are selected, either the operator or the automation can select a final solution.

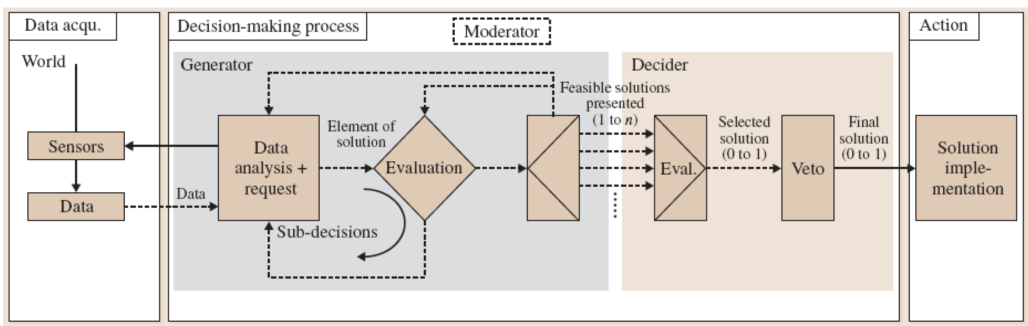


Figure 4. Human-Automation Collaboration Taxonomy Model [24]

The authors of HACT included three distinct roles in the decision-making process: the moderator, generator, and decider. The moderator is responsible for ensuring that each phase in the decision-making process is executed and that the process moves forward. The generator develops feasible solutions and begins to evaluate the solutions. Finally, the decider makes the final selection of the plan and has veto power over this selection. Each of these roles could have different Levels of Collaboration (LOC) between human and computer, rated from -2 where the role is entirely assumed by the automation, to 2 where the human is responsible for the role, as shown in Table 3. A LOC of 0 is a balanced collaboration between the human and automation.

HACT's ability to delineate degrees of collaboration between the human and computer at different points in the decision-making process makes it well suited to model the collaborative replanning method used by the simulation testbed in this thesis. It also provides a basis from which to extend the model to investigate the concept of a dynamic objective function.

Table 3. Moderator, Generator, and Decider Levels in HACT [24]

Level	Who assumes the role of generator and/or moderator?	Who assumes the role of decider?
2	Human	Human makes final decision, automation cannot veto
1	Mixed, but more human	Human or automation can make final decision, human can veto, automation cannot veto
0	Equally shared	Human or automation can make final decision, human can veto, automation can veto
-1	Mixed, but more automation	Human or automation can make final decision, human cannot veto, automation can veto
-2	Automation	Automation makes final decision, human cannot veto

In summary, a number of different taxonomies for determining role allocation between humans and automation have been developed. More recently, these taxonomies have moved away from the rigid "MABA-MABA" framework to take into account the ability of humans and computers to collaborate [52, 55]. Despite the challenges in modeling the impact of uncertainty on collaborative systems, these taxonomies can be useful for modeling collaborative human-automation systems and for predicting the impact of proposed changes to these systems, such as adding a dynamic objective function tool to a collaborative multi-UV control system.

3.2 Application of Theoretical Framework to Simulation Testbed

HACT was chosen to descriptively model human-automation collaboration in the decentralized UV testbed used in this thesis. This section begins by describing the decentralized UV testbed. The HACT model is then applied to descriptively model the interactions between the human operator and automated planner. Finally, the theoretical impact of adding a dynamic objective function is analyzed by extending the HACT model.

3.2.1 Simulation Platform

This thesis utilizes a collaborative, multiple UV simulation environment called Onboard Planning System for UxVs Supporting Expeditionary Reconnaissance and Surveillance (OPS-

USERS), which leverages decentralized algorithms for vehicle routing and task allocation. This simulation environment functions as a computer simulation but also supports actual flight and ground capabilities [18]; all the decision support displays described here have operated actual small air and ground UVs.

Operators are placed in a simulated command center where they control multiple, heterogeneous UVs for the purpose of searching the area of responsibility for new targets, tracking targets, and approving weapons launch. The UVs in the scenario include one fixed-wing UAV, one rotary-wing UAV, one Unmanned Surface Vehicle (USV) restricted to water environments, and a fixed-wing Weaponized Unmanned Aerial Vehicle (WUAV). Once a target is found, it is designated as hostile, unknown, or friendly, and given a priority level by the user. Unknown targets are revisited as often as possible, tracking target movement. Hostile targets are tracked by one or more of the vehicles until they are destroyed by the WUAV. A primary assumption is that operators have minimal time to interact with the displays due to other mission-related tasks.

Participants interact with the simulation via two displays. The primary interface is a Map Display (Figure 5). The map shows both geo-spatial and temporal mission information (i.e., a timeline of mission significant events), and supports an instant messaging "chat" communication tool, which provides high level direction and intelligence. Icons represent vehicles, targets of all types, and search tasks, and the symbology is consistent with MIL-STD 2525 [64].

In the Map Display, operators have two exclusive tasks that cannot be performed by automation: target identification and approval of all WUAV weapon launches. Operators also create search tasks, which dictate on the map those areas the operator wants the UVs to specifically search. The performance plot in Figure 5 gives operators insight into the automated planner performance, as the graph shows expected (red) versus actual (blue) performance. When the automation generates a new plan that is at least five percent "better" than the current plan, the Replan button turns green and flashes, and a "Replan" auditory alert is played. When the Replan button is selected, whether flashing or not, the operator is taken to the Schedule Comparison Tool (SCT), detailed in the next section, for conducting scheduling tasks in collaboration with the automation.

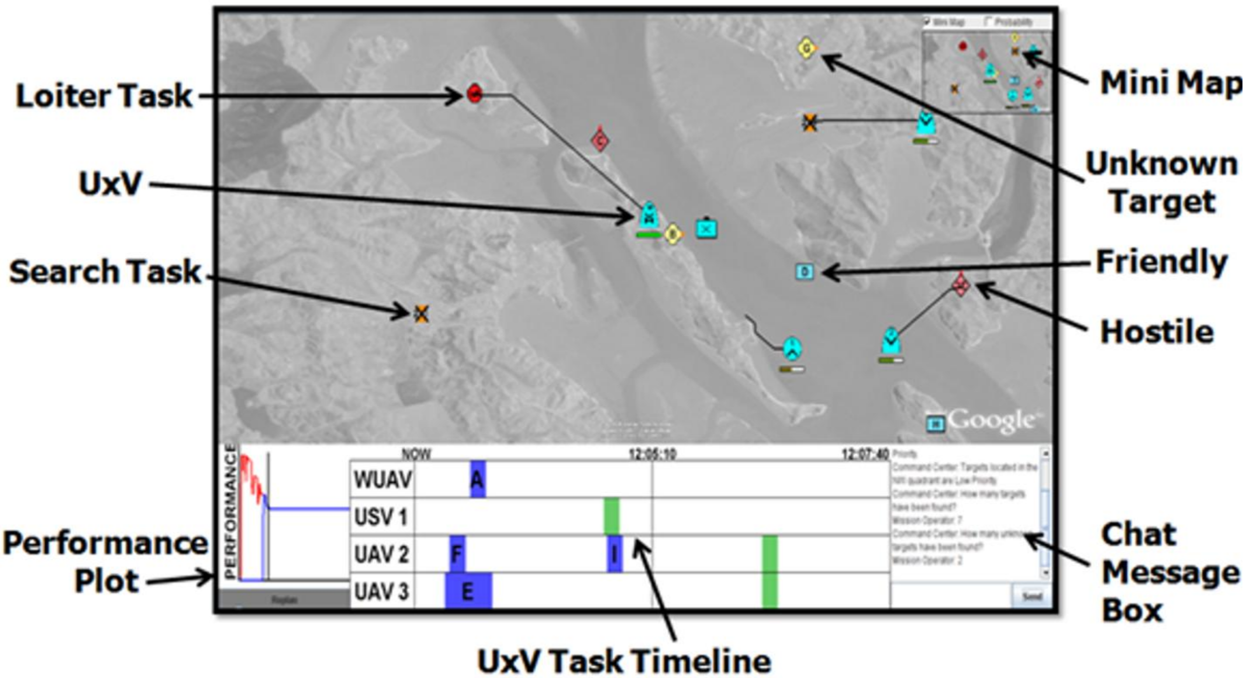


Figure 5. Map Display

3.2.2 Replanning Interface

The SCT display appears when the Replan button is pressed, showing three geometrical forms colored gray, blue, and green at the top of the display (Figure 6). These colors represent configural displays that enable quick comparison of the current, working, and proposed schedules. The left form (gray) is the current UV schedule. The right form (green) is the latest automation proposed schedule. The middle working schedule (blue) is the schedule that results from user modification to the plan. The rectangular grid on the upper half of each shape represents the estimated area that the UVs would search according to the proposed plan. The hierarchical priority ladders show the percentage of tasks assigned in high, medium, and low priority levels.

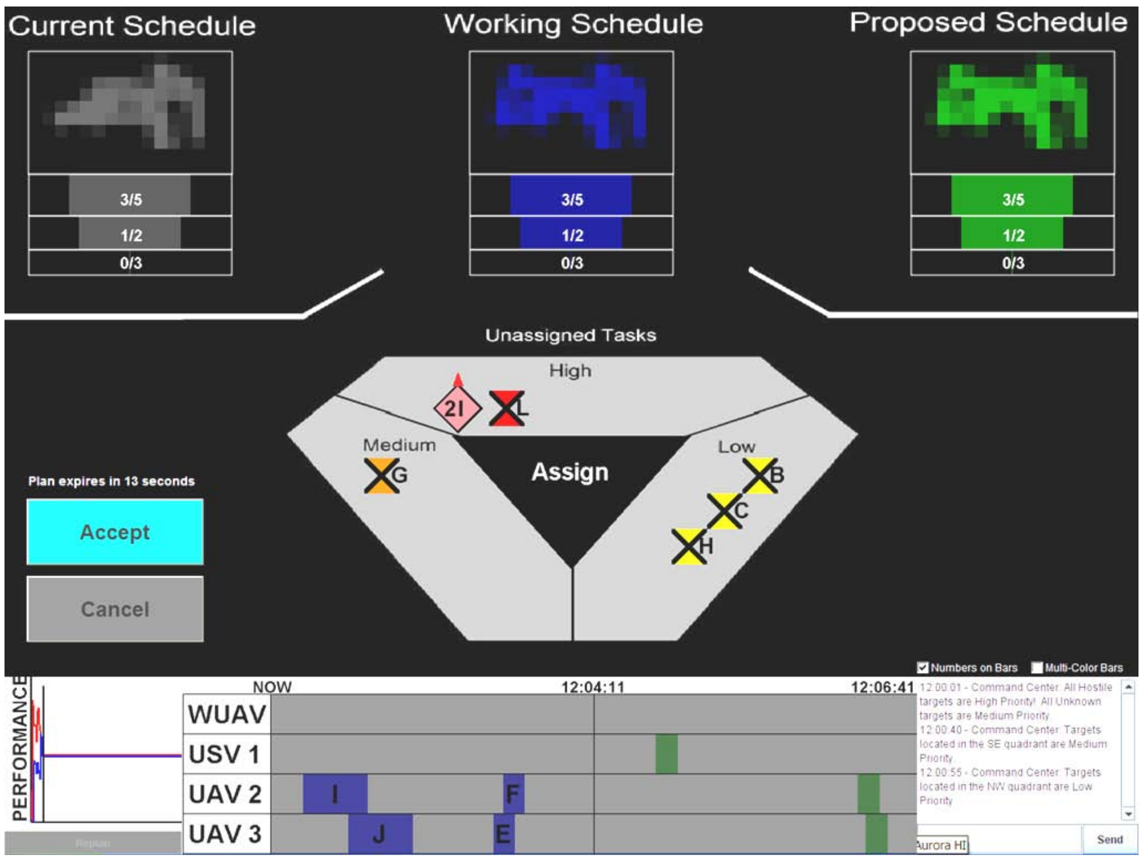


Figure 6. Schedule Comparison Tool

When the operator first enters the SCT, the working schedule is identical to the proposed schedule. The operator can conduct a “what-if” query process by dragging the desired unassigned tasks into the large center triangle. This query forces the automation to generate a new plan if possible, which becomes the working schedule. The configural display of the working schedule alters to reflect these changes. However, due to resource shortages, it is possible that not all tasks can be assigned to the UVs, which is representative of real world constraints. The working schedule configural display updates with every individual query so that the operator can leverage direct-perception interaction [65] to quickly compare the three schedules. This “what-if” query, which essentially is a preview display [40], represents a collaborative effort between the human and automation [66]. Operators adjust team coordination metrics at the task level as opposed to individual vehicle metrics, which has been shown to improve single operator control of a small number of multiple, independent robots [67]. Details of the OPS-USERS interface design and usability testing can be found in Fisher [26].

Operators can either choose to accept the working schedule or cancel to keep the current schedule. Upon accepting a new schedule, the automated planner only communicates to the vehicles via a prioritized task list, and the vehicles sort out the actual assignments amongst themselves. This human-automation interaction scheme is one of high level goal-based control, as opposed to more low-level vehicle-based control.

3.2.3 HACT Application to Testbed

The HACT taxonomy was applied to model the existing simulation testbed prior to the implementation of the dynamic objective function. The testbed was assigned a level 2 moderator because the human operator fully controls the replanning process by deciding when to replan, modify the plan, and accept a final plan. The operator cannot change the criteria to evaluate plans and can only modify the plans by attempting to assign tasks through the “what-if” process. Therefore, the generator role was assigned to level -1, which indicated a mixed role, but with a larger automation presence. Finally, the decider role was assigned to level 1, since the automation presented a final solution to the operator, but the selection of the final solution was completely up to the human operator and the automation did not have veto power.

The HACT framework was extended and slightly modified to illustrate two specific human-automation collaboration methods, as shown in Figure 7. The first is the “what-if” sensitivity analysis tool that already exists in the OPS-USERS system. The second is the proposed dynamic objective function tool for modifying the automated planner.

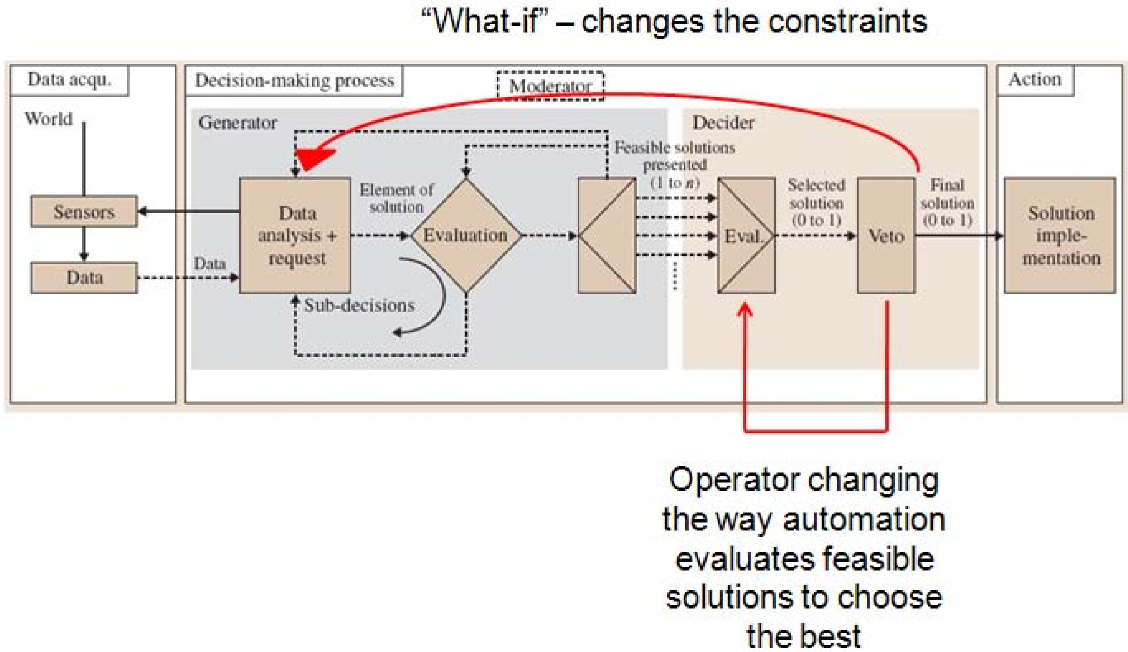


Figure 7. Modified HACT Model with Dynamic Objective Function

The simulation testbed provides a decision support tool that enables an operator to query the automated planner in a "what-if" manner to determine the feasibility and performance consequences of adding a task to the schedule of the UVs. As shown, this process occurs when the human operator is in the decider role, looking at a proposed plan that has been selected by the automated planner. The human operator essentially modifies the constraints placed on the schedule, by specifying that a specific task be assigned in the schedule. These changes send the automated planner back into the generator mode, to recalculate potential solutions to the optimization problem. Many iterations of this "what-if" loop would be required to achieve a solution that the human operator desires, especially if the automated planner is choosing solutions based on an objective function that does not place an emphasis on the quantities of interest to the human operator at that point in the mission.

As illustrated in Figure 7, a dynamic objective function method of human-computer collaboration could result in a shorter loop within the collaborative decision-making process than the “what-if” loop. A dynamic objective function tool would provide the operator with the capability to modify the objective function of the automated planner. This changes the method by which the automated planner would select the best solution, which occurs in the decider role. In terms of the HACT framework, it would change the LOC designation for the decider role from -1 to a more balanced collaborative level of 0. The human operator would have the ability to modify the way that the automation evaluates plans by changing the weightings in the objective function. Positive performance results have been shown in previous research where the human operator could change the search space of the automation [28] or modify the way that the automation evaluates plans [66], even under time-pressure [68]. On the other hand, some researchers have shown that under time-pressure on the order of seconds, human judgment degrades and higher automation roles could be beneficial [69, 70].

In a highly dynamic environment and scenario, less iterations of the longer “what-if” loop would be necessary to achieve a solution that accomplishes what the human operator desires because the objectives of the operator and automated planner would be aligned. Therefore, providing the operator with a dynamic objective function could reduce the number of cognitive steps and amount of time necessary for the combined human-automation team to evaluate and select a new solution. This would reduce the workload of the human operator for replanning, which could positively impact overall mission performance by freeing the operator to focus on making other critical decisions and maintaining SA. As shown in a previous experiment, higher operator workload, especially due to increased rates of replanning, can lead to lower system performance [27].

3.3 Summary

In summary, numerous function allocation taxonomies were reviewed for their applicability to the proposed dynamic objective function. Many of these taxonomies were too rigid, assuming that a function should be performed solely by the human or the computer, instead of allowing for the possibility of human-automation collaboration. Other taxonomies focused exclusively on the LOA concept in order to describe changes in the autonomy of the UVs, instead of allowing for changes in the way that the automated system worked during a collaborative decision-making process. Most of these taxonomies suffer from an inability to sufficiently model the impact of uncertainty on collaborative systems.

Of the reviewed function allocation methods, the HACT was chosen to model the UV simulation testbed used in this thesis. HACT was designed to explicitly take into account levels of collaboration between the human and computer during various stages of the planning and resource allocation decision-making process. The testbed was described, including the Map View for overall operator SA and the SCT for human-automation collaboration in developing schedules for the UVs. HACT was applied to describe the testbed in a manner that would enable theoretical extension to a dynamic objective function capability.

The extended HACT model showed the potential for the dynamic objective function to reduce the workload of the human operator in replanning tasks. The extension revealed the potential for both shorter loops within the collaborative decision-making process and less iterations of the "what-if" loop to reach a satisfactory solution to the human operator. Changing the objectives of the automated planner to match a dynamic mission while potentially reducing the operator's workload could lead to system performance benefits. These theoretical findings were evaluated through human performance experiments, detailed in Chapter 4.

4 Human Performance Experimentation

In order to evaluate the theoretical benefits of a dynamic objective function, derived in the previous chapter, human performance experimentation was conducted using a previously developed multi-UV simulation software package. The experiment tested workload and performance hypotheses using an automated planner with a static objective function and two versions of a dynamic objective function. This chapter describes the experimental objectives and hypotheses, the participants, the apparatus (including the new interfaces designed to enable manipulation of the objective function), the scenarios for the simulation, and the experimental design and procedure.

4.1 Experiment Objectives

The objectives of this experiment focus on providing a human operator who is controlling multiple heterogeneous UVs with the ability to modify the objective function of the automated planner assisting in path planning and task allocation. The specific objective is to test the effectiveness of providing this dynamic objective function manipulation capability for a search, track, and destroy mission. The experiment evaluates the impact of the dynamic objective function on system performance, human cognitive workload, and operator satisfaction. This experiment addresses the gaps in experimental research identified previously, by allowing the operator to collaborate with the automation to plan in a time-critical, dynamic, uncertain environment and by testing different methods to enable the operator to express his or her desires to the automated planner.

4.2 Experimental Hypotheses

4.2.1 Mission Performance

It was hypothesized that the ability to modify the objective function of the automated planner during the mission would enable an operator and the system to achieve higher performance as compared to using a static, a priori coded objective function. Human and system performance were evaluated in three ways. First, performance of the overall mission goals that were provided to operators was evaluated. Second, system performance over time was evaluated through mission efficiency metrics. Finally, as in real-life scenarios, changing external conditions often require the human and the system to adapt, which are experimentally represented through "Rules of Engagement" (ROEs). Mission performance was also measured by adherence to these ROEs and execution of the objectives specified by the ROEs. The following hypotheses describe the expected mission performance:

- *Hypothesis 1*: use of the dynamic objective function is expected to result in significant increases in overall system performance by the end of the mission.
- *Hypothesis 2*: use of the dynamic objective function is expected to result in significant increases in mission efficiency.
- *Hypothesis 3*: the ability to adhere to the ROEs and to perform the specified objectives in the ROEs is expected to improve with use of the dynamic objective function as compared to a static objective function.

4.2.2 Workload

As discussed in the extended HACT model of human-automation collaboration in Chapter 3, collaboration through modification of the objective function of an automated planner could potentially reduce some of the iterations in the "what-if" loop that would typically occur when the human operator's desires do not match up with the objective function of the automated planner. In this mismatch situation, the automated planner would continue to select plans for the operator to view that do not achieve the desired goals of the operator. This can result in a longer time spent attempting to modify the plan manually by assigning tasks individually. Therefore, providing the operator with a dynamic objective function should reduce the amount of time necessary for the combined operator-automated planner team to evaluate and select new plans, as shown in previous research [39]. Workload was measured through an objective utilization metric, through a secondary task to measure spare mental capacity, and through a subjective self-reported workload metric on a five-point Likert scale. The following results were expected:

- *Hypothesis 4*: a reduction in objective and subjective mental workload is expected with use of the dynamic objective function as compared to a static objective function.
- *Hypothesis 5*: use of the dynamic objective function is expected to result in significant reductions in the amount of time spent replanning.

