JOURNAL OF AEROSPACE COMPUTING, INFORMATION, AND COMMUNICATION
Vol. 9, No. 4, December 2012

Operator Objective Function Guidance for a Real-Time Unmanned Vehicle Scheduling Algorithm

Andrew S. Clare,* Mary L. Cummings,† and Jonathan P. How‡
Massachusetts Institute of Technology, Cambridge, Massachusetts 02139
Andrew K. Whitten§
U.S. Air Force Flight Test Center, Edwards Air Force Base, California, 93524
and
Olivier Toupet¶
Zee.Aero, Mountain View, California 94043

DOI: 10.2514/1.I010019

Advances in autonomy have made it possible to invert the typical operator-to-unmanned-vehicle ratio so that a single operator can now control multiple heterogeneous unmanned vehicles. Algorithms used in unmanned-vehicle path planning and task allocation typically have an objective function that only takes into account variables initially identified by designers with set weightings. This can make the algorithm seemingly opaque to an operator and brittle under changing mission priorities. To address these issues, it is proposed that allowing operators to dynamically modify objective function weightings of an automated planner during a mission can have performance benefits. A multiple-unmanned-vehicle simulation test bed was modified so that operators could either choose one variable or choose any combination of equally weighted variables for the automated planner to use in evaluating mission plans. Results from a human-participant experiment showed that operators rated their performance and confidence highest when using the dynamic objective function with multiple objectives. Allowing operators to adjust multiple objectives resulted in enhanced situational awareness, increased spare mental capacity, fewer interventions to modify the objective function, and no significant differences in mission performance. Adding this form of flexibility and transparency to automation in future unmanned vehicle systems could improve performance, engender operator trust, and reduce errors.

I. Introduction

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 (e.g., [1]), have enabled the military to conduct long-duration missions over hostile territory without placing a pilot in harm's way, and have aided in weather research [2], border patrol [3], and forest firefighting [4]. Although these UVs contain advanced technology, they typically require multiple human operators, often more than a comparable manned vehicle would require [5]. The need for many operators per UV causes increased training and operating costs [5] and challenges in meeting the ever-increasing demand for more UV operations [6]. 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 U.S. 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].

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. Although multiple studies have demonstrated the capacity of a single operator to control multiple UVs [12,13], the large amount of data generated by such a system could cause operator cognitive saturation, which has been shown to correlate with poor operator performance [14,15]. To mitigate possible high mental workload in these future systems, operators will be assisted by automated planners, which can be faster and more accurate than humans at path planning [16] and task allocation [17] in a multivariate, dynamic, time-pressured environment.

Outside 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. Although 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 use these tools, especially when working in dynamic, time-critical situations with high uncertainty in information [18].

Human management of these automated planners is crucial, as automated planners do not always perform well 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 [19,20]. In a command and control situation such as supervising multiple UVs—where events are often unanticipated—automated planners have difficulty accounting for and responding to unforeseen problems [21,22]. Additionally, operators can

Received 7 November 2011; revision received 18 July 2012; accepted for publication 26 July 2012.

*Ph.D. Candidate, Department of Aeronautics and Astronautics, Humans and Automation Laboratory; aclare@mit.edu. Student Member AIAA.
†Associate Professor of Aeronautics and Astronautics, Director of the Humans and Automation Laboratory; missyc@mit.edu. Associate Fellow AIAA.
‡Professor of Aeronautics and Astronautics, Director of the Aerospace Controls Laboratory; jhow@mit.edu. Associate Fellow AIAA.
§Flying Qualities Engineer, 773 Test Squadron; andrew.whitten@edwards.af.mil. Member AIAA.
¶Senior Controls Engineer; otoupet@zee.aero. Member AIAA.

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 [18,20,23,24]. To truly assist human supervisors of multiple UVs in dynamic environments, however, automated planners must be capable of dynamic mission replanning. As vehicles move, new tasks emerge and mission needs shift, and the automated planner should adapt to assist in real-time decision making. This will require greater flexibility and transparency in the computer algorithms designed for supporting multi-UV missions.