4.2.3 Subjective Appeal

Subjectively, it was expected that operators controlling multiple UVs in a search, track, and destroy mission would prefer to collaborate with an automated planner featuring a dynamic objective function over working with a static, a priori coded objective function. Increased automation transparency and decreased "brittleness" [21] were hypothesized to contribute to these operator preferences. However, it was acknowledged that there could have been a bias towards the static objective function due to its simplicity and due to the need to train operators in using the dynamic objective function tool. Additionally, to avoid additional training that could lead to operator confusion, operators were not allowed to use both the static and dynamic objective functions. Therefore, operators were not able to directly compare the different methods of collaborating with the automated planner. Operators' subjective appeal was determined by analyzing the participants' responses to a survey at the end of the experiment. The following result was expected:

- *Hypothesis 6*: use of the dynamic objective function is expected to result in greater operator satisfaction with the plans generated by the automated planner and higher self-ratings of confidence and performance.

4.3 Participants

To test these hypotheses, 30 participants were recruited from undergraduate students, graduate students, and researchers at the Massachusetts Institute of Technology (MIT). As the concept of multiple UV supervisory control through a decentralized network is a futuristic concept, without current subject matter experts, it was determined that a general user base should first be used to verify the potential of a dynamic objective function.

The 30 participants consisted of 21 men and 9 women. The age range of participants was 18-38 years with an average age of 21.30 and a standard deviation of 3.98. Only 1 participant had served or was currently serving in the military, but a previous experiment using the OPS-USERS system showed that there was no difference in performance or workload between participants based on military experience [27]. Each participant filled out a demographic survey prior to the experiment that included age, gender, occupation, military experience, average hours of television viewing, video gaming experience, and perception of UAVs. The results of these demographic surveys can be found in Appendix A, and the consent forms and demographic surveys filled out by participants can be found in Appendices B and C.

4.4 Testbed

4.4.1 Apparatus

The human performance experiment to test the dynamic objective function tool was conducted using two Dell 17" flat panel monitors operated at 1280 x 1024 pixels and a 32-bit color resolution. The primary monitor displayed the testbed and the secondary monitor showed a legend of the symbols used in the system (Appendix D). The workstation was a Dell Dimension DM051 with an Intel Pentium D 2.80 GHz processor and a NVIDIA GeForce 7300 LE graphics card. System audio was provided using standard headphones that were worn by each participant during the experiment. All data regarding the human participant's interactions with the system for controlling the simulated UVs was recorded automatically by the system.

4.4.2 Dynamic Objective Function Tool

The automated planner in the original testbed used a static objective function to evaluate schedules for the UVs based on maximizing the number of tasks assigned, weighted by priority, while minimizing switching times between vehicles based on arrival times to tasks. A new dynamic objective function was developed for the automated planner that was used in this experiment. Five non-dimensional quantities were chosen as options for evaluating mission plans. The human operators were given the ability to choose the quantities that were high priority, either with guidance from the ROEs or due to their own choices on which aspects of the mission were most important to them at the time. The five quantities were:

- *Area Coverage*: When this quantity was set to high priority, the vehicles covered as much area as possible. The UVs would ignore operator-generated search tasks in favor of using their algorithms to "optimally" explore the unsearched area for new targets. Previously found targets would also not be actively tracked, to free vehicles for the search.

- *Search/Loiter Tasks*: As opposed to allowing the automation to conduct the search for new targets on its own, operators could create search tasks to direct the automation to send vehicles to explore specific regions of the map. Loiter tasks could also be created to direct the WUAV to circle at a particular spot. This quantity for evaluating mission plans was based on the number of assigned search or loiter tasks in a schedule as compared to all available search or loiter tasks. When this quantity was selected, the vehicles performed search tasks that the operator created and the WUAV went to specific loiter points created by the operator.
- *Target Tracking*: This quantity was based on the number of targets assigned to be tracked in a schedule as compared to all available targets.
- *Hostile Destruction*: This quantity was based on the number of assigned hostile destruction tasks as compared to all actively tracked hostile targets that were eligible for destruction. Once a hostile target was found and tracked by one of the regular UVs, it was eligible to be destroyed by the WUAV. The WUAV was only tasked to destroy these hostiles if this quantity was selected.
- *Fuel Efficiency*: This quantity was based on the fuel efficiency of the UVs. Operators could change the weighting of this quantity in order to vary the velocity of the UVs linearly between the cruise and maximum velocity of each UV. The simulated fuel consumption of each UV varied quadratically with velocity. Guided by the ROEs or their own desires, operators could select this quantity as high priority, so that the vehicles traveled more slowly, but also burned fuel more slowly and did not have to refuel as often. The fuel consumptions and velocities of the four UVs used in this experiment are detailed in Appendix E.

For this experiment, only a binary choice of "on" or "off" was allowed for each quantity, with weightings set in advance for the "on" and "off" condition, as opposed to allowing operators to set a weighting anywhere between 0.0 and 1.0 for each quantity. Tversky and Kahneman [71] explained that a human who estimates a numerical value when starting from different initial values often makes insufficient adjustments based on the initial value, a phenomenon known as the "anchoring and adjustment" heuristic. To avoid this issue, operators were limited to a binary choice on each quantity.

The weightings for the "on" and "off" condition were chosen after pilot testing the system in order to achieve schedule selection and UV behavior that was intuitive to human operators. Selecting a quantity gave it a weighting of 1.0 in the objective function of the automated planner, while de-selecting a quantity gave it a weighting of 0.05. The exception was the hostiles destroyed quantity, which received a weighting of 0 when it was de-selected, to prevent the automation from planning to destroy hostile targets without operator permission.

The ability to modify the objective function was implemented in the Schedule Comparison Tool (SCT) through two different interfaces. The first method for modifying the dynamic objective function was through a Checkbox button interface, shown in Figure 8. Operators could select any of the five quantities, in any combination, through the "Plan Priorities" panel on the right side of the SCT. The second method utilized a Radio button interface, shown in Figure 9. Operators could only select one of the quantities at a time, as their highest priority for evaluating potential UV schedules. These two interfaces, along with the static objective function interface (Figure 6), were the three possible types of SCT that operators could use in the human performance experimentation.

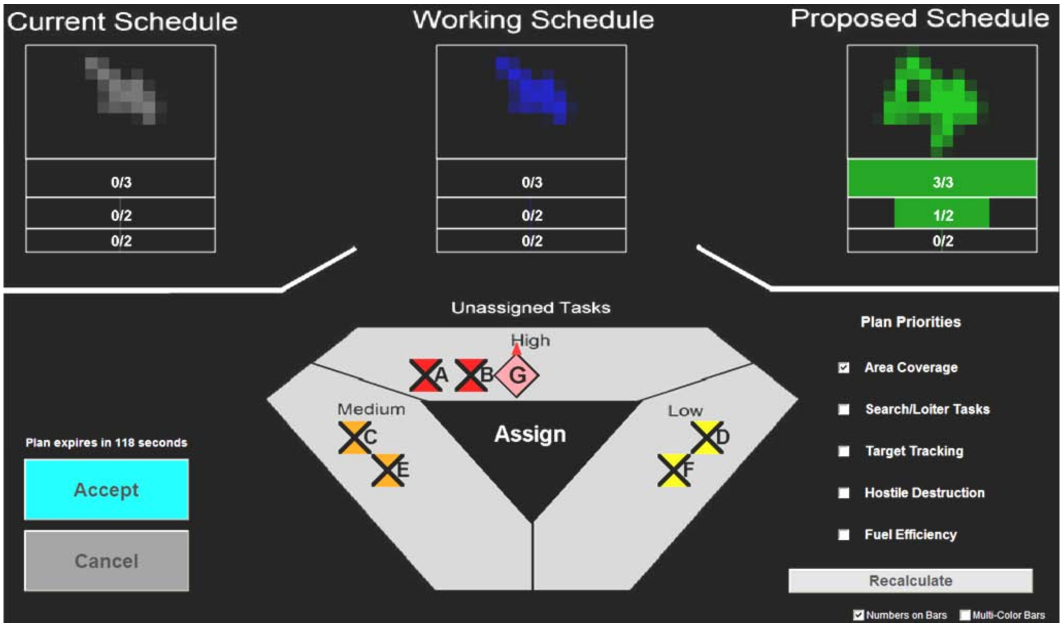


Figure 8. Schedule Comparison Tool with Checkbox Interface

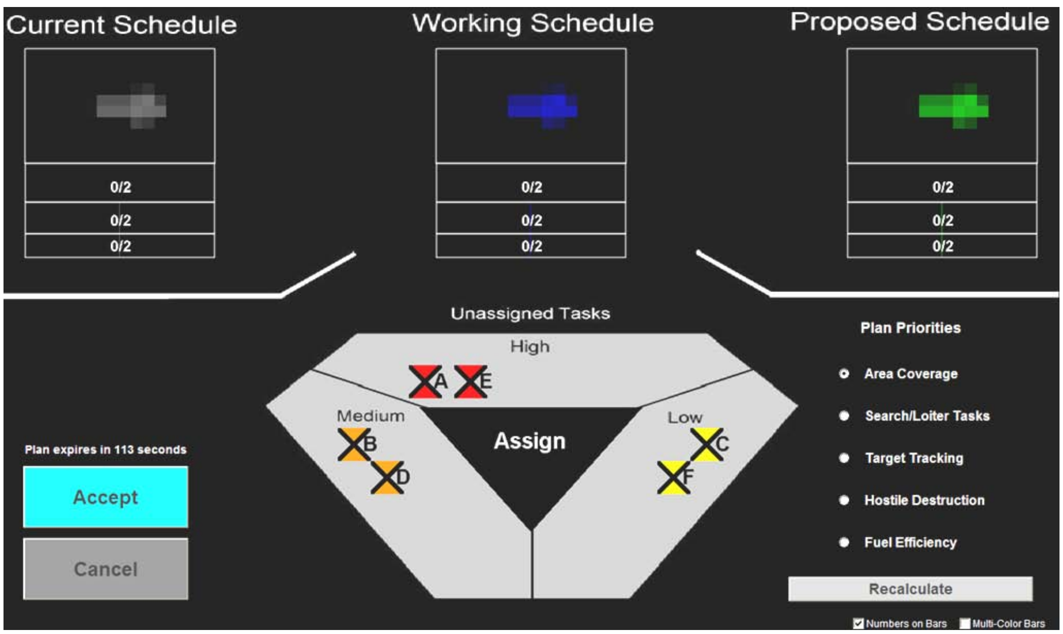


Figure 9. Schedule Comparison Tool with Radio Button Interface

4.5 Experimental Design

Three scenarios were designed for this experiment: a practice scenario and two test scenarios. Each scenario involved controlling four UVs (one of which was weaponized) in a mission to conduct surveillance of an area in order to search for targets, track these targets, and destroy any hostile targets found (when instructed). The area contained both water and land environments and targets could be either tanks on the ground or boats in the water. The vehicles automatically returned to the base when necessary to refuel and were equipped with sensors (either radar or cameras) which would notify the operator when a target was detected so that the operator could view sensor information in order to designate the target and give it a priority level. Perfect sensor operation was assumed, in that there were no false detections or missed target detections.

Each scenario had 10 targets that were initially hidden to the operator. These targets always had a positive velocity and moved on pre-planned paths throughout the environment (unknown to the operator), at roughly 5% of the cruise velocity of the WUAV. Each scenario had three friendly targets, three hostile targets, and four unknown targets. The operator received intelligence information on the unknown targets through the chat window, revealing that two of the targets were friendly and two were hostile. Upon receiving this intelligence, the operator could re-designate the targets. The operator would also be asked by the "Command Center" through the chat window to create search tasks in specified quadrants at various times throughout the mission. The scenarios were all different, but of comparable difficulty, so that operators would not learn the locations of targets between missions.

4.5.1 Independent Variables

The experimental design was a 3x2 repeated measures nested design with two independent variables: the type of objective function used by the automated planner and the type of mission. The objective function type had three levels: "None", "Radio", and "Checkbox." The None level used the original testbed objective function as described earlier in this chapter, which was set a priori and the operator did not have the opportunity to modify it. The Radio level allowed the operator to change the objective function by choosing one of the quantities to be most important at the time. For example, if the operator chose area coverage due to a change in the ROEs, the automated planner optimized the usage of the vehicles for covering the most unsearched area while setting the weights of the other variables to the lowest setting. Finally, in the Checkbox level, the operator was allowed to select any combination of the five quantities to be equally important. This was a between-subjects factor, in that a particular subject only experienced one type of objective function representation, to avoid training biases.

The second independent variable was Mission Type. There were two levels: a Standard Mission and a Dynamic Mission. For the Standard Mission, operators were given a set of ROEs that did not change throughout the mission. The ROEs instructed operators on aspects of the mission that were most important at the time in order to guide their high level decision making. The ROEs also specified when hostile target destruction was permitted. For the Dynamic Mission, every 5 minutes during the 20 minute mission, new ROEs were presented to the operator and the operator needed to decide whether and how to change the objective function under the new ROEs (if they had the interface that allowed for manipulation of the objective function), as well as possibly altering their tasking strategies.

For example, the operator may have received an original ROE stating that they should "Search for new targets and track all targets found." Then, a new ROE may have come in stating "Destroy all Hostile Targets Immediately." Participants could adjust the objective function of the automated planner to reflect the changed ROE, for example by increasing the weighting of the "Destroy Hostiles" quantity or lowering the weightings of other quantities. The ROEs for the Standard and Dynamic missions are listed in Appendix F. This was a within-subjects factor, as each subject experienced both a Standard and Dynamic mission. These missions were presented in a randomized and counterbalanced order to avoid learning effects.

4.5.2 Dependent Variables

The dependent variables for the experiment were mission performance, mission efficiency, primary workload, secondary workload, situational awareness (SA), and subjective ratings of performance, workload, and confidence. Overall mission performance was measured by taking the following four metrics: percentage of area coverage, percentage of targets found, percentage of time that targets were tracked, and number of hostile targets destroyed. Mission efficiency measured the performance metrics over time, which included average time to target detection and average time from hostile detection to destruction. Adherence to the ROEs presented to the operator during the Dynamic Mission (Appendix F) was also measured by the following metrics: 1) number of targets destroyed when hostile target destruction was forbidden, 2) percentage of area covered during the first 5 minutes of the mission, when covering area to find new targets was the highest priority, 3) percentage of targets found during the first 5 minutes of the mission, and 4) percent of time that targets were tracked between 10 and 15 minutes, when tracking all previously found targets was the highest priority.

The primary workload measure was a utilization metric calculating the ratio of the total operator "busy time" to the total mission time. For utilization, operators were considered "busy" when performing one or more of the following tasks: creating search tasks, identifying and designating targets, approving weapons launches, interacting via the chat box, and replanning in the SCT. All interface interactions were via a mouse with the exception of the chat messages, which required keyboard input.

Another method for measuring workload was measuring the spare mental capacity of the operator through reaction times to a secondary task. Secondary workload was measured via reaction times to text message information queries, as well as reaction times when instructed to create search tasks via the chat tool. Such embedded secondary tools have been previously shown to be effective indicators of workload [72].

SA was measured through the accuracy percentage of responses to periodic chat box messages querying the participant about aspects of the mission. Additionally, 4 of the targets were originally designated as unknown. Chat messages would provide intelligence information to the operator about whether these targets were actually hostile or friendly (based on their location on the map). It was up to the operator to re-designate these targets based on this information. Therefore, a second measure of SA was the ratio of correct re-designations of unknown targets to number of unknown targets found.

Finally, a survey was provided at the end of each mission asking the participant for a subjective rating of their workload, performance, confidence, and satisfaction with the plans generated by the automated planner on a Likert scale from 1-5. Subjective ratings are crucial, both for providing an additional measure of workload and for evaluating whether the addition of the dynamic objective function influenced the operator's confidence and trust in the collaborative decision-making process, factors which have been shown to influence system performance [73].

4.6 Procedure

In order to familiarize each subject with the interface, a self-paced, slide-based tutorial was provided (Appendix G). Subjects then conducted a fifteen-minute practice session during which the experimenter walked the subject through all the necessary functions to use the interface. Each subject was given the opportunity to ask the experimenter questions regarding the interface and mission during the tutorial and practice session. Each subject also had to pass a proficiency test, which was a 5-question slide-based test (Appendix H). If the subjects did not pass the proficiency test, they were given time to review the tutorial, after which they could take a second, different proficiency test. All subjects passed on either the first or second test.

The actual experiment for each subject consisted of two twenty-minute sessions, one for each of the two different mission types. The order of the mission types presented to the subject was counterbalanced and randomized to prevent learning effects. During testing, the subject was not able to ask the experimenter questions about the interface and mission. All data and operator actions were recorded by the interface and Camtasia© was used to record the operator's actions on the screen. Finally, a survey was administered at the end of each mission to obtain the participant's subjective evaluation of their workload, performance, and confidence, along with general comments on using the system (Appendix I). Subjects were paid $10/hour for the experiment and a performance bonus of a $100 gift card was given to the individual who obtained the highest mission performance metrics (to encourage maximum effort).

4.7 Summary

Once the experiment was completed, data had been collected for each of the performance, workload, SA, and subjective rating metrics for all 30 participants. In order to evaluate the hypotheses presented in this chapter, the data needed to be formally analyzed using appropriate inferential statistical tests. The statistical tests utilized, and the results of those tests, are presented in Chapter 5.

5 Results

This chapter presents the statistical results of the experiment described in Chapter 4. The experiment included two independent variables: Objective Function Type (None, Radio, or Checkbox) and Mission Type (Standard or Dynamic). Numerous dependent variables were considered in the analysis of the data in order to capture and measure performance, workload, SA, and subjective ratings of performance, workload, and confidence, as described in Chapter 4. First, a system design issue that was identified during the experiment is discussed. Then, an analysis of the dependent variables is presented. Finally, the impact of family-wise error rates is described, along with a summary of the important findings.

5.1 Interface Issue

During the experiment, an issue was uncovered that impacted the performance of operators using the None objective function during the Dynamic mission. For the first 10 minutes of the Dynamic mission, the ROEs stated "Do not destroy any hostiles." Operators using the Radio or Checkbox objective functions were trained to modify the objective function during this time period so that tasks would not be created to destroy hostile targets. Operators using the None objective function type, however, had no way to prevent the automated planner from creating hostile destruction tasks. When the system opened the window shown in Figure 10, requesting permission for the WUAV to destroy a hostile target, the operator was trained to click the "Cancel: Redesignate to Unknown" button if the ROEs did not permit destruction of hostile targets at the time. The result of clicking this button was that the target which was previously designated as hostile was then changed in designation to unknown.

Results from the experiment showed that all 10 participants using the None objective function clicked the "Cancel" button at least once during the Dynamic mission, with 9 of the 10 operators clicking it at least twice. Of all of the trials using the Radio and Checkbox objective function, there was only one “Cancel.”

Figure 10. Hostile Destruction Approval Window

Operators had the option to re-designate these targets back to hostile at any point in the mission, especially once the ROEs changed to permit the destruction of hostile targets. Some of the operators using the None objective function did perform this action successfully, however, many did not, due to inadequate system design and training. Therefore, it was decided that the total hostile targets destroyed and the hostile destruction efficiency metrics would only be used to compare the performance of the operators using the Checkbox and Radio objective functions during Dynamic Missions. Total hostile targets destroyed and hostile destruction efficiency metrics were still used to evaluate the performance of all operators during the Standard Mission.

5.2 Statistical Analysis Overview

All dependent variables were recorded by the computer simulation. For all metrics other than those noted below, a 3 x 2 repeated measures Analysis of Variance (ANOVA) model was used for parametric dependent variables ($\alpha = 0.05$). Unless otherwise noted, all metrics met the homogeneity of variance and normality assumptions of the ANOVA model. For dependent variables that did not meet ANOVA assumptions, non-parametric analyses were used.

Due to the confusion among the operators using the “None” Objective Function type during hostile destruction tasks in the Dynamic mission, a separate analysis was done for the Standard and Dynamic missions for all metrics related to the destruction of hostile targets. A single factor repeated measures ANOVA model was used for parametric dependent variables related to hostile destruction ($\alpha = 0.05$). For analyzing the results of the Dynamic missions, results were only compared between the Radio and Checkbox Objective Function types.

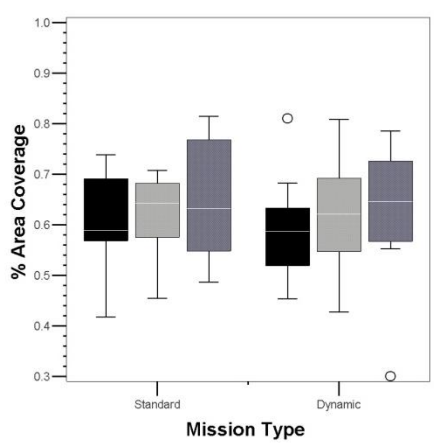
5.3 Mission Performance

As outlined in Section 4.5.2, performance was measured by 1) overall mission performance metrics, computed at the end of the mission; 2) satisfaction of the ROEs that were presented to the operator at 5 minute intervals during the Dynamic Mission; and 3) by mission efficiency metrics, which measure performance over time.

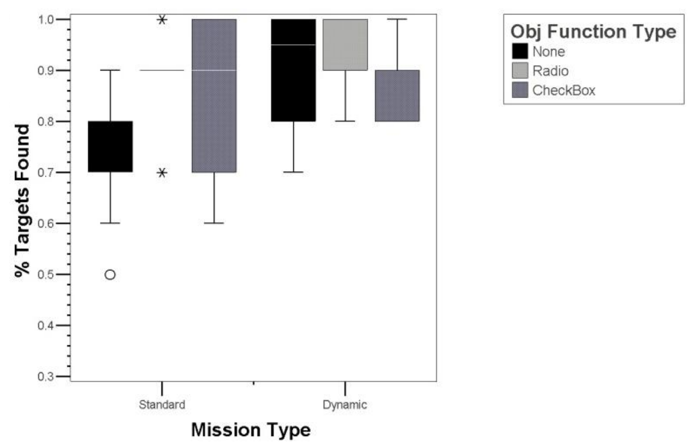
5.3.1 Overall Mission Performance

The four overall mission performance metrics were percentage of area coverage, percentage of targets found, percentage of time that targets were tracked, and number of hostile targets destroyed. The omnibus area coverage test was not significant for Mission Type, $F(1,27) = 0.328$, $p = 0.571$, nor for Objective Function Type, $F(2,27) = 0.344$, $p = 0.712$. For the percentage of targets found, non-parametric tests were needed. The Mann-Whitney dependent test on the percentage of targets found showed a significant difference across Mission Type, $Z = -2.795$, $p = 0.005$, where more targets were found in the Dynamic Mission Type. The Kruskal-Wallis omnibus test on the percentage of targets found was not significant for Objective Function Type, $\chi^2(2, N=60) = 3.599$, $p = 0.165$. The omnibus percentage of time that targets were tracked test was not significant for Mission Type: $F(1,27) = 1.115$, $p = 0.300$, nor for Objective Function Type, $F(2,27) = 1.961$, $p = 0.160$.

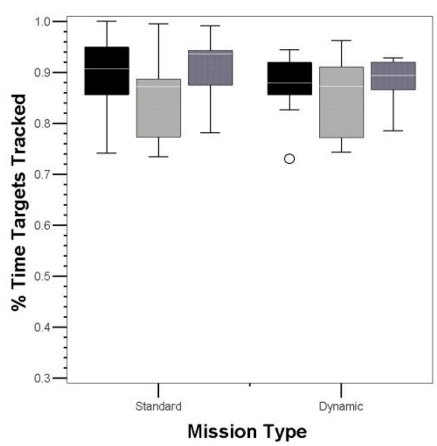
For the number of hostile targets that were destroyed, non-parametric tests were needed. A separate analysis was performed for the Standard and Dynamic Mission Types, where the Dynamic Mission excluded the "None" Objective Function Type. For the Standard Mission, the omnibus Kruskal-Wallis test was not significant for Objective Function Type, $\chi^2(2, N=30) = 3.729$, $p = 0.155$. For the Dynamic Mission, the Mann-Whitney independent test was not significant for Objective Function Type, $Z = -1.592$, $p = 0.111$. The boxplots in Figure 11 illustrate the results for the performance metrics, and Table 4 summarizes the key statistics.



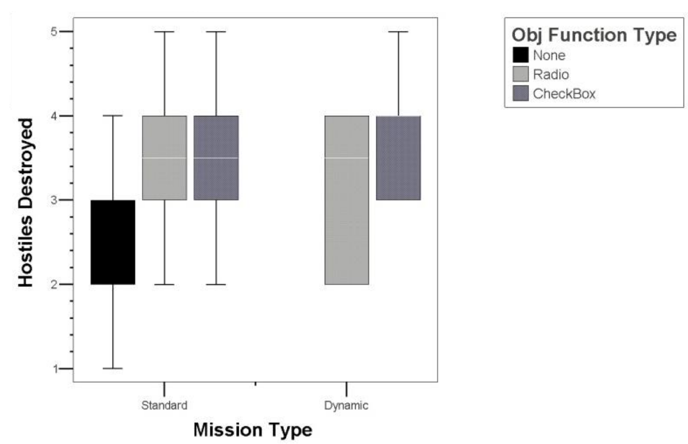
(a) Area Coverage



(b) Targets Found



(c) Time Targets Tracked



(d) Hostiles Destroyed

Figure 11. Performance Metrics Comparison

Table 4. Performance Metrics Summary

Metric	Mission Type	Objective Function	Mean	Median	Std Dev
% Area Coverage	Standard	None	60.6%	58.9%	9.7%
		Radio	62.0%	64.3%	8.1%
		Checkbox	64.3%	63.2%	11.2%
	Dynamic	None	59.5%	58.7%	10.3%
		Radio	61.1%	62.1%	11.8%
		Checkbox	62.7%	64.6%	14.0%
% Targets Found	Standard	None	72.0%	70.0%	11.4%
		Radio	87.0%	90.0%	9.5%
		Checkbox	84.0%	90.0%	15.1%
	Dynamic	None	90.0%	95.0%	11.6%
		Radio	91.0%	90.0%	7.4%
		Checkbox	89.0%	90.0%	7.4%
% Time Targets Tracked	Standard	None	89.5%	90.7%	7.6%
		Radio	84.8%	87.2%	87.2%
		Checkbox	91.1%	93.6%	5.9%
	Dynamic	None	87.5%	87.8%	6.3%
		Radio	85.0%	87.3%	7.7%
		Checkbox	88.5%	89.3%	4.4%
Hostiles Destroyed	Standard	None	2.7	3.0	0.9
		Radio	3.5	3.5	0.9
		Checkbox	3.4	3.5	1.0
	Dynamic	None	-	-	-
		Radio	3.2	3.5	0.9
		Checkbox	3.9	4	0.7

5.3.2 Satisfaction of Rules of Engagement in Dynamic Mission

As described in Section 4.5.2, satisfaction of the ROEs was measured by 1) number of targets destroyed when hostile target destruction was forbidden, 2) percentage of area covered during the first 5 minutes of the mission, when covering area to find new targets was the highest priority, 3) percentage of targets found during the first 5 minutes of the mission, and 4) percent of time that targets were tracked between 10 and 15 minutes, when tracking all previously found targets was the highest priority. No significant differences were found for the percentage of area covered during the first 5 minutes and for the percent of time that targets were tracked between 10 and 15 minutes.

With regards to the restriction during the first ten minutes of the Dynamic mission that no hostile targets were to be destroyed, it was found that of the 30 trials of the Dynamic mission, 3 test subjects violated this ROE and destroyed a hostile target before it was permitted. All 3 of these test subjects used the Radio Objective Function.

The percentage of all targets found in the first 5 minutes of the Dynamic mission was analyzed, as the highest priority of operators during this time period was to search for new targets. The omnibus test on targets found in the first 5 minutes was significant for Objective Function Type, $F(2,27) = 4.517$, $p = 0.02$. Tukey pairwise comparisons showed that the Radio Objective Function was different from Checkbox and None Objective Functions ($p = 0.02$ and $p = 0.012$, respectively), but the Checkbox and None Objective Functions were not statistically different ($p = 0.823$). Operators who used the Radio Objective Function found more targets in the first 5 minutes of the Dynamic mission. The boxplots in Figure 12 illustrate the results for number of targets found in the first 5 minutes, and Table 5 summarizes the key statistics.

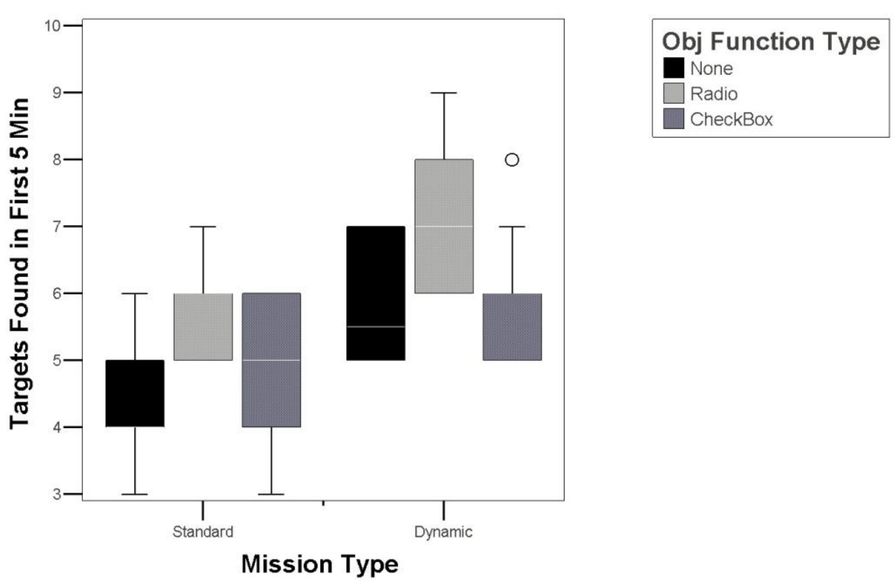


Figure 12. Targets Found in the First 5 Minutes Comparison

Table 5. Targets Found in the First 5 Minutes Summary

Objective Function	Mean	Median	Std Dev
None	5.8	5.5	0.9
Radio	7.0	7.0	1.0
Checkbox	5.9	6.0	1.0

5.3.3 Mission Efficiency

Finally, performance was measured by efficiency metrics, which characterize performance over time. The efficiency metrics were the average time to target detection and the average time from when a hostile target was detected to its destruction, as calculated using the formulas shown in Equations 1 and 2 respectively. Each metric is calculated in a three step process. First, for each target, the time to either find the target or to destroy the target after it was designated as hostile is divided by the amount of time the target was available. In the case of finding a target, it was available to be found the entire simulation. If a target was not found or not destroyed, it is given a ratio of 1. Second, the ratios are summed and divided by the total number of targets found or hostiles destroyed. This metric shows both speed and quantity of either targets found or hostiles destroyed, where a lower score is better. Third, the metric is normalized by dividing by the total number of targets available (10 targets) or hostiles available (5 targets). To make the efficiency metric such that a higher score is better, it is subtracted from 1, so the maximum value is 1 and the minimum value is 0.

$$\mathbf{1 - \frac{1}{N^*}\left[\frac{\sum_{i=1}^{N^*}\frac{f_i}{T}}{N}\right]} \tag{1}$$

$$1 - \frac{1}{H^*}\left[\frac{\sum_{i=1}^{H^*}\left(\frac{d_i - r_i}{T - r_i}\right)}{H}\right] \tag{2}$$

Where:

- T = Total Simulation Time (1200 seconds)
- N* = total number of targets available in simulation (10 targets)
- H* = total number of hostile targets available in simulation (5 targets)
- N = total number of targets found during simulation
- H = total number of hostile targets found during simulation
- f_i = Time in seconds that target i was found (set to 1200 if never found)
- r_i = Time in seconds that a hostile target was re-designated as hostile (set to 0 if never found or re-designated)
- d_i = Time in seconds that a hostile target was destroyed (set to 1200 if never destroyed)

The omnibus target finding efficiency test was significant for Mission Type, $F(1,26) = 32.687$, $p < 0.001$ and also significant for Objective Function Type, $F(2,26) = 3.776$, $p = 0.036$. Tukey pairwise comparisons showed that the Radio Objective Function was different from the None Objective Function ($p = 0.011$), but there was no significant difference between the Checkbox and either the Radio or None Objective Functions ($p = 0.134$ and $p = 0.230$, respectively). Operators using the Radio Objective Function had the highest target finding efficiency and all operators had a higher target finding efficiency during the Dynamic Mission. The omnibus hostile destruction efficiency test was not significant for Objective Function Type in the Standard Mission, $F(2,26) = 0.971$, $p = 0.392$, and was only marginally significant in the Dynamic Mission, $F(1,18) = 4.329$, $p = 0.052$. The boxplots in Figure 13 illustrate the results for the efficiency metrics, and Table 6 summarizes the key statistics.

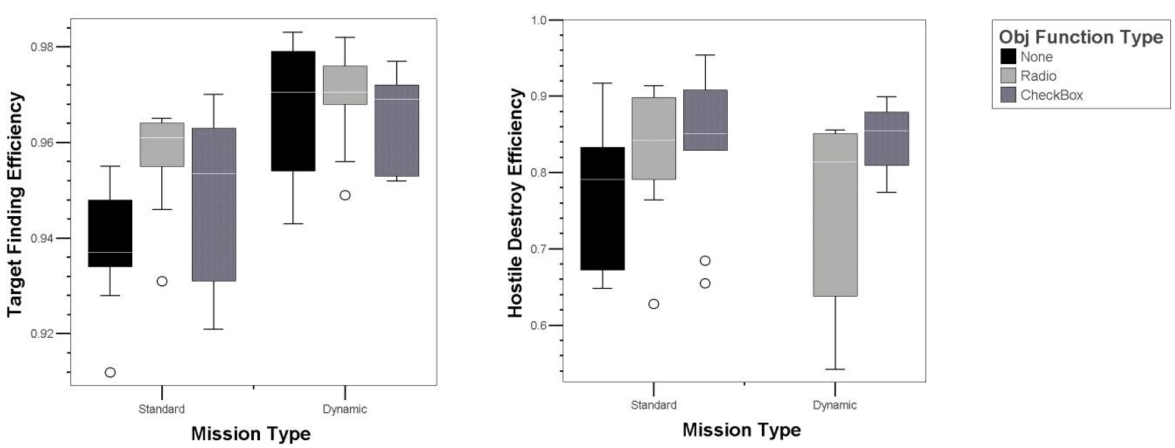


Figure 13. Target Finding and Hostile Destruction Efficiency Comparison

Table 6. Target Finding and Hostile Destruction Efficiency Summary

Metric	Mission Type	Objective Function	Mean	Median	Std Dev
Target Finding Efficiency	Standard	None	0.938	0.937	0.013
		Radio	0.957	0.961	0.011
		Checkbox	0.948	0.954	0.018
	Dynamic	None	0.965	0.970	0.014
		Radio	0.969	0.971	0.010
		Checkbox	0.965	0.969	0.009
Hostile Destruction Efficiency	Standard	None	0.779	0.791	0.104
		Radio	0.826	0.842	0.086
		Checkbox	0.838	0.851	0.098
	Dynamic	None	-	-	-
		Radio	0.758	0.814	0.122
		Checkbox	0.844	0.855	0.045

5.4 Workload

Primary workload was measured through utilization, calculating the ratio of the total operator "busy time" to total mission time. Time spent replanning in the SCT was evaluated as a component of workload. In addition to these primary workload metrics, secondary workload was measured via reaction times to text message information queries, as well as reaction times when instructed to create search tasks via the chat tool.

5.4.1 Utilization

The omnibus utilization test was significant for Mission Type, $F(1,27) = 5.216$, $p = 0.030$, but was not significant for Objective Function Type, $F(2,27) = 1.122$, $p = 0.340$. Operator utilization was higher during the Dynamic mission than the Standard mission. The boxplot in Figure 14 illustrates the results for utilization, and Table 7 summarizes the key statistics.

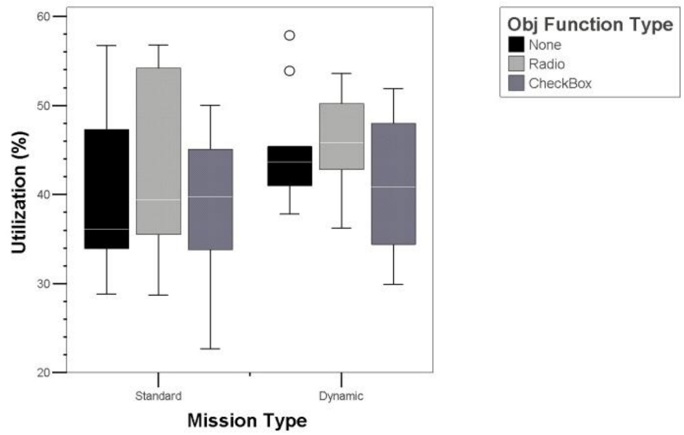


Figure 14. Utilization Comparison

Table 7. Utilization Summary

Mission Type	Objective Function	Mean (%)	Median (%)	Std Dev (%)
Standard	None	40.1	36.1	9.8
	Radio	42.8	39.4	9.6
	Checkbox	38.5	39.8	8.8
Dynamic	None	44.9	43.7	6.4
	Radio	45.8	45.8	5.4
	Checkbox	40.8	40.9	7.4

5.4.2 Time Spent in the Schedule Comparison Tool (SCT)

Operators using either of the dynamic objective functions (Checkbox or Radio) potentially had more to do while in the Schedule Comparison Tool (SCT), such as modifying the weightings of the objective function. There was, however, no significant difference in average time spent in the SCT among the three types of objective function, $F(2,27) = 2.039$, $p = 0.150$. As can be expected due to the increased complexity of the Dynamic Mission as compared to the Standard Mission, there was a significant difference in the average time spent in the SCT between the two mission types, $F(1,27) = 20.786$, $p < 0.001$. Operators spent more time, on average, in the SCT during the Dynamic Mission as compared to the Standard Mission.

5.4.3 Secondary Workload

For the Standard Mission, there were no significant differences in chat message response time or in reaction time to creating a search task when prompted. For the Dynamic Mission, there were four measures of secondary workload: a chat message question requiring a response at 235 seconds, a prompt to create a search task at 300 seconds, another prompt to create a search task at 725 seconds, and finally, a chat message question requiring a response at 1104 seconds.

The omnibus test for the reaction time to the chat question at 235 seconds was significant for Objective Function Type, $F(2,26) = 8.839$, $p = 0.001$. Tukey pairwise comparisons showed that the None Objective Function was different from Checkbox and Radio Functions ($p = 0.001$ and $p = 0.002$, respectively), but the Checkbox and Radio Objective Functions were not statistically different ($p = 0.703$). Operators using the None Objective Function had slower reaction times to answer the chat question at 235 seconds.

The omnibus test for the reaction time to the chat question at 1104 seconds was significant for Objective Function Type, $F(2,26) = 3.411$, $p = 0.048$. Tukey pairwise comparisons showed that the Checkbox Objective Function was different from the None Objective Function ($p = 0.022$), but there were no significant differences between the Radio and either the Checkbox or None Objective Functions ($p = 0.056$ and $p = 0.712$, respectively). Generally, operators using the Checkbox objective function had faster reaction times to answer the chat question at 1104 seconds.

All other reaction times were not significantly different. Figure 15 illustrates the reaction times for the four secondary workload measures during the Dynamic mission, showing the average reaction times to each prompt. Table 8 summarizes the key statistics for the two chat message reaction times analyzed above.

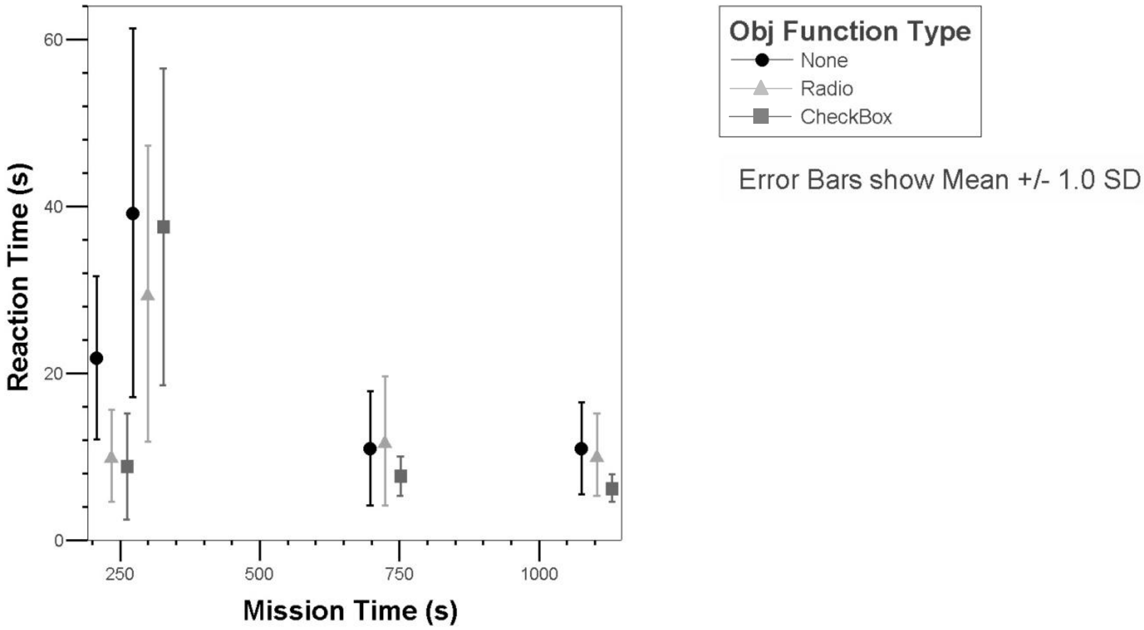


Figure 15. Secondary Workload Metrics for Dynamic Mission Comparison

Table 8. Secondary Workload Metrics for Dynamic Mission Summary

Metric	Objective Function	Mean (s)	Median (s)	Std Dev (s)
Chat Message Reaction Time at 235 seconds	None Radio Checkbox	21.81 8.94 8.82	24.69 7.88 6.21	9.80 4.36 6.38
Chat Message Reaction Time at 1104 seconds	None Radio Checkbox	11.50 10.23 6.21	9.14 7.98 5.69	5.62 4.94 1.63

5.5 Situational Awareness

SA was measured through two metrics: the accuracy of responses to periodic chat box messages querying the participant about aspects of the mission and the accuracy of re-designations of unknown targets based on chat intelligence information. For both metrics, non-parametric tests were needed.

The Mann-Whitney dependent test on chat accuracy showed no significant differences across Mission Type, Z = 0.0, p = 1.0. The Kruskal-Wallis omnibus test on chat accuracy was significant for Objective Function Type, $\chi^2(2, N=60) = 6.167$, p = 0.046. Further Mann-Whitney independent pairwise comparisons showed that the Checkbox Objective Function was different from the None Objective Function (p = 0.013) and marginally significantly different from the Radio Objective Function (p = 0.057). There was no significant difference between the Radio and None Objective Functions (p = 0.551). Operators using the Checkbox Objective Function had higher chat accuracy than the None and Radio Objective Function users.

The Mann-Whitney dependent test on re-designation accuracy showed a significant difference across Mission Type, Z = -2.482, p = 0.013, where operators had higher re-designation accuracy during the Dynamic Mission. The Kruskal-Wallis omnibus test on the re-designation accuracy was also significant for Objective Function Type, $\chi^2(2, N=60) = 10.392$, p = 0.006. Further Mann-Whitney independent pairwise comparisons showed that the None Objective Function was different from Checkbox and Radio Objective Functions (p = 0.003 and p = 0.019 respectively), but the Checkbox and Radio Objective Functions were not statistically different (p = 0.342). Operators using the None Objective Function had lower re-designation accuracy than operators using either the Checkbox or Radio Objective Function. The boxplots in Figure 16 illustrate the results for chat accuracy and re-designation accuracy, and Table 9 summarizes the key statistics for both chat accuracy and re-designation accuracy.

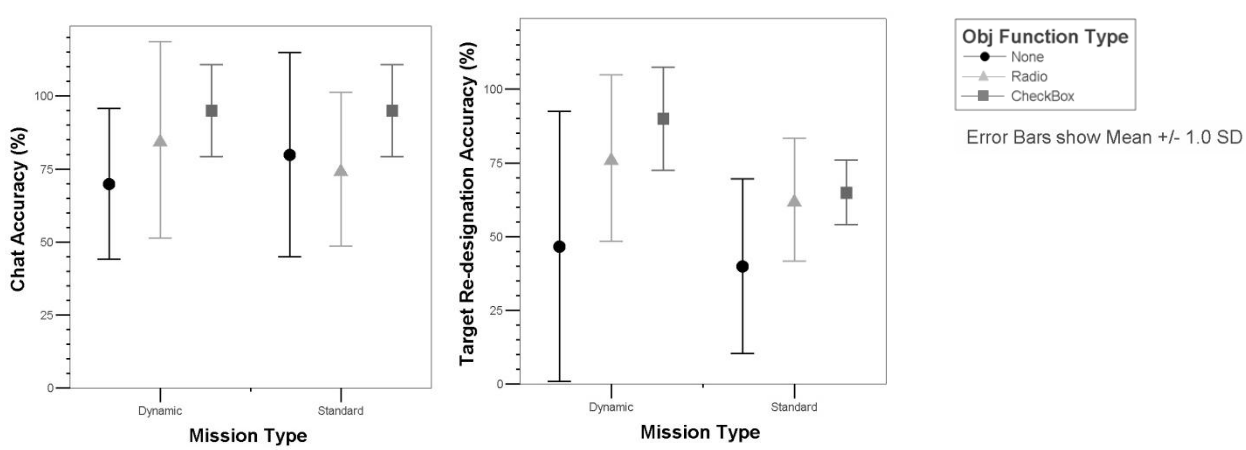


Figure 16. Chat Accuracy and Target Re-designation Comparison

Table 9. Chat Accuracy and Target Re-designation Summary

Metric	Mission Type	Objective Function	Mean (%)	Median (%)	Std Dev (%)
Chat Question Accuracy	Standard	None	80%	100%	35.0%
		Radio	75%	75%	26.4%
		Checkbox	95%	100%	15.8%
	Dynamic	None	70%	50%	25.8%
		Radio	85%	100%	33.7%
		Checkbox	95%	100%	15.8%
Target Re-designation Accuracy	Standard	None	40.0	33.3	29.6
		Radio	62.5	66.7	20.9
		Checkbox	65.0	66.7	11.0
	Dynamic	None	46.7	45.9	45.8
		Radio	76.7	87.5	28.2
		Checkbox	90.0	100	17.5

5.6 Subjective Responses

A survey was provided at the end of each mission asking the participant for a subjective rating of his or her workload, performance, confidence, and satisfaction with the plans generated by the automated planner on a Likert scale from 1-5 (1 low, 5 high). Non-parametric tests were needed for this Likert scale data. There were no significant differences among the ratings of workload and satisfaction with the plans generated by the automated planner.