To address this lack of flexibility and transparency in multi-UV control algorithms, this paper investigates the impact of human–computer collaboration in the context of dynamic objective function manipulation. 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 predetermined variables with set weightings. Predetermined variables are those quantifiable metrics chosen in advance by the designers of the algorithm as crucial to the goals of the mission. In this effort, human operators are given the ability to modify the weightings of optimization variables during a mission. Given the potential for high operator workload and possible negative performance consequences, we investigate operator workload as well as human and system performance as a function of providing this additional level of human–computer collaboration.

II. Background

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 [25–27] or trade space exploration for large-scale design optimization [28]. 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 [29]. This automation bias can cause complacency, degradation in skills and performance, and potential loss of situational awareness (SA) [30].

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 contribute to the formulation and analysis of solutions [31]. For example, a mixed-initiative tool to solve an overconstrained 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 [32]. Scott et al. showed that, in experiments with humans using mixed-initiative systems for vehicle routing, operator intervention led to better results, but there was variation in the way that operators interacted with the system and in their success in working with the automation [26]. Howe et al. developed a mixed initiative scheduler for the U.S. Air Force satellite control network, implementing a satisficing algorithm, which recommended plans despite the fact that a solution that satisfied all constraints did not exist [33]. The user chose the 'best' plan despite constraint violations and modified 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 nonlinear evaluations made by the human and the unavailability of all information necessary for the algorithm to make a decision [33].

A number of other studies have developed algorithms and architectures for control and coordination of multiple semi-autonomous satellite swarms [34–38]. Unmanned spacecraft often have a similar or even higher level of automation as compared to modern UVs, and there are some similarities between the two domains, including the potential desire to use these vehicles for surveillance missions, the necessity of an algorithm to coordinate the movement of the vehicles, and the human–computer interaction necessary for mission success. Although some studies have considered the role of the human controller in supervising multiple spacecraft or other space autonomous agents [39–44], few have live experiments with human operators to investigate the most appropriate way for operators to interact with or modify these control algorithms in real-time operations, thus warranting further research in this important area.

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 [45]. They also observed that operators tended to think less in terms of numerical optimization when planning UAV routes but more in abstract terms about the overall goals or tactical objectives that they wanted to accomplish. The authors argued 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 use [46]. This system limited the human operators' interactions with the automation to selecting predetermined 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 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 to aid in accomplishing mission objectives. In these previous studies, the assumption was that the planning algorithms were static and unchanging throughout the period in which the human interacted with the automation. Operator SA was typically low, and operators complained about the lack of transparency in how the automation generated plans [12,18,33,45]. 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 focused on providing the human operator with the ability to modify the way the automated planner works for collaborative decision-making. In one study, a customizable heuristic search algorithm, where the human operator could choose and rank criteria that adjusted the weights of variables in the objective function, was used to aid operators in a multivariate resource allocation task [47,48]. The associated decision-support interface allowed the human operator to manually adjust the solution after using the heuristic search algorithm to develop an initial solution. Results showed no statistical difference in performance between this method of collaborative human–automation planning as compared to a more manual method of planning. However, this collaborative interface using the customizable search algorithm required significantly fewer steps than the manual interface, thus reducing overall workload. Although lower workload was achieved, the mission was not time-critical on the order of seconds.

Finally, Forest et al. conducted an experiment during which operators created a schedule for multiple UAVs with the assistance of a human-guided algorithm [24]. The subjects were presented with different interfaces to preplan a mission based on preexisting targets with given values and risks. In some instances, subjects could modify the weights on five factors that the objective function used to calculate scores for the plans including total target value, risk, percentage of available missiles used (utilization), distance, and mission time. Although subjects could use any of these factors to evaluate plans, the mission instructions encouraged them 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 [24,49]. 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, although automation is excellent at optimizing a solution for specific goals, it 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 highlights the difficulty of human–automation collaboration when humans have different internal objective functions from that of the automation. In subjective ratings, participants gave the highest rating to the interface where they had the most control of the objective function [49]. They found it more 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 preplanning scenario, where algorithm searches took 20–30 s, and the entire planning process could take up to 15 min. Although these experiments show that dynamic objective functions can result in improved collaboration between humans and automation, only six participants were involved in the study.