The Mann-Whitney dependent test on subjective performance rating was not significant for Mission Type, $Z = -0.215$, $p = 0.830$. The Kruskal-Wallis omnibus test on the performance rating was, however, significant for Objective Function Type, $\chi^2(2, N=60) = 15.779$, $p < 0.001$. Further Mann-Whitney independent pairwise comparisons showed that the Checkbox Objective Function was different from None and Radio Objective Functions ($p < 0.001$ and $p = 0.008$ respectively), but the None and Radio Objective Functions were not statistically different ($p = 0.224$). Operators using the Checkbox Objective Function had the highest self-ratings of performance.

Similar results were obtained for subjective ratings of confidence. The Mann-Whitney dependent test on the confidence rating was not significant for Mission Type, Z = -1.057, p = 0.291. The Kruskal-Wallis omnibus test on the confidence rating was, however, significant for Objective Function Type, $\chi^2(2, N=60) = 12.540$, p = 0.002. Further Mann-Whitney independent pairwise comparisons showed that the Checkbox Objective Function was different from None and Radio Objective Functions (p = 0.001 and p = 0.011 respectively), but the None and Radio Objective Functions were not statistically different (p = 0.430).

Operators using the Checkbox Objective Function rated their performance and confidence as higher than operators using the other objective functions. It should be noted that there was a significant effect on confidence ratings for the order that the Mission Types were shown to the operator (p = 0.026). Confidence ratings were higher when operators saw the Standard Mission prior to the Dynamic mission, as opposed to seeing the Dynamic Mission prior to the Standard Mission.

The plots in Figure 17 illustrate the self-rating results and Table 10 summarizes the key statistics for performance and confidence self-ratings.

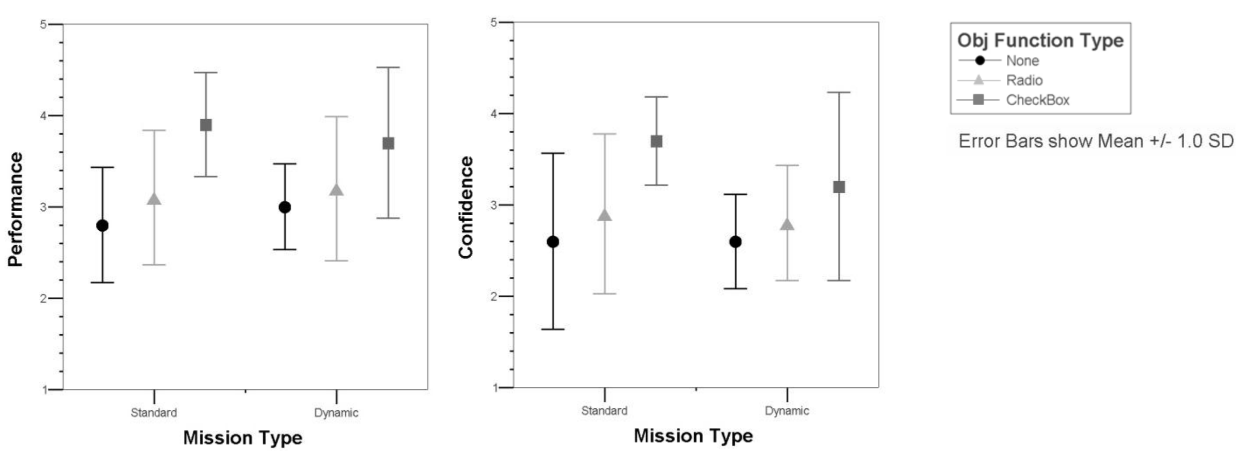


Figure 17. Performance and Confidence Self-ratings Comparison

Table 10. Performance and Confidence Self-ratings Summary

Metric	Mission Type	Objective Function	Mean	Median	Mode	Std Dev
Performance self-rating	Standard	None	2.8	3	3	0.6
		Radio	3.1	3	3	0.7
		Checkbox	3.9	4	4	0.6
	Dynamic	None	3.0	3	3	0.5
		Radio	3.2	3	3	0.8
		Checkbox	3.7	4	3, 4	0.8
Confidence self-rating	Standard	None	2.6	3	3	1.0
		Radio	2.9	3	2	0.9
		Checkbox	3.7	4	4	0.5
	Dynamic	None	2.6	3	3	0.5
		Radio	2.8	3	3	0.6
		Checkbox	3.2	3	4	1.0

5.7 Operator Strategy and Top Performer Analysis

A further analysis of the strategies of the participants was conducted, focusing on those participants who used either the Radio or Checkbox objective function. In addition, a set of analyses were performed to determine if there were additional trends in the data that predicted high performance, based on operator strategy or demographic factors.

5.7.1 Operator Strategies with Dynamic Objective Function

Investigating the number of objective function modifications made by operators using the dynamic objective functions, we find a significant difference between the strategies adopted by participants using the Checkbox versus the Radio objective function. Radio operators made more total modifications to the objective function than Checkbox operators, $F(1,17) = 26.094$, $p < 0.001$. In fact, Radio operators modified the objective function more than double the amount that Checkbox operators did, with an average of 28.3 modifications over the 20 minute simulation as compared to 12.4 modifications for the Checkbox operators.

Of all of their SCT sessions, Radio operators made at least one modification to the objective function 66.8% of the time, as compared to 35.5% of SCT sessions for Checkbox operators. Radio operators modified the objective function more times per SCT session as well ($F(1,17) = 23.395$, $p < 0.001$), making on average of 0.85 modifications per session, as compared to 0.45 modifications per session for Checkbox operators. All of these values were calculated with combined data from the Standard and Dynamic Mission Types.

5.7.2 Top Performers

A set of linear regression analyses was performed to see if there were any significant predictor variables for high (or low) system performance and operator workload. The linear regression estimates coefficients of a linear equation, with one or more predictor variables, that best predict the value of the dependent variable. The system performance and operator workload dependent variables were percentage of area coverage, percentage of targets found, percentage of time that targets were tracked, number of hostile targets destroyed, and operator utilization.

As there would be 5 linear regressions, the typical $\alpha = 0.05$ significance level was reduced to $\alpha = 0.01$ using the Bonferroni correction [74]. A backwards elimination linear regression was utilized, which removed predictor variables that did not meet a significance level of $\alpha = 0.01$, so that the most parsimonious model was derived for predicting the dependent variables. Potential predictor variables included both demographic information and strategy information derived from experimental data. These variables were age, gender, gaming experience, perception of UAVs, comfort with computers, recent amount of sleep, occupation/education level, total number of objective function modifications, number of objective function modifications per SCT session, and the percentage of SCT sessions with at least one objective function modification.

Table 11 shows the results from the 5 backwards elimination linear regressions, including the variables which were significant predictors of the performance and workload metrics. The normality, homogeneity of variance, linearity, and independence assumptions of a linear regression were met by the 3 regressions that found significant predictor variables. There were no significant predictor variables for the number of targets found and the number of hostiles destroyed.

Table 11. Linear Regression Results

Dependent Variable	R^2	β_0	Education Level	Total Objective Function Mods	Mods per SCT Session	Percent of SCT Sessions with a Mod
Area Coverage	0.459	$\beta = 0.738$ $p < 0.001$	$\beta = -0.067$ $p = 0.001$	$\beta = 0.012$ $p < 0.001$	$\beta = -0.382$ $p < 0.001$	-
Targets Found	0	$\beta = 0.876$ $p < 0.001$	-	-	-	-
Time Targets Tracked	0.172	$\beta = 0.938$ $p < 0.001$	-	-	-	$\beta = -0.132$ $p = 0.010$
Hostiles Destroyed	0	$\beta = 3.474$ $p < 0.001$	-	-	-	-
Utilization	0.261	$\beta = 0.346$ $p < 0.001$	-	$\beta = 0.004$ $p = 0.001$	-	-

For the area coverage regression, 3 significant predictor variables were found. The first is education level, where the test participants reported whether they were an undergraduate, master's, or Ph.D student. These categories were numbered 1, 2, or 3, respectively, along with a category of 4 for "Non-student/other" (demographic data can be found in Appendix A). A negative relationship was found between increasing education level and total area coverage. For example, moving from undergraduate to a master's level student would result in a 6.7% decrease in area coverage through this linear model.

The second significant predictor variable for area coverage was the total number of objective function modifications. A positive relationship was found between increasing number of total modifications with area coverage percentage. The linear model predicted a 1.2% increase in area coverage for each additional modification to the objective function. Finally, the third significant predictor variable was the average number of modifications per SCT session. A negative relationship was found between increasing average modifications per SCT session with area coverage, in that an increase of 1 in the average number of modification per SCT session would predict a 38.2% decrease in area coverage.

For the linear regression on the percentage of time that targets were tracked, the only significant predictor variable was the percent of SCT sessions with an objective function modification. A negative relationship was found between these two quantities, in that an increase of 1% in the percent of SCT sessions with an objective function modification would result in a 0.00132% reduction in the percentage of time that targets were tracked.

Finally, for the linear regression on utilization, the only significant predictor variable was the total number of objective function modifications. A positive relationship was found between these two quantities, in that the linear model predicted a 0.4% increase in utilization for each additional modification to the objective function.

5.8 Summary

Results from the human performance experiment led to a range of results. The analysis indicated that operators using the Checkbox and Radio objective functions had superior results in some of the metrics, while there were no significant differences in other metrics. These results can aid in evaluating the impact of the dynamic objective function based on theoretical predictions in Chapter 3. Results also indicated that operators generally performed better in the Dynamic mission over the Standard Mission. All of the results are summarized in Table 12, where the conditions with superior results are shown in bold.

Table 12. Summary of Experimental Findings

Category	Metric	Objective Function Type	Mission Type
System Performance	% Area Coverage	Indistinguishable (p = 0.571)	Indistinguishable (p = 0.712)
	% Targets Found	Indistinguishable (p = 0.165)	**Dynamic Mission (p = 0.005)**
	% Time Targets Tracked	Indistinguishable (p = 0.160)	Indistinguishable (p = 0.300)
	Hostiles Destroyed	Indistinguishable (p = 0.155 & p = 0.111)	N/A
Adherence to ROEs	Hostiles Destroyed when restricted	**Checkbox and None (0 errors)**	N/A
	% Area Coverage during first 5 min	Indistinguishable (p = 0.687)	N/A
	Targets Found during first 5 min	**Radio (p = 0.020)**	N/A
	% Time Targets Tracked between 10-15 min	Indistinguishable (p = 0.107)	N/A
Mission Efficiency	Target Finding Efficiency	**Radio (p = 0.011)**	**Dynamic Mission (p < 0.001)**
	Hostile Destruction Efficiency	Indistinguishable (p = 0.392 and p = 0.052)	N/A
Primary Workload	Utilization	Indistinguishable (p = 0.340)	**Dynamic Mission (p = 0.030)**
	Time spent in SCT	Indistinguishable (p = 0.150)	**Dynamic Mission (p < 0.001)**
Secondary Workload	Chat reaction time at 235 seconds	**Checkbox and Radio (p = 0.001 and p = 0.002)**	N/A
	Chat reaction time at 1104 seconds	**Checkbox (p = 0.048)**	N/A
Situational Awareness	Target re-designation accuracy	**Checkbox and Radio (p = 0.003 and p = 0.019)**	**Dynamic (p = 0.013)**
	Chat question accuracy	**Checkbox (p = 0.046)**	Indistinguishable (p = 1.000)
Subjective Ratings	Performance	**Checkbox (p < 0.001)**	Indistinguishable (p = 0.830)
	Confidence	**Checkbox (p = 0.002)**	Indistinguishable (p = 0.291)
	Workload	Indistinguishable (p = 0.413)	Indistinguishable (p = 782)
	Satisfaction with AP plans	Indistinguishable (p = 0.254)	Indistinguishable (p = 0.197)

It should be noted that a large number of statistical tests were used in the analysis of data from this experiment, due to the number of dependent variables and 3x2 nested experimental design. In a generous accounting of the number of tests, where each omnibus test is counted as a single "test", and where n pairwise comparisons after a significant omnibus test are only counted as n-1 tests, there were approximately 36 tests. There is an inherent danger in conducting a large amount of statistical tests, as it has an impact on the family-wise error rate. As opposed to having 95% confidence in the conclusions of each test, when $\alpha = 0.05$, the actual confidence level goes to 0% with 36 tests. A confidence level of 95% implies that there is a 1 in 20 chance of a Type I error, thus with 36 tests, it is likely that 2 tests will be false positives.

Utilizing the Bonferroni procedure [74], it can be shown that to obtain a family confidence coefficient of at least 95%, each test must achieve a confidence coefficient of $1-\alpha/g$, where g is the number of tests. Thus, only tests which are significant at the $\alpha=0.0014$ level should be considered significant. Taking this into account, only a few of the statistical tests on the dependent variables would remain significant. The results would still show that operators using the Checkbox interface rated their performance and confidence higher than operators using the None interface ($p < 0.001$ in both cases). Also, there would still be significant differences in target finding efficiency and average time spent in the SCT between the Dynamic and Standard missions ($p < 0.001$). Finally, operators using the Radio objective function made significantly more total modifications to the objective function and made more modifications per SCT session as compared to operators using the Checkbox objective function ($p < 0.001$). For this analysis, however, statistical tests that were significant at the $\alpha = 0.05$ level will still be recognized.

Chapter 6 provides further discussion of these results and evaluation of the results in the context of the experiment hypotheses.

6 Discussion

This chapter discusses the results presented in Chapter 5 and compares them to the hypotheses outlined in Chapter 4. Performance, workload, and situational awareness results are compared across the different Objective Function Types and analyzed in relation to the model presented in Chapter 3. Subjective responses gathered through surveys are reported and evaluated. The effect of changing Rules of Engagement is analyzed. Throughout the chapter, operator strategy and demographic predictors of performance are discussed.

6.1 Performance and Situational Awareness

Performance was characterized by overall mission performance metrics, adherence to the ROEs, and by efficiency metrics, as described in section 4.2.1. Situational awareness was measured by the accuracy of responses to chat box queries and the accuracy of re-designating unknown targets based on chat message information.

The results did not indicate any statistically significant differences in the overall mission performance metrics among the different types of objective function at the $\alpha = 0.05$ level. In comparing the situational awareness of the operators, which has been shown to be an important attribute in operator performance [40, 75], the results show that operators using the Checkbox objective function had significantly higher target re-designation accuracy and chat accuracy than the operators using the None objective function. While the addition of the capability to modify the objective function did not significantly increase system performance, as predicted in hypothesis 1, it may in fact have enhanced SA.

It is likely that the use of the Checkbox interface, which supports multi-objective optimization and provides the operator with a choice of objectives to optimize, enhanced operator SA. Level 1 SA, perception of changes in the environment, is supported by the multi-objective function because it encourages operators to maintain awareness of changes to either the environment or the mission goals to align the objective function with these changes. Level 3 SA, projection of future states, is also supported by the multiple objective function because the use of this objective function best aids operators in understanding what UV actions will result from a selected plan.

In terms of the efficiency metrics that characterize performance over time, the results indicated that operators using the Radio objective function had significantly better target finding efficiency as compared to operators using the None objective function, whereas there was no significant difference for hostile destruction efficiency. This supports hypothesis 2, which predicted that there would be an increase in mission efficiency with the use of a dynamic objective function. A similar result was found in terms of following the ROEs, which guide the operator's high level decision-making by indicating what is most important to accomplish and what is restricted during each time period. Operators using the Radio objective function found more targets in the first 5 minutes of the Dynamic mission, which was one of the primary goals set by the ROEs. These results support hypothesis 3, which predicted that providing the operator with a dynamic objective function would enhance the operator's ability to perform the specified objectives in the ROEs.

It is likely that the Radio objective function, which requires the operator to choose a single objective to optimize, is best for adhering to a single mission goal, such as finding targets as fast as possible. By providing the capability to directly modify the goals of the optimization algorithm, the objectives of the automated planner and the operator were aligned towards this single mission. The plans that the automated planner selected for the operator to review were likely very focused on this single objective, removing several mental steps from the human-automation collaboration process discussed in Chapter 4 and resulting in superior pursuit of the mission objective.

There was, however, a tradeoff between performing the specified mission goals in the ROEs and adherence to the restrictions of the ROEs. During the Dynamic mission, the only 3 operators who violated the ROEs by destroying a hostile target during the first 10 minutes of the mission were operators using the Radio objective function. It is unclear whether these mistakes were due to lack of experience with the system, insufficient training, poor system design, or the increased number of modifications to the objective function necessary when using the Radio objective function.

Additionally, a number of significant predictor variables for performance metrics were found based on demographics and operator strategy. In terms of demographics, lower educational levels predicted higher area coverage. It is possible that undergraduate students were more familiar with mathematical optimization algorithms and therefore were more comfortable with manipulating objective functions to achieve greater area coverage. It is also possible that students above the undergraduate level were exhibiting automation bias [32], through poor understanding of how the automation generated plans or whether a plan would lead to better area coverage. In terms of objective function manipulation strategy, operators that were more parsimonious with the number of objective function modifications that they made per SCT session had higher area coverage and higher percentage of time that targets were tracked. By modifying the objective function fewer times per SCT session, operators likely had more time in the Map View to observe the vehicles and targets, leading to better decision-making. In contrast, the overall number of objective function modifications did predict higher area coverage, as human guidance of automated planners has been shown to enhance search [28].

It has been shown in these results that providing the operator with the ability to modify the objective function of the automated planner could enhance performance, especially if the operator has multi-objective optimization choices, but could also increase the likelihood of mistakes if the operator is limited to single-objective optimization.

6.2 Workload

Workload was measured via an objective workload metric of operator utilization, a secondary workload metric that measured spare mental capacity, and a subjective workload measure intended to capture the mental workload that participants associated with each mission. There were no significant differences among the different objective function types in operator utilization or in the participants' self-rating of how busy they were. It was found that there was no significant difference in average time spent in the SCT among the three types of objective function, contradicting hypothesis 5, which predicted less time spent replanning when using a dynamic objective function.

It should be noted that Radio objective function operators had a higher percentage of SCT sessions where they modified the objective function at least once, made double the total number of changes to the objective function, and had a higher average number of modifications per SCT session. Based on these metrics, it appears that operators may have been working harder, although this workload difference was not reflected in the time spent replanning. Although it has been shown that time spent on a task can be an effective predictor of mental workload [16, 40], it is not a perfect correlation, in that a task can require more cognitive resources without a change in task execution time.

Also, although subjective workload measures have been used effectively in previous human supervisory control experiments [59, 76] where they have been shown to be a reliable indicator of cognitive workload, these measures are difficult to employ because people rate their own workload differently. The objective function type was a between-subjects factor in this experiment, adding to the difficulty in comparing subjective workload evaluations.

An additional method of measuring cognitive workload was through reaction times to accomplish embedded secondary tasks. The results show that at two points during the dynamic mission, operators using the Checkbox objective function had significantly faster reaction times to a secondary task than the operators using the None objective function. At one of those points, the operators using the Radio objective function were also significantly faster. As shown in previous research [72], an embedded secondary tool can provide an effective indicator of workload by measuring the spare mental capacity of the operator. These results could indicate that at certain points during the mission, operators with access to a dynamic objective function were less overloaded than operators using a static objective function. This higher level of spare mental capacity could indicate that the dynamic objective function reduced the operator's mental workload, which is consistent with hypothesis 4, predicting a reduction in mental workload with use of a dynamic objective function.

6.3 Subjective Responses

Participants were asked to rate their performance, confidence, and satisfaction with the plans generated by the automated planner on a Likert scale from 1-5. Participants were also given open-ended questions to prompt them to give general feedback (Appendix I). The responses pertaining directly to collaboration with the automated planner through a dynamic objective function, as well as other comments about the experiment and interface as a whole, are discussed here.

Results indicated that operators using the Checkbox objective function had significantly higher confidence and performance self-ratings than both the Radio and None objective function. These results are consistent with hypothesis 6, which stated that use of a dynamic objective function is expected to result in greater operator satisfaction with the plans generated by the automated planner and higher self-ratings of confidence and performance. There was, however, no significant difference in the ratings for operator satisfaction with the plans generated by the automated planner. All of these measures are between-subjects, as each participant only interacted with a single objective function. Therefore, the subjective self-ratings were isolated evaluations of the objective functions instead of a direct comparison. Despite this issue, the use of a dynamic objective function likely contributed to increased automation transparency and decreased "brittleness," which led to these operator preferences. Although the potential for bias towards the static objective function due to its simplicity was acknowledged as a possibility in section 4.2.3, this bias was not apparent in the results.

The Radio objective function limited operators to choosing only one of the five quantities (area coverage, search/loiter tasks, target tracking, hostile destruction, fuel efficiency) at a time to be their highest priority for evaluating plans. The Checkbox objective function enabled operators to choose any combination of these quantities as high priority. By providing operators using the Checkbox objective function with multi-objective optimization and the capability to communicate their goals to the automated planner, it reduced the number of times that the operator had to modify the objective function of the automated planner. The operators using the limited Radio objective function only had single objective optimization capabilities and were forced to perform numerous "what-if's" on the objective function, more than double the modifications of Checkbox operators, to obtain acceptable plans from the automated planner.

This may indicate why operators using the Checkbox objective function generally rated their confidence and performance higher.

Beyond quantitative subjective data, qualitative evaluations of the system and experiment were also obtained from all participants. Ninety-seven percent of participants indicated that they understood the changes in the ROEs and how to manipulate the system to adhere to the new rules. Also, 87% of participants felt that the automated planner was fast enough for this dynamic, time-pressured mission. Four of the 10 participants who used the Radio objective function complained in writing about the restriction to only select one variable as their top priority and more complained verbally during training. This feeling of restriction in objective function choice is likely related to the lower subjective ratings of the Radio objective function.

As was shown in previous experiments [13], a common complaint from participants was a desire for increased vehicle-level control, as opposed to only task-level control. Fifty-three percent of all subjects wrote about wanting to manually assign vehicles to certain tasks because they disagreed with an assignment made by the automated planner. These comments could be due to the fact that the automated planner was taking into account variables that the human did not comprehend, such as the need to refuel soon, or the speed or capabilities of the vehicle. The participants were also frustrated because of sub-optimal automation performance, as one participant wrote, “the automated planner is fast, but doesn’t generate an optimal plan” while another wrote, “I did not always understand decisions made by the automated planner…namely it would not assign tasks…while some vehicles were seemingly idle.” Finally, one participant wrote, “the automated planner makes some obviously poor decisions…I feel like a lot is hidden from me in the decision making…I felt like I had to trick it into doing things.”

Three of the 20 participants who used one of the dynamic objective functions noted that although they were told that the weightings of each variable were the same if that variable was checked, the automated planner seemed to favor certain variables over others. This could once again be due to sub-optimal automation performance or design and should be investigated in further research.

6.4 Changing Rules of Engagement

Although not a primary focus of this research, it was shown that the second independent variable in the experiment, Mission Type, was a significant factor in the analysis of many of the dependent variables. For the Standard Mission, the ROEs were presented to the operator once at the start of the mission and did not change. For the Dynamic Mission, every 5 minutes during the 20 minute mission, new ROEs were presented to the operator. These ROEs gave the operator guidance on what was most important to accomplish during that time period and what actions they were restricted from taking.

As can be expected, operators conducting the more complicated Dynamic mission had significantly higher utilization and spent significantly more time in the SCT on average. An interesting and unexpected result was that regardless of the objective function used, operators found significantly more targets and had higher target finding efficiency in the Dynamic mission as compared to the Standard mission. Additionally, operators had significantly higher accuracy in the re-designation of unknown targets in the Dynamic mission, which is a measure of SA.

Despite the fact that operators were working harder during the Dynamic mission, they also performed better. It is possible that the scenarios designed for each mission, which had different target locations and paths, were of different perceived difficulty levels despite the fact that they were designed to be of comparable difficulty. Another possibility is that more frequent reminders of mission goals, through the changing ROEs, could have played a role in this increase in performance. The ROE changes provided more specific goals to the operator, guiding them in how to conduct the mission, which led to higher performance. The ROE changes influenced the internal objective function of the human operator, who then communicated his or her objectives to the automated planner, which generated new plans for the vehicles, subject to the operator's approval. Further research is necessary to evaluate whether more frequent reminders of goals can lead to higher performance in an unmanned vehicle supervisory control setting.

6.5 Summary

Results from the human performance experiment provided insight into methods of collaboration between a human operator and automated planner for conducting supervisory control of a network of decentralized UVs. The results indicated that the original hypotheses were generally correct, in that providing an operator with the ability to modify the weightings of the variables in the objective function of an automated planner resulted in enhanced SA, increased spare mental capacity, and increased subjective ratings of the human-automation collaboration. There were caveats to these results, including the fact that target finding efficiency and adherence to changing mission objectives increased with use of a single-objective optimization function, but some operators violated the ROEs while using this single-objective function.

One potential confound in this experiment is that by the nature of the experiment, operators should have been able to adhere to changing mission objectives better with a dynamic objective function. Theoretically, a static objective function would be inferior if the mission goals and ROEs were changing throughout the mission. Therefore, a comparison between a static and dynamic objective function in terms of adherence to changing mission goals may be unfair. It is clear, however, that the dynamic objective function with multiple objective optimization capabilities resulted in superior SA, spare mental capacity, and subjective ratings.

In addition, the results provided new information on the impact of changing mission goals on human-automation collaboration. While it was expected that changing mission goals would cause a higher cognitive workload, the results indicated that operators also had higher SA and performed better in terms of finding new targets. Further research is necessary to analyze the impact of changing mission goals on the human operator and how they influence overall system performance.

Two methods of implementing a dynamic objective function were implemented and compared, one with single objective optimization and one with multiple objective optimization. By providing the operator with more choice in communicating his or her goals to the automation, through multi-objective optimization, the operator could communicate to the automation faster, did not have to work as hard, and felt more confident about his or her actions.

Finally, the results have shown an interesting trend that increasing levels of education predicted lower system performance. A controlled experiment investigating the impact of education level on multiple UV supervisory control would need to be run to draw any substantial conclusions on this topic. It is, however, of interest to current military operations, where the demand for increased UAV missions is driving a trend towards placing enlisted military personnel in UAV operator roles.

Chapter 7 will discuss the implications these results have on the initial research objectives and the design of future collaborative UV systems.

7 Conclusions

There is an increasing demand to use UVs for a variety of civilian and military purposes. To keep up with this demand, as well as reduce the expense of operating UVs and enhance the capabilities of UVs through better coordination, human operators will need to supervise multiple UVs simultaneously. In order to successfully conduct this form of supervisory control, operators will need the support of significant embedded collaborative autonomy. Automated planners are useful in this mission, as they are more effective than humans at certain aspects of path planning and resource allocation in time-pressured, multivariate environments. While reducing the need for manual control and allowing the operator to focus on goal-based control, automated planners can also be "brittle" when dealing with uncertainty, which can cause lower system performance or higher workload as the operator manages the automation. Therefore, this research was motivated by the desire to reduce mental workload and maintain or improve overall system performance in supervisory control of multiple UVs.

The design and testing of an interface to provide an operator with the ability to modify the objective function of the automated planner demonstrated the potential for new methods of human-automation collaboration in UV control. A dynamic objective function increases the transparency and reduces the "brittleness" of the automated planner, which enhances the ability of a human operator to successfully work with the automation. It provides the operator with a convenient method to communicate his or her goals to the automation, especially in light of changing mission goals.

7.1 Research Objectives and Findings

The objectives of this research were to determine the motivating principles for dynamic objective function manipulation, develop an interface to provide operators with this capability, and to evaluate the effectiveness of real-time human manipulation of the objective function of a scheduling and resource allocation algorithm. The goal was to address these objectives through the following methods:

- Review current research in human-computer collaboration for scheduling, resource allocation, and path planning, in order to develop a theoretical model of dynamic objective function manipulation (Chapter 3).
- Design a dynamic objective function tool and integrate the tool into an existing multi-UV mission simulation testbed (Chapter 4).
- Use a human performance experiment to evaluate the impact of real-time human manipulation of a dynamic objective function on system performance, workload, and subjective appeal (Chapters 4-6).

The review of previous research in Chapter 2 motivated this research by revealing gaps in the human-automation collaboration literature, including the lack of experiments featuring a dynamic and uncertain environment, time-pressure for decision-making, and methods for enabling an operator to express his or her desires to the automated planner. The human-automation collaboration model that was extended in Chapter 3 to include the concept of objective function manipulation illustrated the many cognitive steps that are involved in generating, evaluating, and selecting plans for multiple UV control. The model also showed the potential for a reduction in the number of cognitive steps required to evaluate plans through the use of a dynamic objective function. Chapter 4 introduced the dynamic objective function tool that was developed and integrated into an existing simulation testbed. The impact of real-time human manipulation of a dynamic objective function was evaluated through a human performance experiment.

The results of this experiment established that a dynamic objective function with a single objective improved adherence to changing mission priorities, but also led to ROE violations. It is possible that the single objective method assisted in causing the violations, either because the operators were focused on a single objective or because the method required extensive interaction to achieve an acceptable plan, increasing the chance of error. Secondary results of the experiment indicated that changing mission goals, as expected, caused higher cognitive workload, but unexpectedly resulted in superior performance and higher SA. Additionally, an undergraduate education was shown to be a predictor of higher system performance over higher levels of education.

Finally, operators using a dynamic objective function with multi-objective capabilities needed fewer modifications to the objective function to achieve an acceptable plan, had enhanced SA, and had increased spare mental capacity, indicating lower workload. One of the most revealing results of the experiment were the subjective ratings of the interfaces, showing that operators clearly preferred the dynamic objective function with multi-objective capabilities, which gave them the most flexibility in communicating their goals and desires to the automated planner. Developing an appropriate level of trust between the human and automated planner is crucial for successful human-automation collaboration [77], and providing the capability to modify the objective function for multi-objective optimization can aid in developing this trust.

7.2 Recommendations and Future Work

Though the results of this thesis indicate that dynamic objective function manipulation shows potential for improved performance with reduced mental workload and increased subjective appeal in a human-automation collaboration for multi-UV control, further investigation is required. The following are recommendations for future work based on the research presented in this thesis:

- As described in Section 5.1, a system-level re-design of the interface for the OPS-USERS testbed is required to incorporate the concept of changing ROEs. The interface was originally designed assuming that the destruction of hostile targets would always be permitted, which is why the only options provided to an operator when asked to approve the destruction of a hostile target are to either approve the destruction or re-designate the target as unknown. Adding the capability to designate a hostile target as "ineligible for destruction" or a way to remind the operator that a target was re-designated from hostile to unknown would be helpful.
- An additional design recommendation for the OPS-USERS testbed, based on suggestions from participants, is to develop additional methods to provide feedback to the operator about why a task could not be assigned. Often times, constraints in available UVs, the time required to travel to a task's location, re-fueling constraints, or the time required to conduct a task causes the automation to reject a task that the operator attempted to assign in a "what-if" query. If the reason for the rejection could be communicated to the operator visually and/or verbally, it would decrease operator frustration with the automation.
- A direct method of obtaining subjective user feedback that directly compares the various objective function types should be considered. This would result in a within-subjects experimental design where each participant conducts multi-UV missions with each of the objective functions.

- Further investigation of the types of dynamic objective functions that can be implemented is warranted. More options for manipulating the values of the weightings in the objective function should be investigated, as opposed to just allowing goal manipulation at a binary level of "on" or "off." For example, rating each value as "high," "medium," or "low" or ranking the values in priority order could be explored.
- It is unclear from this thesis whether the changing ROEs guided the human in how to conduct the mission, leading to enhanced performance, or whether it was simply the act of reminding the operator of his or her goals that led to superior performance. An experiment could be run to determine whether more frequent reminders of goals leads to enhanced performance.
- It remains an open question whether the participants simply set the objective function weightings better than the a priori coded objective function, or whether the operator's manipulations of the objective function actually took the system performance beyond a level that could be achieved autonomously. Further investigation is necessary to determine the optimal settings for the objective function of the automated planner. This would require, for example, Monte Carlo simulations using a recently developed human operator model [78] to work with the automated planner. This would be difficult to pursue, however, for 2 reasons: 1) the definition of "optimal" will be very difficult to define in a complex command and control scenario and 2) the dynamic and uncertain nature of the simulation may prevent the development of an optimal policy for the objective function weightings.

Appendix A: Demographic Descriptive Statistics

Category	N	Min	Max	Mean	Std. Dev.
Age (years)	30	18	38	21.30	3.98
Rating of past 2 nights of sleep (1-4)	30	1	4	2.23	0.82
Rating of TV watching (1-5)	30	1	5	2.30	0.99
Rating of gaming experience (1-5)	30	1	5	2.37	1.25
Rating of comfort level with computers (1-4)	30	2	4	3.40	0.68
Rating of perception of unmanned vehicles (1-5)	30	2	5	3.80	0.85
Occupation (Student/Other)	Undergraduate: 18 Masters: 6 Ph.D: 4 Non-student: 2	-	-	-	-
Military experience (Y/N)	1/29	-	-	-	-
Gender (M/F)	21/9	-	-	-	-

Appendix B: Consent to Participate Form

CONSENT TO PARTICIPATE IN NON-BIOMEDICAL RESEARCH

OPS-USERS Dynamic Human-Computer Collaboration in Real-time Unmanned Vehicle Scheduling

You are asked to participate in a research study conducted by Professor Mary L. Cummings Ph.D and Andrew Clare, S.B., from the Aeronautics & Astronautics Department at the Massachusetts Institute of Technology (M.I.T.). Results from this study will contribute to a Master's Thesis by Andrew Clare. You were selected as a possible participant in this study because the expected population this research will influence is expected to contain men and women between the ages of 18 and 55 with an interest in using computers. You should read the information below, and ask questions about anything you do not understand, before deciding whether or not to participate.

- **PARTICIPATION AND WITHDRAWAL**

Your participation in this study is completely voluntary and you are free to choose whether to be in it or not. If you choose to be in this study, you may subsequently withdraw from it at any time without penalty or consequences of any kind. The investigator may withdraw you from this research if circumstances arise which warrant doing so.

- **PURPOSE OF THE STUDY**

The purpose of this study is to investigate the effect of different methods of collaboration between a human operator and an automated mission planner on overall mission performance.

- **PROCEDURES**

If you volunteer to participate in this study, we would ask you to do the following things:

- Participate in training and practice sessions to learn a video game-like software environment that will have you control a team of simulated unmanned vehicles. The team you will control will be assigned with the task of finding, identifying, and tracking targets in an area of interest within a 20 minute period.
- Practice on the software environment will be performed until an adequate level of performance is achieved, which will be determined by your demonstration of basic proficiency in operating the unmanned vehicles and replanning the mission. (one 15 minute training session)
- Execute two trials consisting of the same tasks as above (Estimated time 50 mins).
- After each trial you will be assigned a score for the trial based on the number of targets you successfully find, how long they are successfully tracked thereafter,

and what percentage of the total area of interest is searched, and how well you follow instructions provided by the chat box.

- All testing will take place at MIT in room 35-220 or in 37-301.
- Total time: 1.5 hours, depending on skill level.

- **POTENTIAL RISKS AND DISCOMFORTS**

There are no anticipated physical or psychological risks.

- **POTENTIAL BENEFITS**

While you will not directly benefit from this study, the results from this study will assist in the design of interfaces for human/unmanned vehicle systems.

- **PAYMENT FOR PARTICIPATION**

You will be paid $15 for your participation in this study, which will be paid upon completion of your debrief. Should you elect to withdraw during the study, you will be compensated for your time spent in the study. The subject with the best performance will be given a reward of a $100 Best Buy Gift Card.

- **CONFIDENTIALITY**

Any information that is obtained in connection with this study and that can be identified with you will remain confidential and will be disclosed only with your permission or as required by law. You will be assigned a subject number which will be used on all related documents to include databases, summaries of results, etc.

- **IDENTIFICATION OF INVESTIGATORS**

If you have any questions or concerns about the research, please feel free to contact the Principal Investigator, Mary L. Cummings, at (617) 252-1512, e-mail, missyc@mit.edu, and her address is 77 Massachusetts Avenue, Room 33-311, Cambridge, MA, 02139. The student investigator is Andrew Clare. He may be contacted at (617) 253-0993 or via email at aclare@mit.edu.

- **EMERGENCY CARE AND COMPENSATION FOR INJURY**

If you feel you have suffered an injury, which may include emotional trauma, as a result of participating in this study, please contact the person in charge of the study as soon as possible.

In the event you suffer such an injury, M.I.T. may provide itself, or arrange for the provision of, emergency transport or medical treatment, including emergency treatment and follow-up care, as needed, or reimbursement for such medical services. M.I.T. does not provide any other form of compensation for injury. In any case, neither the offer to provide medical assistance, nor the actual provision of medical services shall be considered an admission of fault or acceptance of liability. Questions regarding this policy may be directed to MIT's Insurance Office, (617) 253-2823. Your insurance carrier may be billed for the cost of emergency transport or medical treatment, if such services are determined not to be directly related to your participation in this study.

- **RIGHTS OF RESEARCH SUBJECTS**

You are not waiving any legal claims, rights or remedies because of your participation in this research study. If you feel you have been treated unfairly, or you have questions regarding your rights as a research subject, you may contact the Chairman of the Committee on the Use of Humans as Experimental Subjects, M.I.T., Room E25-143B, 77 Massachusetts Ave, Cambridge, MA 02139, phone 1-617-253 6787.

SIGNATURE OF RESEARCH SUBJECT OR LEGAL REPRESENTATIVE

I understand the procedures described above. My questions have been answered to my satisfaction, and I agree to participate in this study. I have been given a copy of this form.

____
Name of Subject

____
Name of Legal Representative (if applicable)

______________________________ ______________
Signature of Subject or Legal Representative Date

SIGNATURE OF INVESTIGATOR

In my judgment the subject is voluntarily and knowingly giving informed consent and possesses the legal capacity to give informed consent to participate in this research study.

______________________________ ______________
Signature of Investigator Date

Appendix C: Demographic Survey

Pre-experiment Survey

Page 1

1. Subject number:_____

2. Age:_____

3. Gender: *M* *F*

4. Occupation:_______________________________

   if student, (circle one): *Undergrad* *Masters* *PhD*

5. Military experience (circle one): *No* *Yes*

   If yes, which branch:________

   Years of service:________

6. Give an overall rating of your past two nights of sleep.

   Poor *Fair* *Good* *Great*

7. On average, how much TV do you watch daily?

   Never watch TV *Infrequently watch TV* *About 1 hour* *About 2 hours* *More than 2 hours*

8. How often do you play computer games?

   Rarely play games *Play games once a month* *Weekly gamer* *A few times a week gamer* *Daily gamer*

   Types of games played:

9. Rate your comfort level with using computers.

   Not comfortable *Somewhat comfortable* *Comfortable* *Very Comfortable*

10. What is your perception toward unmanned vehicles?

    Intense dislike *Dislike* *Neutral* *Like* *Really Like*

Pre-experiment Survey

Page 2

1. Subject number:_____

2. How confident were you about the plans you created?

   Not Confident *Somewhat Confident* *Confident* *Very Confident* *Extremely Confident*

   Comments:

3. How did you feel you performed overall?

   Very Poor *Poor* *Satisfactory* *Good* *Excellent*

4. How busy did you feel during the practice mission?

   Extremely Busy *Busy* *Not Busy* *Idle*

5. Do you understand how to create search tasks?

   No *Somewhat* *Yes*

6. Do you understand how to use the target identification window?

   No *Somewhat* *Yes*

7. Do you understand how to approve a weapon launch on hostile targets?

   No *Somewhat* *Yes*

8. Do you understand how to use the Schedule Comparison Tool (SCT)?

   No *Somewhat* *Yes*

9. Do you understand that you must accept a plan in order for the unmanned vehicles to perform new search, track and destroy tasks?

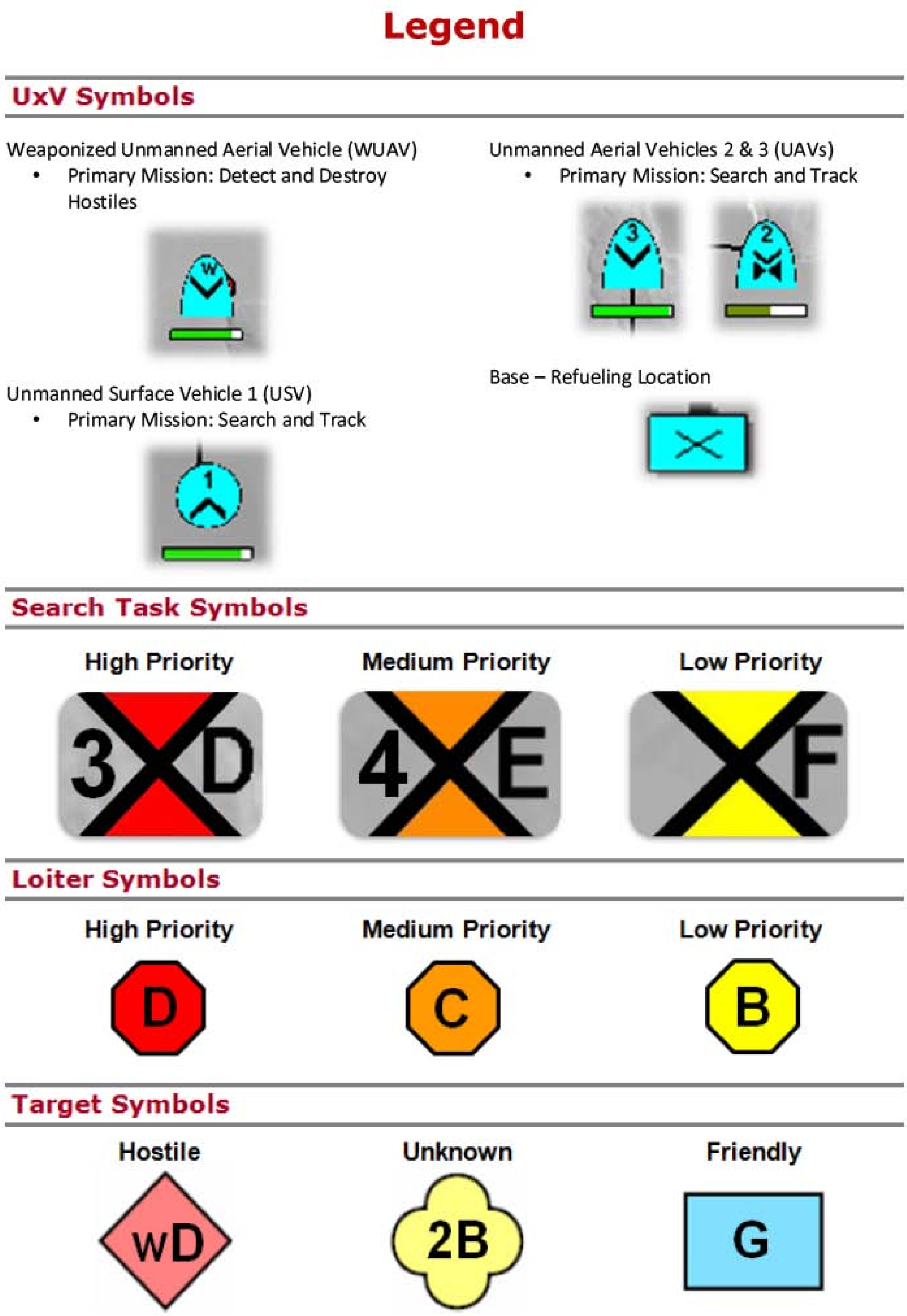
   No *Somewhat* *Yes*

10. Do you understand how to modify the objective function of the automated planner?

    No *Somewhat* *Yes*

Now is the time to ask the experiment administrator any questions you have about the mission or interface.

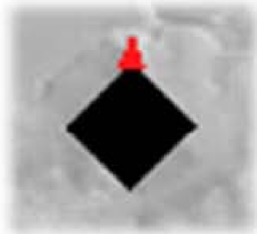
Appendix D: Experiment Legend



Legend

Lost Target

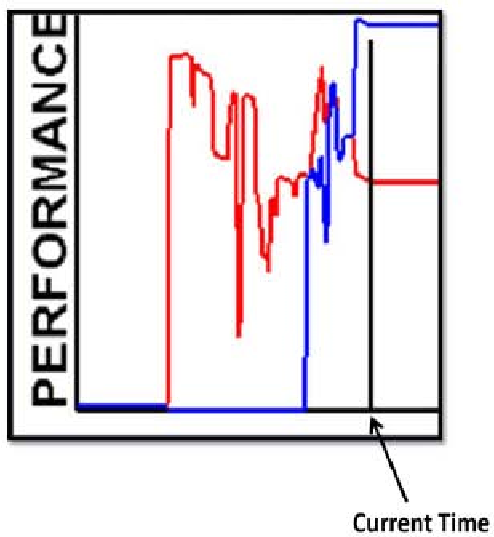
Destroyed Hostile Target



Performance Plot

Red is **Proposed Performance** calculated by the computer algorithm

Blue is **Actual Performance**



Red above Blue, you have work to do.

Blue above Red, you're getting ahead!

Your Mission Score

Overall Mission Score will be calculated by:

- % Area Covered by Mission End
- % Targets Found
- % Time Targets Tracked
- % Hostile Targets Destroyed
- Reaction Time for Replan Prompt
- Chat Box Response Time and Accuracy

Appendix E: Unmanned Vehicle Velocity and Fuel Consumption

Data was obtained on the MQ-1 Predator to aid in setting the cruise and maximum velocities, and cruise and maximum fuel consumption for the UAVs used in the simulation for this experiment. The cruise speed of a Predator is 84 miles per hour, the maximum speed is 135 miles per hour, and the fuel capacity is 100 gallons [79]. The range of the predator is 2,302 miles [80]. The maximum speed to cruise speed ratio for the Predator is approximately 1.6.

The general equation for the drag of a solid object moving through a fluid [81] is:

$$D = \frac{C_d \rho A V^2}{2}$$

D = Drag Force
C_d = Coefficient of Drag
ρ = density of fluid (air in this case)
A = cross-sectional area of the object
V = velocity of the object

This equation reveals that drag increases with the square of speed. Based on the aerodynamics assumption that fuel consumption increases linearly with drag, and the fact that the maximum speed to cruise ratio of the Predator is 1.6, we can calculate that the maximum fuel consumption of the Predator should be approximately 2.5 times the cruise fuel consumption. Speeds and fuel consumptions were set for the UAVs in the simulation to match this 1.6 ratio between cruise and maximum speed and the 2.5 ratio between cruise and maximum fuel consumption, as shown in Table 13. Note that the units of these numbers are based on the simulation environment and not on any real-life units.

Table 13. Velocities and Fuel Consumption for Unmanned Vehicles

Unmanned Vehicle Type	Cruise Velocity	Max Velocity	Cruise Fuel Consumption	Max Fuel Consumption	Fuel Capacity
WUAV	100	160	0.01	0.025	7
USV	25	50	0.01	0.025	3
Fixed-wing UAV	75	120	0.01	0.025	4
Helicopter UAV	75	120	0.01	0.025	4

Appendix F: Rules of Engagement

F.1 Standard Mission

The following Rules of Engagement were sent through the Chat Window to the operator as soon as the mission began and did not change during the 20 minute mission:

- *Track all found targets and destroy all hostile targets found.*

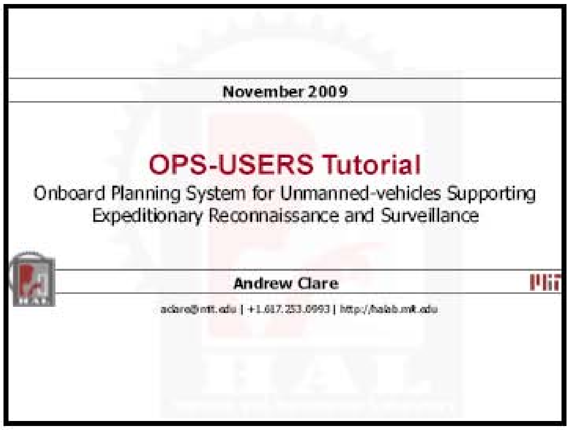
F.2 Dynamic Mission

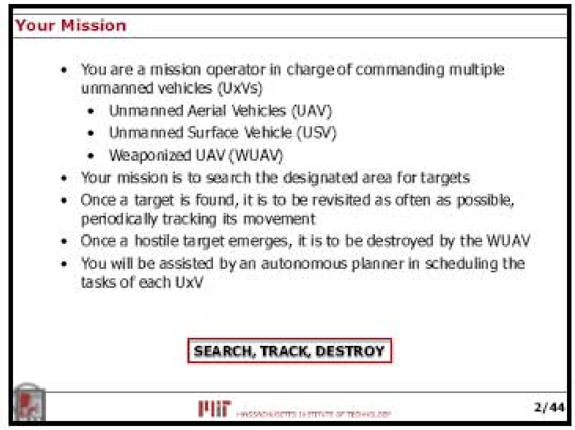
The following Rules of Engagement were sent through the Chat Window to the operator at the specified times:

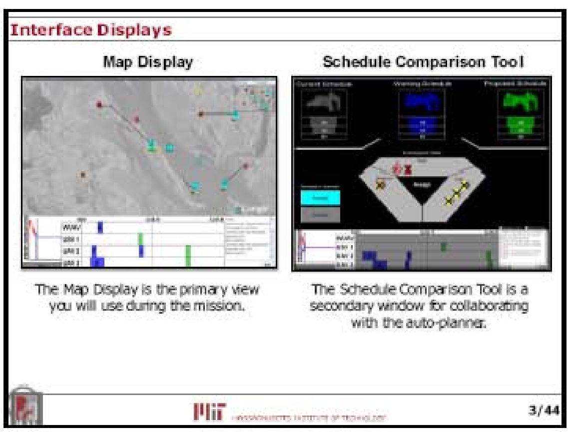
- START: *Cover as much area as possible to find new targets. Tracking found targets is low priority. Do not destroy any hostiles*.
- FIVE MINUTES: *Conduct search tasks in SE and SW Quadrants. 2nd priority: Track all targets previously found. Do not destroy any hostiles*.
- TEN MINUTES: *Track all targets closely - it is important not to lose any targets! 2nd priority: conserve fuel. 3rd priority: destroy hostile targets*.
- FIFTEEN MINUTES: *All Hostile Targets are now high priority - destroy all hostiles!*

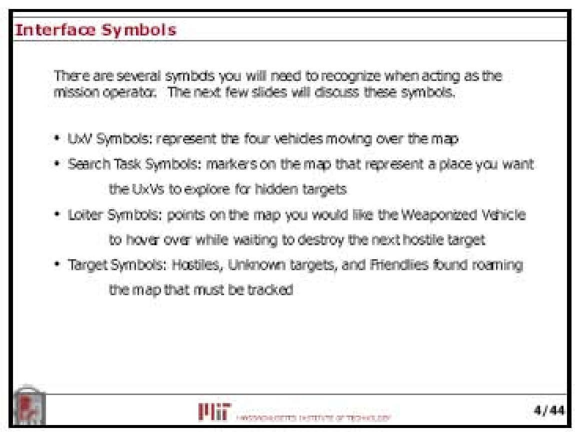
Appendix G: Experiment PowerPoint Tutorials

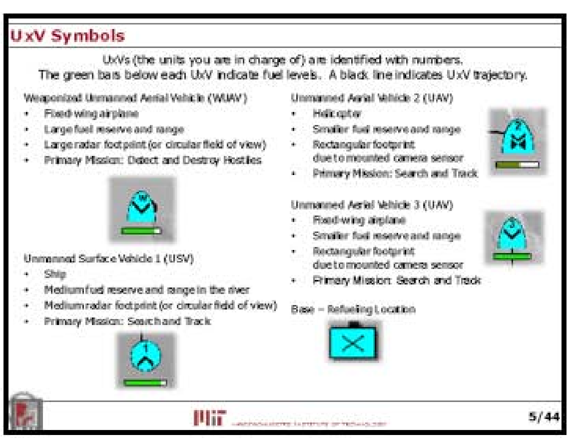
G.1 Static (None) Objective Function Tutorial

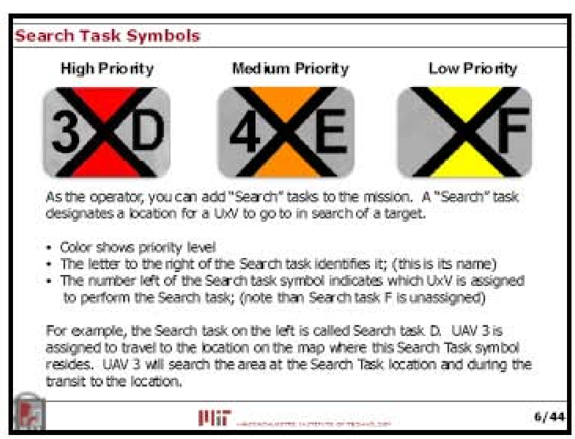












Loiter Symbols

The loiter symbol for the WUAV resembles a stop sign.
The color indicates priority level.

The Weaponized UAV does not attend to search tasks. The auto-planner will not schedule the WUAV to track targets. The mission of the WUAV is to **detect and destroy hostile targets**. The WUAV can be sent to loiter, or hover over a particular position, while waiting to destroy hostile targets. The WUAV can only search the map (increasing the coverage area) on the route to a loiter task. The WUAV will not search on its own like the other UxVs.

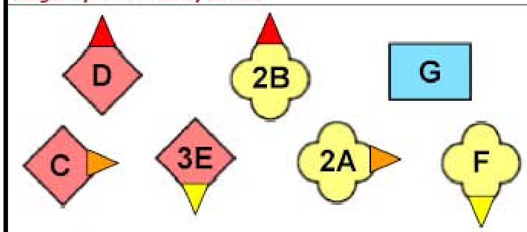
Target Symbols

The UxVs must **periodically track**, or revisit, the targets that have been found. The Weaponized UAV must **destroy hostile targets**.

- Red diamonds are hostile targets
- Yellow clovers are unknown targets
- Blue rectangles are friendly and are not tracked
- The letter on the right identifies the target
- The character on the left indicates which UxV is assigned to the target

For example, according to the hostile symbol on the left, the Weaponized UAV is assigned to destroy hostile target D. Next, the center symbol shows that UAV 2 will track Unknown Target B. UAV 2 will travel to the location where this target symbol is positioned on the map and begin following the target. If UAV 2 has another task to perform or must go back to base to refuel, the algorithm will calculate an estimated new position for the target based on the target's last known position and velocity. Friendly target G, on the right, is not assigned because friendlies are not to be tracked.

Target Symbol Priority Levels

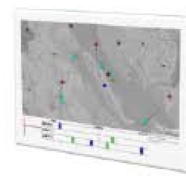


Flags designate target priority level.

- Red vertical flag on top of the target symbol specifies high priority
- Orange horizontal flag beside the target symbol specifies medium priority
- Yellow downward flag below the target symbol specifies low priority
- Friendlies do not have a priority level flag because they do not need to be tracked
  - Only hostile and unknown targets are tracked

Overview of the Map Display

- Getting Started
- Parts of the Map Display
- Selecting a UxV
- Creating or Editing a Search Task
- Selecting an Existing Search Task
- Chat Message Box
- Mini Map
- Identifying a Target
- Selecting a Target
- Lost Targets
- Approving Weapon Launch
- Timeline
- Coverage Area Overlay
- Performance Plot
- Replanning

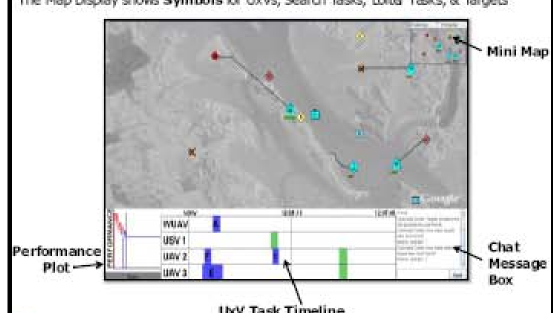


Getting Started

Click the **green start button** in the center of the interface to begin the scenario.

Map Display

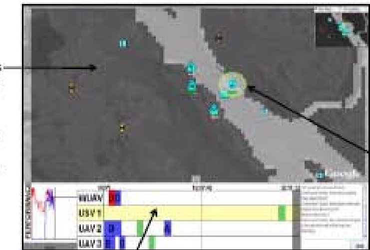
The Map Display shows **Symbols** for UxVs, Search Tasks, Loiter Tasks, & Targets



Selecting a UxV

To select a UxV, left click the UxV's symbol.

UxV's **Obstacles** appear as a dark overlay where the UxV cannot go



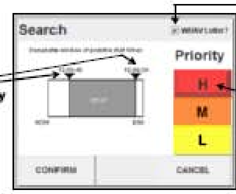
UxV's **Sensor range** appears in lime green

UxV's **Timeline** is highlighted in light yellow

Creating or Editing a Search Task

Right clicking a location on the Map View brings up the Search Task Creation Window. Right click an existing Search Task to edit it using the same window.

Drag the ends of the timeline to **specify window of opportunity** for when you want the task to begin.



Use this same window to **create or edit a WUAV Loiter Task.** Select the WUAV Loiter checkbox.

Designate the priority level by clicking the appropriate button. The priority is used by the auto-planner to add the task to the schedule.

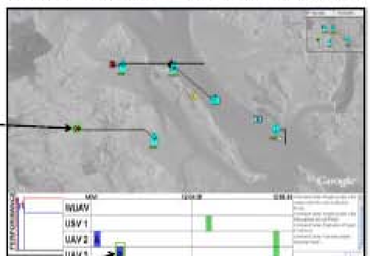
You may choose to create a Search Task at anytime. Sometimes the Chat Message Box will prompt you to create a search task and will specify the window of opportunity.

If a Search Task expires or becomes completed, a yellow message in the lower-right corner of the Map View will notify you of that event.

Selecting an existing Search Task

Click on an existing Search Task Symbol to select it.

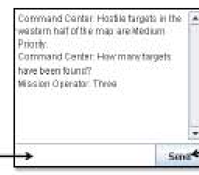
Task **symbol outlined** in lime green when selected



Task **outlined in the timeline** of the UxV assigned to perform the task (if assigned)

Chat Box: Intelligence Information

Command Center sends intelligence information through chat box instant messaging. The Chat Box gives important information dictating priority levels for targets.



message input window

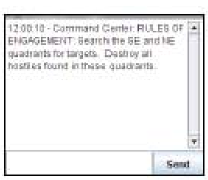
Send button

When a message comes in, you will hear a tone and the black outline of the box will blink. You can **click in the Chat Box** to stop it from blinking.

Some messages require responses. Click in the **message input window**. Type your message. Then hit enter or click the **send button.** One-word answers or numbers are preferred.

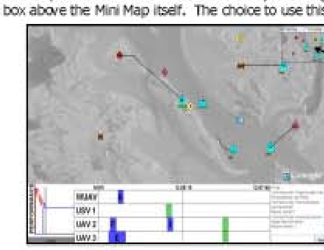
Chat Box: Rules of Engagement

Command Center also sends changes to the **Rules of Engagement** through chat box instant messaging. These Rules of Engagement (ROE) dictate what is most important for you to accomplish. You should adjust your strategy for managing and tasking your vehicles whenever you receive new ROE. Your performance score depends on how well you execute the mission described by the ROE. If you receive new ROE, the old ROE no longer apply.



Mini Map

The upper right-hand corner of the Map Display is equipped with a Mini Map that shows the symbols for UxVs, search and loiter tasks, and targets as they appear on the map. Since you can zoom in on the big map the using roller wheel on your mouse, it is convenient to glance at the Mini Map for a quick view of the overall picture. The Mini Map feature can be turned on or off by checking or un-checking the Mini Map box above the Mini Map itself. The choice to use this feature is yours.



Mini Map

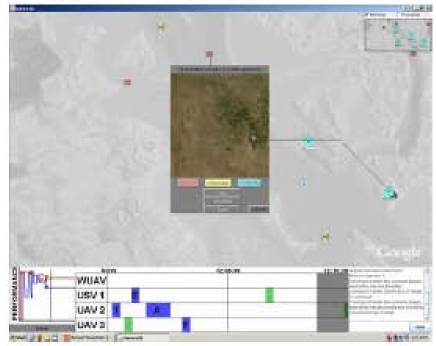
Identifying a Target

In real search & track missions, when a target is found, an operator must identify what type of target has been found. To simulate this, a Target ID Window pops up automatically when a target is found. The Chat Box will provide information on target priority level. You can drag the Target ID Window around the screen.

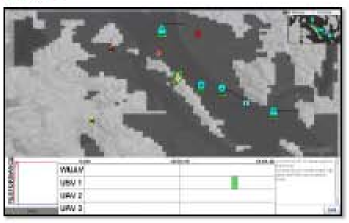
Initially target may not be visible. | Click and drag over the area to pan for the target symbol. | Click the appropriate target designation button. | Click a priority level button.

Unknown targets must first be marked as "Unknown," but the designation can be edited afterwards as more information arrives from the Chat Box.

Target ID Video—Click on the Picture to Watch



Selecting a Target



Left clicking a target will cause the following:

- Target's obstacles appear (targets can be land or water vehicles)
- Target's likely position is shown by blue concentric probability rings (if the target has not been visited in a while)
- Target highlighted in the timeline of the UxV assigned to perform it (if assigned)

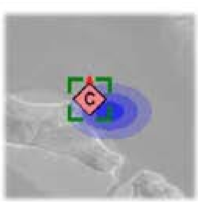
Right clicking a Target will cause the following:

- Target ID Window appears for editing designation or priority level

Changing Target Designation Video

Right clicking a target allows you to edit the target designation and priority level. The video below shows how to change an unknown target into a hostile, high priority target. The unknown target has been right clicked to bring up the Target ID Window. **Click the picture to watch.**

Target Concentric Rings for possible new location



Blue Concentric rings appear automatically for targets that have not been visited in a while that might soon become lost targets. The blue concentric rings shown next to the target symbol represent the algorithm's best estimate of where the target might be. The darker the blue, the more probable the position of the target. A UxV should be assigned a search task near this location ASAP, lest the target become lost.

Lost Targets

Targets become lost if UxVs do not track them for a while.
Lost targets appear dimmer. The probability distribution (blue concentric rings) appear before the target becomes lost and dims.

Compare the gray color of lost Target L on the left to Target A on the right that is not lost.

Lost Target

Target (not lost)

It can take anywhere from 30 to 120 seconds for a target to become lost, depending on target's speed and vehicle type. Revisiting targets more often prevents targets from becoming lost.

A lost target is removed from the map and the schedule queue after the UxVs have revisited the lost icon five times with no success of finding the lost target.

Approving Weapon Launch

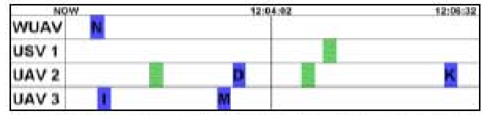
Weapon Launch Approval Window

Destroyed Hostile Target

Destroyed targets appear as black symbols on the Map View.

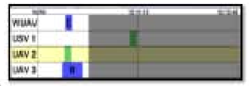
Operator approval must be given before the WUAV is allowed to destroy a hostile target. The Missile Launch Approval Window pops up automatically when the WUAV sights the hostile target for destruction and a second UxV has the hostile in its sights as the "eyes on." Pan the screen for a direct view of the target, and click the red **Approve Launch** button to destroy the target.

Timeline



The timeline gives temporal information for each UxV for the next **ten minutes into the future** indicated in the format hours:minutes:seconds.

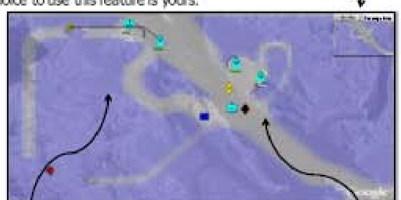
Green bars in the Timeline indicate times of refueling, and blue bars indicate times performing a task. The letter of the task (whether Search Task or Target Track Task) appears in the blue bar. Each UxV is limited to two task assignments at a time. White space indicates idle time or time traveling between tasks. The timeline marches forward as time goes on.



The timeline will begin to gray-out, indicating the end of the mission. The mission end time occurs at the border from white to gray.

Coverage Area Overlay

The Probability Distribution Overlay can be turned on by selecting the **Coverage Area checkbox**. The choice to use this feature is yours.



The **periwinkle blue** coloring over the map shows the area that has not been visited recently. These blue areas are more likely to have hidden targets because a UxV has not visited there yet.

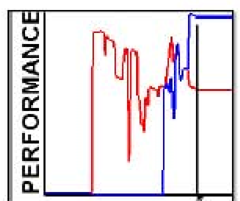
The typical **gray** coloring shows a path that was just visited by a UxV, and this coverage area most likely does not contain hidden targets.

Performance Plot

The Performance Plot shows the auto-planner's computation of the current schedule's performance in comparison to the actual overall performance of the current schedule. The plot marches forward as time goes on, much like the Timeline.

Red is **Proposed Performance** calculated by the computer algorithm

Blue is **Actual Performance**

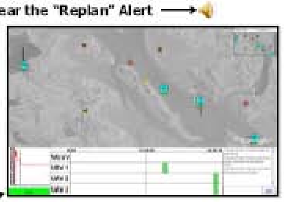


Red above Blue, you have work to do. Blue above Red, you're getting ahead!

Current Time

Replan

The Replan button in the lower left corner of the Map Display will turn green when a new plan is available from the auto-planner. You will hear the auditory "Replan" alert. **Click to hear the "Replan" Alert**



As the Replan button turns green and you hear the alert, a new proposed schedule is available that requires you to hit the replan button to see it.

The auto-planner prompts you to replan when it has generated a better schedule for the UxVs. New targets and tasks in the scenario often result in a new proposed schedule.

Overview of the Schedule Comparison Tool

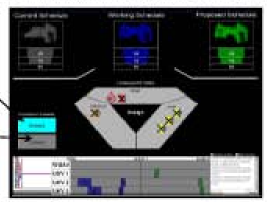
- Schedule Comparison Tool Display
- Configural Display
- Number on Bars Feature
- Multi-Color Bars Feature
- Assigning Unassigned Tasks
- Nothing to Assign
- Selecting a Task
- Switching Schedules

Schedule Comparison Tool (SCT)

Once you click the replan button, the Schedule Comparison Tool (SCT) appears, showing a performance overview for three schedules: the **Current Schedule** (gray), the **Working Schedule** (blue), and the **Proposed Schedule** given by the automation (green).

After using the SCT, you must click **Accept** in order to implement the working schedule you created and to return to the Map Display.

You may always **Cancel** and return to the Map Display. However, the proposed plan will not be communicated to the UxVs.

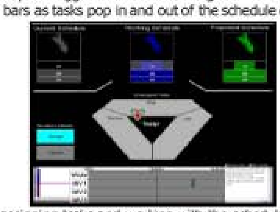


In the SCT interface, the timeline has been grayed out to indicate that it is the timeline for the "Current Schedule" and will only correspond to the "Current Schedule."

You may click the Replan button and go to the SCT anytime you like, even without being prompted.

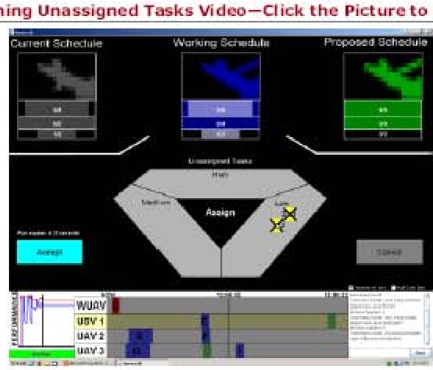
Assigning Unassigned Tasks

Assign a task by dragging it into center "Assign" area. Doing this queries the autonomous planner to see if the task can be assigned or not. If the task does not pop back out, the auto-planner was able to schedule the task. Sometimes a different task pops out because the auto-planner had to un-assign that task in order for the task you dragged in to become assigned. This will be reflected in the hierarchical bars as tasks pop in and out of the schedule query "Assign" area.



After you finish assigning tasks and working with the schedule, click Accept to implement the plan (or Cancel), and you will return to the Map Display.

Assigning Unassigned Tasks Video—Click the Picture to Watch



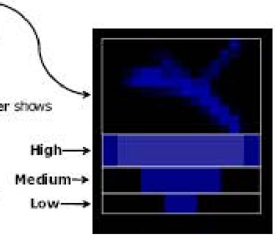
Configural Display

The **top rectangle** represents the "map" area that will be covered for a given schedule.

The more colorful the area, the better searched it is.

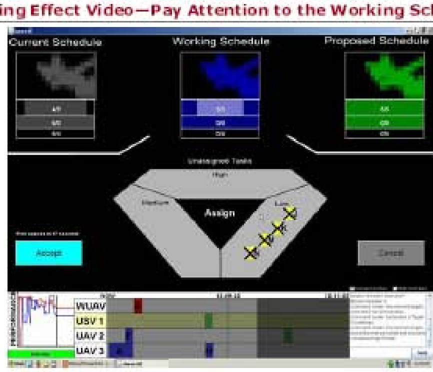
The **bottom hierarchical ladder** shows the percentages of high, medium, and low priority tasks to be completed for a given schedule.

The more color-filled a rectangle is, the more of that task-priority is being done.

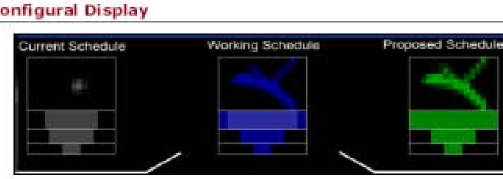


When a task is assigned, the corresponding ladder rung rectangle changes size in a "**ghosting effect**" to visually show what has changed. The darker color is the new size, whereas the lighter overlay is the old size.

Ghosting Effect Video—Pay Attention to the Working Schedule



Configural Display



Area Covered

Comparing these three schedules in the above example: both "Working" and "Proposed" cover more area than "Current."

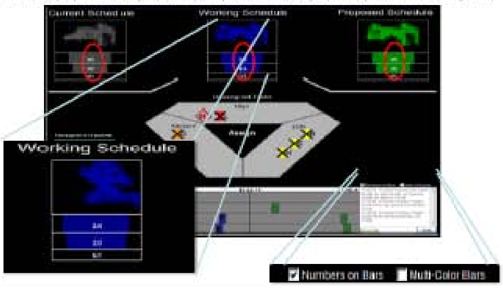
Tasks Completed

Comparing these three schedules: both "Working" and "Proposed" perform more high priority tasks than "Current." "Working" has just been altered (as shown by the ghosting change in the high priority ladder rung).

"Working" starts out identical to "Proposed" and changes as you work on it.

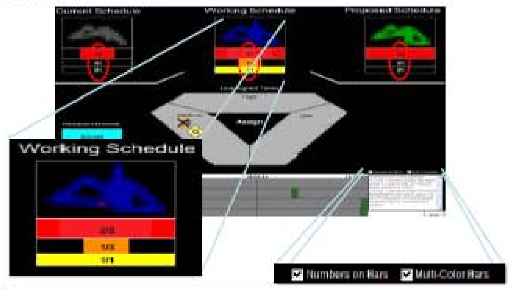
Configural Display "Numbers on Bars" Feature

The ratio of tasks assigned to total tasks will show up on the bars of the hierarchical ladder when the optional "Numbers on Bars" checkbox is selected. For example, on the high priority bar, 3 of the 4 high priority tasks are assigned.

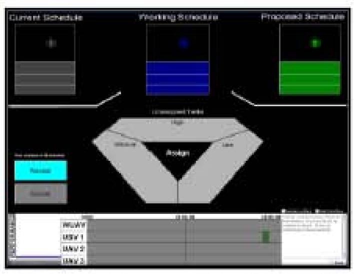


Configural Display "Multi-Color Bars" Feature

The priority level of the bars on the hierarchical ladder will show up for quick priority recognition when the "Multi-Color Bars" checkbox is selected. It is your choice to use this feature.



Nothing to Assign

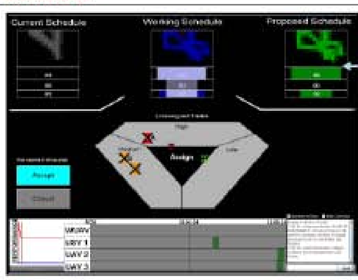


Sometimes the Auto-planner is able to assign all tasks. The SCT will be empty. Click **Accept** to proceed with the completely assigned schedule generated by the auto-planner. If you do not choose Accept, the UxVs will not receive the communication of tasks they need to perform. But, you still have the option to Cancel.

Selected Task shown on configural display

The red X appearing in the "map" of each Configural Display corresponds to the location of the currently **selected task**.

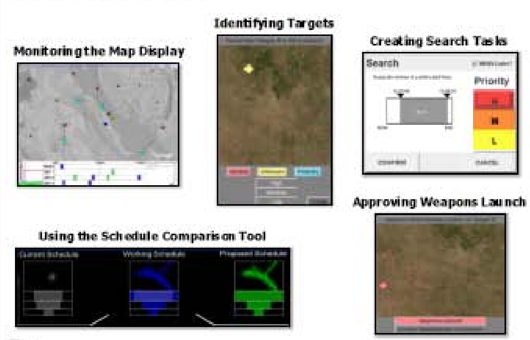
Switching Schedules

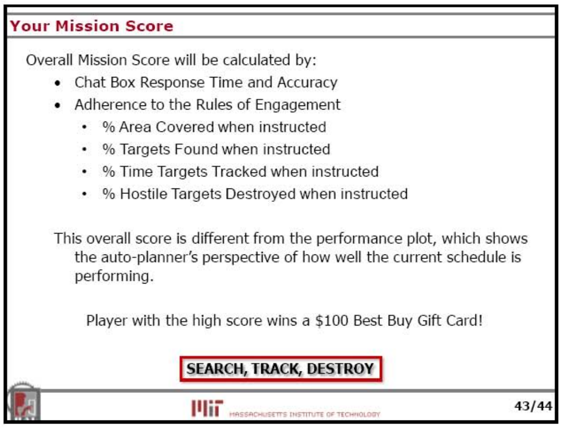


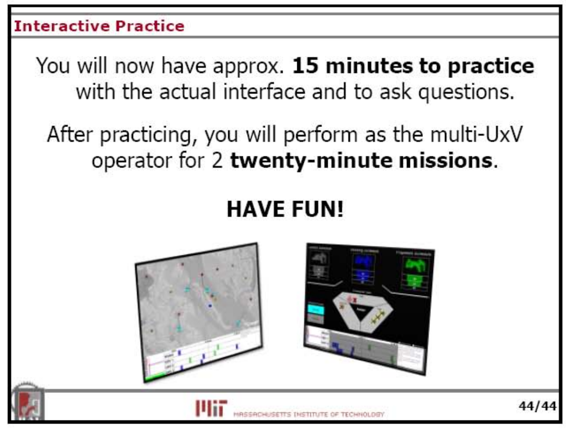
Click to Switch

Tasks viewed in the unassigned central area always correspond to the "Working" Schedule. To view unassigned tasks for either the "Current" or "Proposed" Schedules, clicking on each schedule's configural display will set the "Working" Schedule as equivalent to the selected schedule (erasing the old "Working").

Main parts of your role

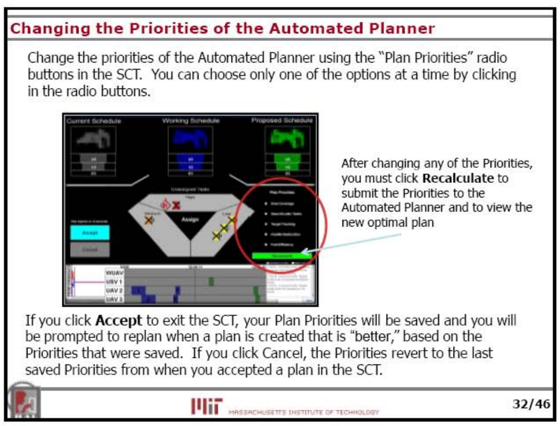


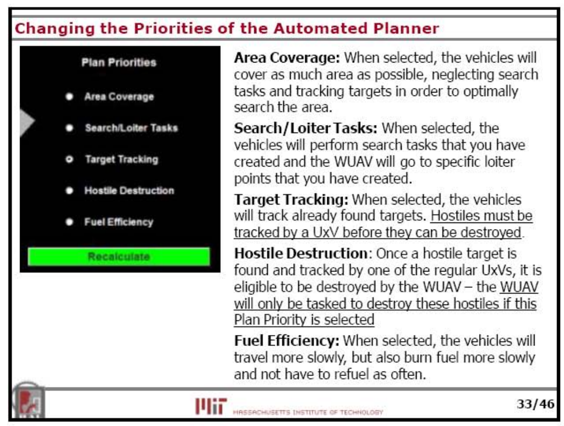




G.2 Radio Button Objective Function Tutorial

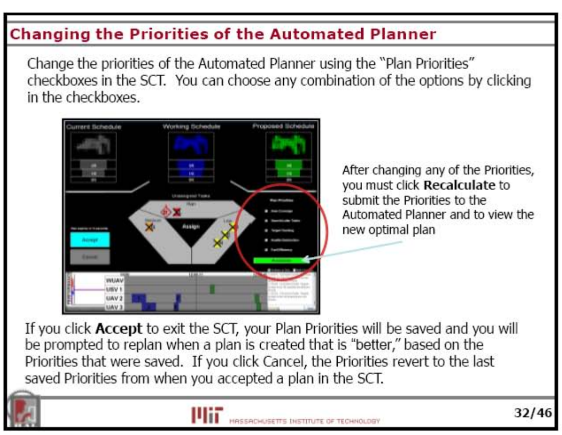
Extra slides specific to the Radio Objective Function:

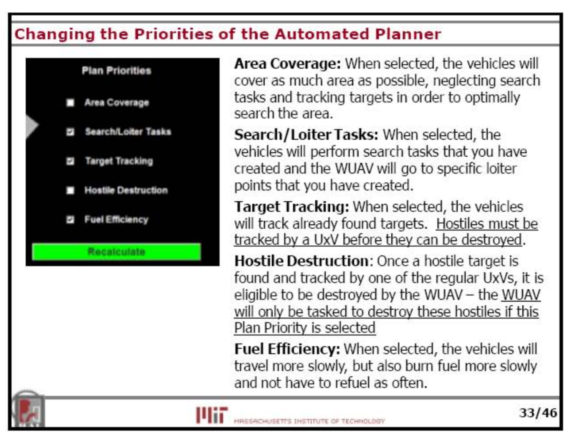




G.3 Checkbox Button Objective Function Tutorial

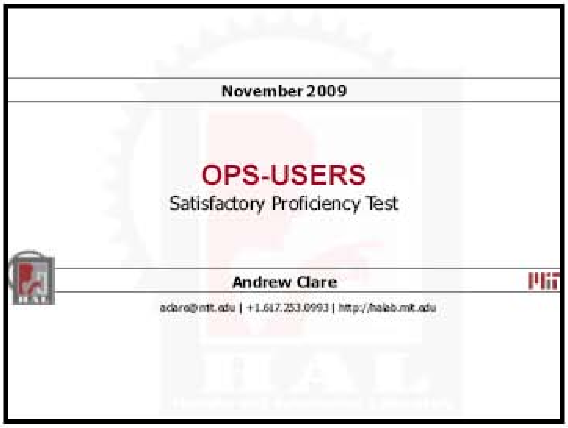
Extra slides specific to the Checkbox Objective Function:

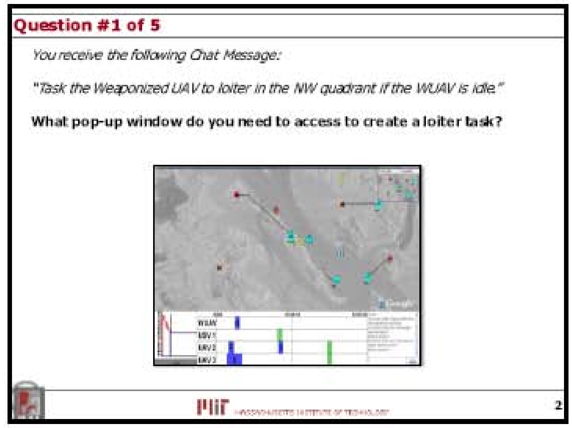


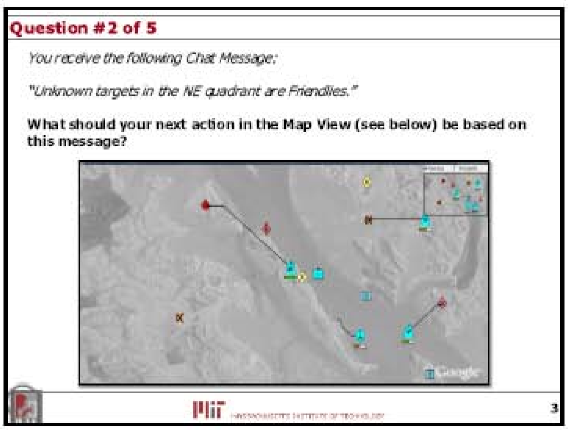


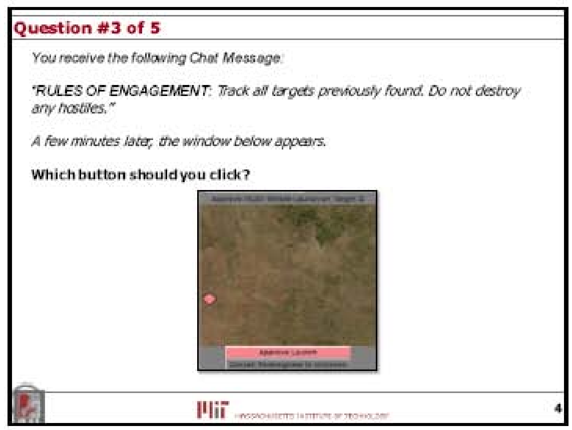
Appendix H: Proficiency Tests

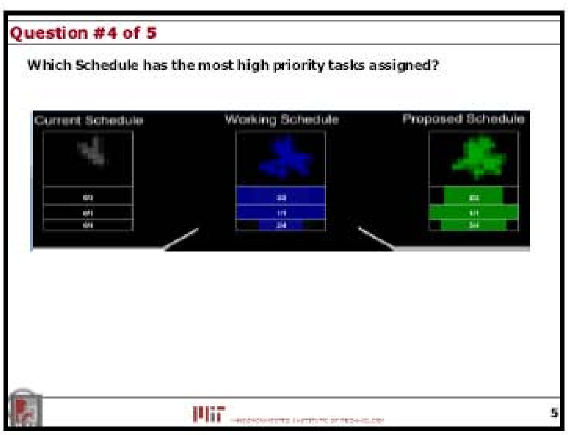
H.1 Static (None) Objective Function Test

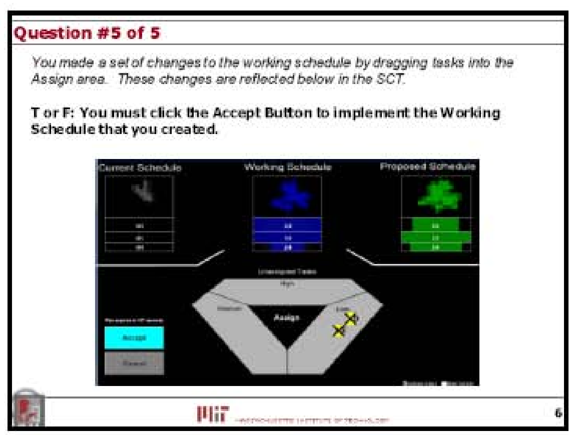




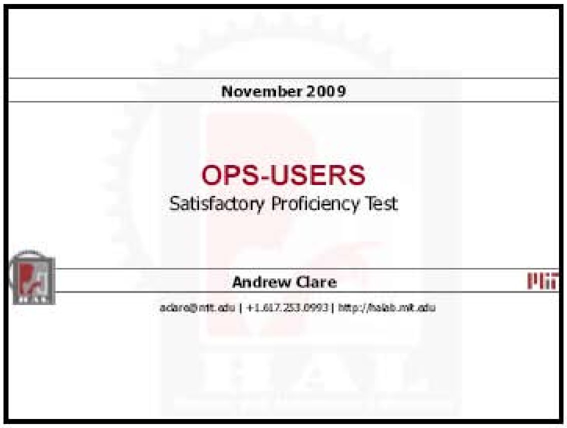


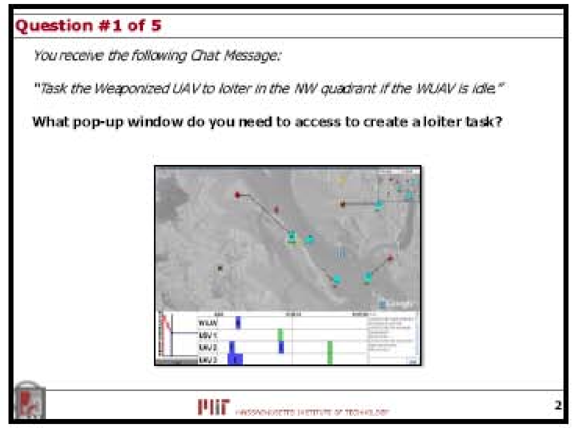


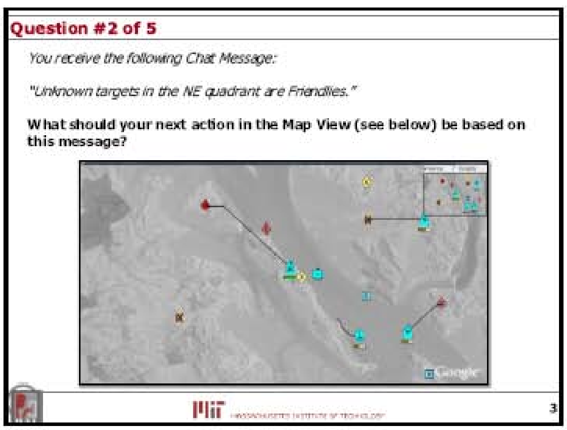


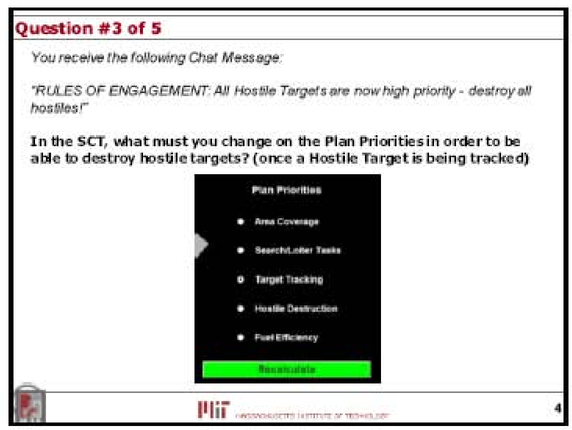


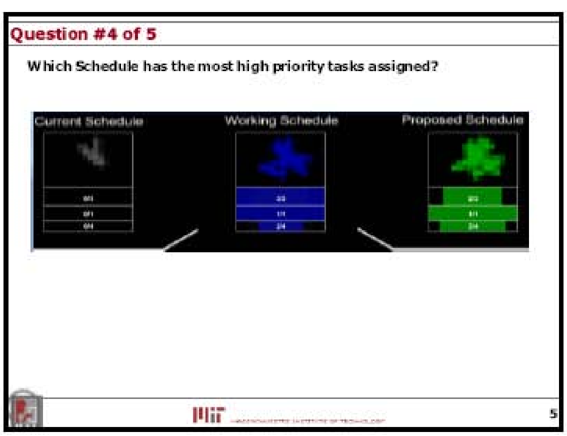
H.2 Radio Button Objective Function Test

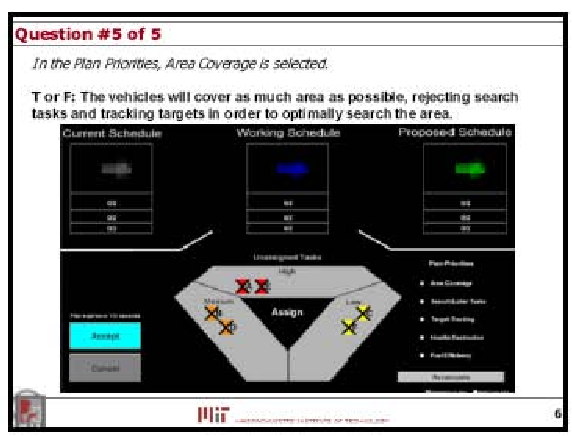




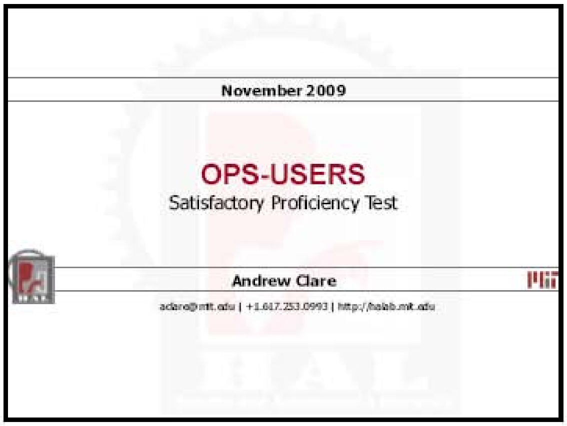


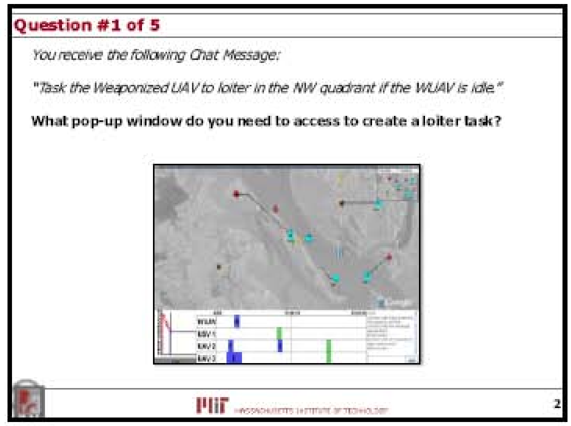


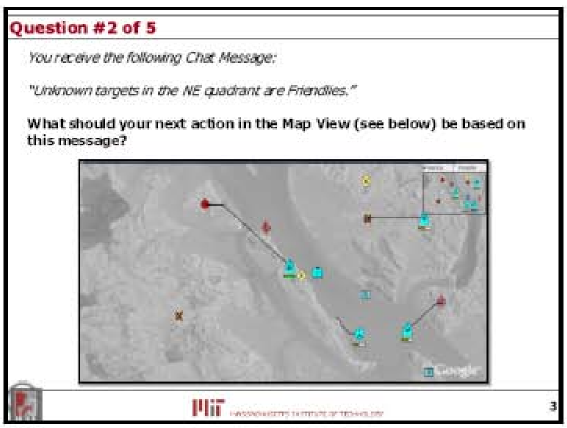


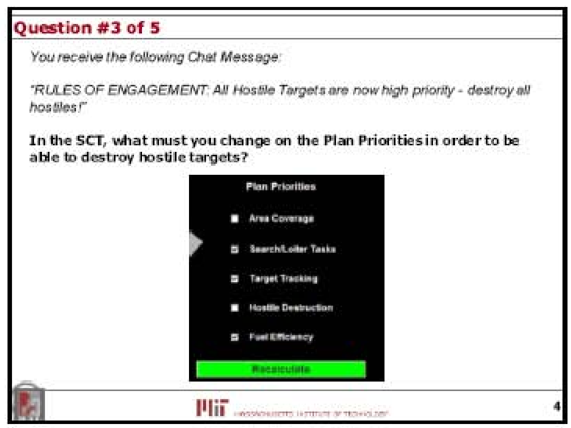


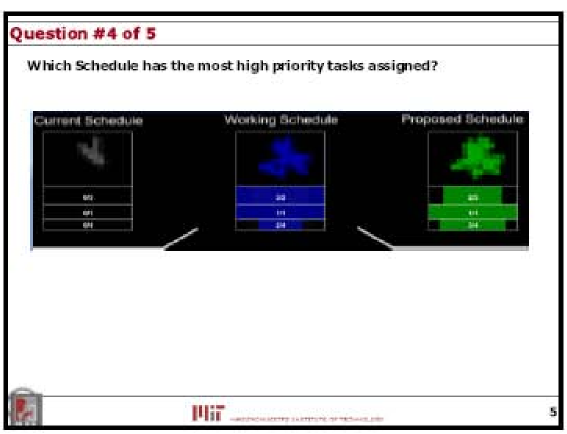
H.3 Checkbox Button Objective Function Test

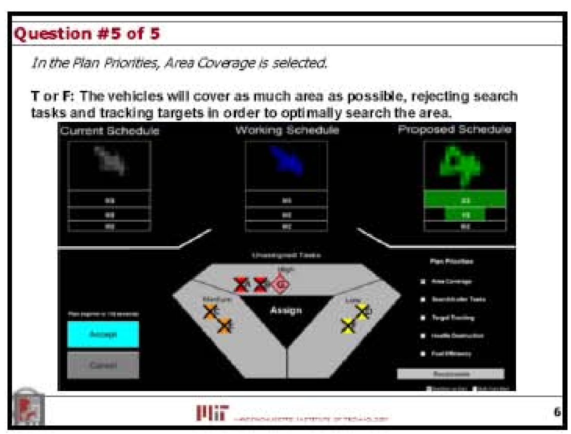












H.4 Answer Key

PASSING = 4 out of 5 correct!
<u>OPS-USERS QUIZ ANSWER KEY – Quiz #1</u>

Static
1. Search Task Window
2. Right click on Unknown Target G, re-designate to friendly
3. “Cancel: Re-designate to Unknown” Button
4. Working Schedule
5. True

Checkbox
1. Search Task Window
2. Right click on Unknown Target G, re-designate to friendly
3. Check “Hostile Destruction”
4. Working Schedule
5. True

Radio
1. Search Task Window
2. Right click on Unknown Target G, re-designate to friendly
3. Click “Hostile Destruction”
4. Working Schedule
5. True

Appendix I: Questionnaires

Scenario Feedback Survey

Round 1

1. Subject number:_____

2. How confident were you about the plans that you created?

   Not Confident   *Somewhat Confident*   *Confident*   *Very Confident*   *Extremely Confident*

   Comments:

3. How did you feel you performed overall?

   Very Poor   *Poor*   *Satisfactory*   *Good*   *Excellent*

4. How busy did you feel during the mission?

   Idle   *Not Busy*   *Busy*   *Very Busy*   *Extremely Busy*

5. How satisfied were you with the plans created by the Automated Planner?

   Very Unsatisfied   *Unsatisfied*   *Satisfied*   *Very satisfied*   *Extremely satisfied*

Scenario Feedback Survey

Round 2

1. Subject number:_____

2. How confident were you about the plans that you created?

   Not Confident   *Somewhat Confident*   *Confident*   *Very Confident*   *Extremely Confident*

   Comments:

3. How did you feel you performed overall?

   Very Poor   *Poor*   *Satisfactory*   *Good*   *Excellent*

4. How busy did you feel during the mission?

   Idle   *Not Busy*   *Busy*   *Very Busy*   *Extremely Busy*

5. How satisfied were you with the plans created by the Automated Planner?

   Very Unsatisfied   *Unsatisfied*   *Satisfied*   *Very satisfied*   *Extremely satisfied*

Questions about the Experiment Overall

1. Were there aspects of the interface that you particularly liked or disliked?

2. Did you understand the changes in the Rules of Engagement? Did you feel like you could implement those changes via the interface?

3. Did you feel that the Automated Planner was fast enough?

4. Did you feel that you understood how manipulating the Plan Priorities affected the plans created by Automated Planner?

5. Other comments:

References

[1] J. Westwood, "Global Prospects for AUVs," presented at Offshore Technology Conference, Houston, TX, 2001.

[2] J. S. Norris, M. W. Powell, M. A. Vona, P. G. Backes, and J. V. Wick, "Mars Exploration Rover Operations with the Science Activity Planner," presented at IEEE International Conference on Robotics and Automation, Barcelona, Spain, 2005.

[3] S. Odedra, S. D. Prior, and M. Karamanoglu, "Investigating the Mobility of Unmanned Ground Vehicles," presented at International Conference on Manufacturing and Engineering Systems, Huwei, Yunlin, Taiwan, 2009.

[4] B. C. Sweeten, D. Royer, and S. Keshmiri, "The Meridian UAV Flight Performance Analysis Using Analytical and Experimental Data," presented at AIAA Infotech@Aerospace Conference and Exhibit, Seattle, Washington, 2009.

[5] A. R. Girard and J. K. Hedrick, "Border Patrol and Surveillance Missions Using Multiple Unmanned Air Vehicles," presented at 43rd IEEE Conference on Decision and Control, Paradise Islands, The Bahamas, 2004.

[6] D. W. Casbeer, D. B. Kingston, R. W. Beard, T. W. McLain, S. M. Li, and R. Mehra, "Cooperative Forest Fire Surveillance Using a Team of Small Unmanned Air Vehicles," *International Journal of Systems Science*, vol. 37, no. 6, pp. 351-360, 2006.

[7] M. L. Cummings, S. Bruni, S. Mercier, and P. J. Mitchell, "Automation Architecture for Single Operator, Multiple UAV Command and Control," *The International Command and Control Journal*, vol. 1, no. 2, pp. 1-24, 2007.

[8] M. Alighanbari and J. P. How, "Robust Decentralized Task Assignment for Cooperative UAVs," presented at AIAA Guidance, Navigation, and Control Conference and Exhibit, Keystone, Colorado, 2006.

[9] L. F. Bertuccelli, H.-L. Choi, P. Cho, and J. P. How, "Real-Time Multi-UAV Task Assignment in Dynamic and Uncertain Environments," presented at AIAA Guidance, Navigation, and Control Conference, Chicago, IL, 2009.

[10] Office of the Secretary of Defense, "Unmanned Aircraft Systems (UAS) Roadmap, 2005-2030," DoD, Washington D.C., 2005.

[11] Naval Studies Board, "Autonomous Vehicles in Support of Naval Operations," National Research Council, Washington D.C., 2005.

[12] U.S. Navy, "Dragon Warrior Communications Relay." [Online]. Available: http://cs.itd.nrl.navy.mil/work/dragon_warrior/index.php. [Accessed: May 2010].

[13] A. S. Clare, C. S. Hart, and M. L. Cummings, "Assessing Operator Workload and Performance in Expeditionary Multiple Unmanned Vehicle Control," presented at 48th AIAA Aerospace Sciences Meeting, Orlando, FL, 2010.

[14] M. L. Cummings and P. J. Mitchell, "Predicting Controller Capacity in Supervisory Control of Multiple UAVs," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 38, no. 2, pp. 451-460, 2008.

[15] H. A. Ruff, S. Narayanan, and M. H. Draper, "Human Interaction with Levels of Automation and Decision-Aid Fidelity in the Supervisory Control of Multiple Simulated Unmanned Air Vehicles," *Presence*, vol. 11, no. 4, pp. 335-351, 2002.

[16] M. L. Cummings and S. Guerlain, "Developing Operator Capacity Estimates for Supervisory Control of Autonomous Vehicles," *Human Factors*, vol. 49, no. 1, pp. 1-15, 2007.

[17] J. Van Den Berg, D. Ferguson, and J. Kuffner, "Anytime Path Planning and Replanning in Dynamic Environments," presented at IEEE International Conference on Robotics and Automation (ICRA), Orlando, FL, 2006.
[18] J. P. How, C. Fraser, K. C. Kulling, L. F. Bertuccelli, O. Toupet, L. Brunet, A. Bachrach, and N. Roy, "Increasing Autonomy of UAVs," *IEEE Robotics & Automation Magazine*, vol. 16, no. 2, 2009.
[19] J. J. Marquez, "Human-Automation Collaboration: Decision Support for Lunar and Planetary Exploration," Ph.D thesis, Massachusetts Institute of Technology, Cambridge, MA, 2007.
[20] B. G. Silverman, "Human-Computer Collaboration," *Human-Computer Interaction*, vol. 7, no. 2, pp. 165-196, 1992.
[21] P. Smith, E. McCoy, and C. Layton, "Brittleness in the Design of Cooperative Problem-solving Systems: The Effects on User Performance," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 27, no. 3, pp. 360-371, 1997.
[22] S. A. Guerlain, "Using the Critiquing Approach to Cope With Brittle Expert Systems," presented at Human Factors and Ergonomics Society 39th Annual Meeting, San Diego, CA, 1995.
[23] P. Polson and N. Smith, "The Cockpit Cognitive Walkthrough," presented at 10th Symposium on Aviation Psychology, Columbus, OH, 1999.
[24] M. L. Cummings and S. Bruni, "Collaborative Human-Automation Decision Making," in *Handbook of Automation LXXVI*, S. Y. Nof, Ed. New York: Springer, 2009, pp. 437-448.
[25] L. M. Forest, A. Kahn, J. Thomer, and M. Shapiro, "The Design and Evaluation of Human-Guided Algorithms for Mission Planning," presented at HSIS'07 - Human Systems Integration Symposium, Annapolis, MD, 2007.
[26] S. Fisher, "Replan Understanding for Heterogenous Unmanned Vehicle Teams," Masters of Engineering thesis, Massachusetts Institute of Technology, Cambridge, MA, 2008.
[27] M. L. Cummings, A. S. Clare, and C. S. Hart, "The Role of Human-Automation Consensus in Multiple Unmanned Vehicle Scheduling," *Human Factors*, vol. 52, no. 1, 2010.
[28] D. Anderson, E. Anderson, N. Lesh, J. Marks, B. Mirtich, D. Ratajczak, and K. Ryall, "Human-Guided Simple Search," presented at AAAI 2000, Austin, TX, 2000.
[29] S. D. Scott, N. Lesh, and G. W. Klau, "Investigating Human-Computer Optimization," presented at Conference on Human Factors in Computer Systems (CHI), Minneapolis, MN, 2002.
[30] E. Bitton and K. Goldberg, "Hydra: A Framework and Algorithms for Mixed-Initiative UAV-Assisted Search and Rescue," presented at IEEE International Conference on Automation Science, Washington D.C., 2008.
[31] D. Carlsen, M. Malone, J. Kollat, and T. W. Simpson, "Evaluating the Performance of Visual Steering Commands for User-Guided Pareto Frontier Sampling During Trade Space Exploration," presented at ASME 2008 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference, Brooklyn, NY, 2008.
[32] M. L. Cummings, "Automation Bias in Intelligent Time Critical Decision Support Systems," presented at AIAA 3rd Intelligent Systems Conference, Chicago, IL, 2004.
[33] M. H. Burstein and D. V. McDermott, "Issues in the Development of Human-Computer Mixed-Initiative Planning Systems," in *Cognitive Technology: In Search of a Humane*

Interface, B. Gorayska and J. Mey, Eds. Amsterdam: Elsever Science B.V., 1996, pp. 285-303.
[34] L. A. Kramer and S. F. Smith, "Optimizing for Change: Mixed-Initiative Resource Allocation with the AMC Barrel Allocator," presented at 3rd International NASA Workshop on Planning and Scheduling for Space, Houston, TX, 2002.
[35] A. E. Howe, L. D. Whitley, L. Barbulescu, and J. P. Watson, "Mixed Initiative Scheduling for the Air Force Satellite Control Network," presented at 2nd International NASA Workshop on Planning and Scheduling for Space, San Francisco, CA, 2000.
[36] M. L. Hanson, E. Roth, C. M. Hopkins, and V. Mancuso, "Developing Mixed-Initiative Interaction with Intelligent Systems: Lessons Learned from Supervising Multiple UAVs," presented at AIAA 1st Intelligent Systems Technical Conference, Chicago, IL, 2004.
[37] C. Miller, H. Funk, P. Wu, R. Goldman, J. Meisner, and M. Chapman, "The Playbook Approach to Adaptive Automation," presented at Human Factors and Ergonomics Society 49th Annual Meeting, Orlando, FL, 2005.
[38] S. Bruni and M. Cummings, "Human Interaction with Mission Planning Search Algorithms," presented at Human Systems Integration Symposium, Arlington, VA, 2005.
[39] M. L. Cummings and S. Bruni, "Human-Automation Collaboration in Complex Multivariate Resource Allocation Decision Support Systems," *Intelligent Decision Technologies*, vol. 4, no. 2, pp. 101-114, 2010.
[40] C. D. Wickens and J. G. Hollands, *Engineering Psychology and Human Performance*, 3rd ed. Upper Saddle River, N.J.: Prentice Hall, 2000.
[41] J. L. Thorner, "Trust-Based Design of Human-Guided Algorithms," Master of Science thesis, Massachusetts Institute of Technology, Cambridge, MA, 2007.
[42] J. Rieskamp and U. Hoffrage, "Inferences Under Time Pressure: How Opportunity Costs Affect Strategy Selection," *Acta Psychologica*, vol. 127, no. 2, pp. 258-276, 2008.
[43] Y. Boussemart, "Effects of Time Pressure on the Use of an Automated Decision Support System," Massachusetts Institute of Technology, Humans and Automation Lab, 2007.
[44] J. Rasmussen, "Skills, rules, and knowledge: Signals, signs, and symbols, and other distinctions in human performance models," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 13, no. 3, pp. 257-266, 1983.
[45] J. Rasmussen, "Outlines of a Hybrid Model of the Process Plant Operator," in *Monitoring Behavior and Supervisory Control*, T. B. Sheridan and G. Johannsen, Eds. New York, NY: Plenum Press, 1976.
[46] A. Stentz, "Optimal and Efficient Path Planning for Partially-known Environments," presented at IEEE International Conference on Robotics and Automation, San Diego, CA, 1994.
[47] P. M. Fitts, "Human Engineering for an Effective Air Navigation and Traffic Control System." Washington, D.C.: National Research Council, 1951.
[48] H. E. Price, "The Allocation of Functions in Systems," *Human Factors*, vol. 27, no. 1, pp. 33-45, 1985.
[49] T. B. Sheridan and W. Verplank, "Human and Computer Control of Undersea Teleoperators," Man-Machine Systems Laboratory, Department of Mechanical Engineering, MIT, Cambridge, MA, 1978.
[50] R. Parasuraman, T. B. Sheridan, and C. D. Wickens, "A Model for Types and Levels of Human Interaction with Automation," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 30, no. 3, pp. 286-297, 2000.

[51] T. B. Sheridan, "Function Allocation: Algorithm, Alchemy, or Apostasy?," *International Journal of Human-Computer Studies*, vol. 52, pp. 203-216, 2000.

[52] J.-M. Hoc, "From Human-machine Interaction to Human-machine Cooperation," *Ergonomics*, vol. 43, no. 7, pp. 833-843, 2000.

[53] M. T. Older, P. E. Waterson, and C. W. Clegg, "A Critical Assessment of Task Allocation Methods and Their Applicability," *Ergonomics*, vol. 40, no. 2, pp. 151-171, 1997.

[54] S. W. Dekker and P. C. Wright, "Function Allocation: A Question of Task Transformation not Allocation," in *ALLFN'97 Revisiting the Allocation of Functions Issue: New Perspectives*, E. Fallon, L. Bannon, and J. McCarthy, Eds. Louisville, KY: IEA Press, 1997, pp. 31-40.

[55] S. W. Dekker and D. D. Woods, "MABA-MABA or Abracadabra? Progress on Human-Automation Co-ordination," *Journal of Cognition, Technology and Work*, vol. 4, pp. 240-244, 2002.

[56] M. A. Goodrich and A. C. Schultz, "Human-Robot Interaction: A Survey," *Foundations and Trends in Human-Computer Interaction*, vol. 1, no. 3, pp. 203-275, 2007.

[57] R. Parasuraman, T. Bahri, J. E. Deaton, J. G. Morrison, and M. Barnes, "Theory and Design of Adaptive Automation in Adaptive Systems," Naval Air Warfare Center, Aircraft Division, Warminster, PA, Progress Report No. NAWCADWAR-92033-60, 1992.

[58] V. Riley, "A General Model of Mixed-Initiative Human-Machine Systems," presented at Human Factors and Ergonomics Society 33rd Annual Meeting, Denver, CO, 1989.

[59] D. B. Kaber, M. R. Endsley, and E. Onal, "Design of Automation for Telerobots and the Effect on Performance, Operator Situation Awareness and Subjective Workload," *Human Factors and Ergonomics in Manufacturing*, vol. 10, no. 4, pp. 409-430, 2000.

[60] A. E. Kirkpatrick, B. Dilkina, and W. S. Havens, "A Framework for Designing and Evaluating Mixed-Initiative Optimization Systems," presented at International Conference on Automated Planning and Scheduling, Monterey, CA, 2005.

[61] M. L. Cummings and S. Bruni, "Collaborative Human-Computer Decision Making in Network Centric Warfare," presented at TTCP HUM TP7 Workshop on Aerospace Human Factors Issues in Network-Centric Warfare, Salisbury, UK, 2005.

[62] S. Bruni, J. J. Marquez, A. Brzezinski, C. Nehme, and Y. Boussemart, "Introducing a Human-Automation Collaboration Taxonomy (HACT) in Command and Control Decision-Support Systems," presented at International Command and Control Research and Technology Symposium, Newport, RI, 2007.

[63] R. Parasuraman, "Designing Automation for Human Use: Empirical Studies and Quantitative Models," *Ergonomics*, vol. 43, no. 7, pp. 931-951, 2000.

[64] U. S. D.O.D., "Interface Standard, Common Warfighting Symbology," MIL-STD-2525B, 30 January, 1999.

[65] J. J. Gibson, *The Ecological Approach to Visual Perception*. Boston, MA: Houghton Mifflin, 1979.

[66] C. Layton, P. J. Smith, and C. E. McCoy, "Design of a Cooperative Problem-solving System for En-route Flight Planning - An Empirical Evaluation," *Human Factors*, vol. 36, no. 1, pp. 94-116, 1994.

[67] M. A. Goodrich, J. Johansen, T. W. McLain, J. Anderson, J. Sun, and J. W. Crandall, "Managing Autonomy in Robot Teams: Observations from Four Experiments," presented

at 2nd ACM/IEEE International Conference on Human Robot Interaction, Washington D.C., 2007.
[68] S. Ganapathy, "Human-Centered Time-Pressured Decision Making In Dynamic Complex Systems," Ph.D thesis, Wright State University, Dayton, OH, 2006.
[69] K. Johnson, L. Ren, J. K. Kuchar, and C. M. Oman, "Interaction of Automation and Time Pressure in a Route Replanning Task," presented at International Conference on Human-Computer Interaction in Aeronautics (HCI-Aero), Cambridge, MA, 2002.
[70] J. A. Maule and A. C. Edland, "The Effect of Time Pressure on Human Judgment and Decision Making," in *Cognitive Models and Explanations*, W. R. Crozier, R. Ranyard, and O. Svenson, Eds. New York: Routledge, 1997.
[71] A. Tversky and D. Kahneman, "Judgment Under Uncertainty: Heuristics and Biases," *Science*, vol. 185, no. 4157, pp. 1124-1131, 1974.
[72] M. L. Cummings and S. Guerlain, "Using a Chat Interface as an Embedded Secondary Tasking Tool," in *Human Performance, Situation Awareness and Automation: Current Research and Trends. HPSAA II*, vol. 1, D. A. Vincenzi, M. Mouloua, and P. A. Hancock, Eds. London: Routledge, 2004, pp. 240-248.
[73] R. Parasuraman and V. Riley, "Humans and Automation: Use, Misuse, Disuse, Abuse," *Human Factors*, vol. 39, no. 2, pp. 230-253, 1997.
[74] M. Kutner, C. Nachtsheim, J. Neter, and W. Li, *Applied Linear Statistical Models*, 5th ed. Boston: McGraw-Hill Irwin, 2004.
[75] M. R. Endsley, "Toward a Theory of Situation Awareness in Dynamic Systems," *Human Factors*, vol. 37, no. 1, pp. 32-64, 1995.
[76] T. Lamoureux, "The Influence of Aircraft Proximity Data on the Subjective Mental Workload of Controllers in the Air Traffic Control Task," *Ergonomics*, vol. 42, no. 11, pp. 1482-1491, 1999.
[77] J. D. Lee and K. A. See, "Trust in Automation: Designing for Appropriate Reliance," *Human Factors*, vol. 46, no. 1, pp. 50-80, 2003.
[78] D. N. Southern, "Human-Guided Management of Collaborating Unmanned Vehicles in Degraded Communication Environments," Masters of Engineering thesis, Massachusetts Institute of Technology, Cambridge, MA, 2010.
[79] U.S. Air Force, "Factsheets: MQ-1 Predator." [Online]. Available: http://www.af.mil/information/factsheets/factsheet.asp?fsID=122. [Accessed: May, 2010].
[80] K. Munson, "Jane's Unmanned Aerial Vehicles and Targets," vol. 20. Coulson, UK: Jane's Information Group, 2003.
[81] J. D. Anderson, *Fundamentals of Aerodynamics*, 3rd ed. New York, NY: McGraw-Hill Companies, Inc., 2001.