This previous literature on human–algorithm interaction reveals three key areas that warrant further research. First, most of the previous experiments in human–automation collaboration occurred in fairly static environments with high certainty. Typically, the experiments involved mission preplanning, 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 where operators must make decisions in real time, so collaborative control methods need to be developed that can operate in replanning environments, as opposed to planning environments that occur before a mission.

A related second area is experimental investigation of human–automation collaboration under time pressure. Many of the collaborative systems discussed previously were developed for preplanning scenarios, when operators have minutes, hours, or even days to make decisions. In real-time replanning scenarios, the time scale for decision making will be reduced dramatically, likely 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 [50,51]. Although 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.

Finally, there is a need to increase our understanding of just how operators could and should express their desires to an automated planner to ensure alignment of the objective functions of the human and automation. A number of participants in previous experiments complained of a mismatch between their own goals and the plans generated by the automated planner. Just how to implement a system that allows such objective function expression between a human operator and an automated planner has not been investigated in detail.

Although there are numerous potential methods to address these areas, this paper seeks to further this body of knowledge 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. To evaluate this kind of collaborative control, dynamic objective functions were implemented in an existing multiple UV simulation test bed described in the following section.

III. Simulation Test Bed

This paper uses a collaborative, multiple-UV simulation environment called onboard planning system for UVs 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 [17]; all the decision-support displays described here have operated actual small air and ground UVs in real time.

Operators were placed in a simulated command center where they controlled multiple heterogeneous UVs for the purpose of searching the area of interest for new targets, tracking targets, and approving weapons launch. The UVs in the scenario included 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). The UAVs and USV were responsible for searching for targets. Once a target was found, the operator was alerted to perform a target identification task (i.e., hostile, unknown, or friendly) along with assigning an associated priority level (i.e., high, medium, low). Then, hostile targets were tracked by one or more of the vehicles until the human operator approved WUAV missile launches. A primary assumption was that operators had minimal time to interact with the displays due to other mission-related tasks.

Operators had two exclusive tasks that could not be performed by automation: target identification and approval of all WUAV weapon launches. Operators created search tasks, which dictated on the map those areas the operator wanted the UVs to specifically search. Operators also had scheduling tasks, but these were performed in collaboration with the automation; when the autonomous planner recommended schedules, operators accepted, rejected, or modified these plans. Details of the autonomous planner are provided in the next section.

A. Path-Planning and Task-Allocation Algorithm

The OPS-USERS system architecture is specifically designed to meet the challenges associated with an automated decision-making system integrated with a human operator on-the-loop. Two key challenges are 1) balancing the roles and responsibilities of the human operator and the automated planner, and 2) optimizing resource allocation to accomplish mission objectives. The system relies on the relative strengths of both humans and automation in that a human operator provides valuable intuition and field experience, while the automation provides raw numerical power and rapid optimization capability.

In OPS-USERS, decision making responsibility is layered to promote goal-based reasoning such that the human guides the autonomy but that the automation assumes the bulk of computation for optimization of task assignments. The automated planner is responsible for decisions requiring rapid calculations or optimization, and the human operator supervises the planner for high-level goals such as where to search and overall resource allocation (i.e., which tasks get included in the overall plan) as well as for tasks that require strict human approval, such as approving weapons release.

To allow the human and the automation to collaborate for task execution, the basic system architecture is divided into two major components, as shown in Fig. 1. The first is the distributed tactical planner, which is a network of onboard planning modules (OPMs) [17] that provides coordinated autonomy between the agents. Each agent carries a processor, which runs an instance of the OPM. The second is the ground control station (GCS), which consists of a centralized strategic planner called the central mission manager, and the operator interface.

A decentralized implementation was chosen for the tactical planner to allow rapid reaction to changes in the environment [52]. When appropriate, the decentralized task planner may modify the task assignment without affecting the overall plan quality (i.e., agents switch tasks), and it is able to make these local repairs faster through interagent communication than it could if it had to wait for the next update from the GCS. Furthermore, plans can be carried out even if the communication link with the GCS is intermittent or lost [53]. The architecture is scalable because

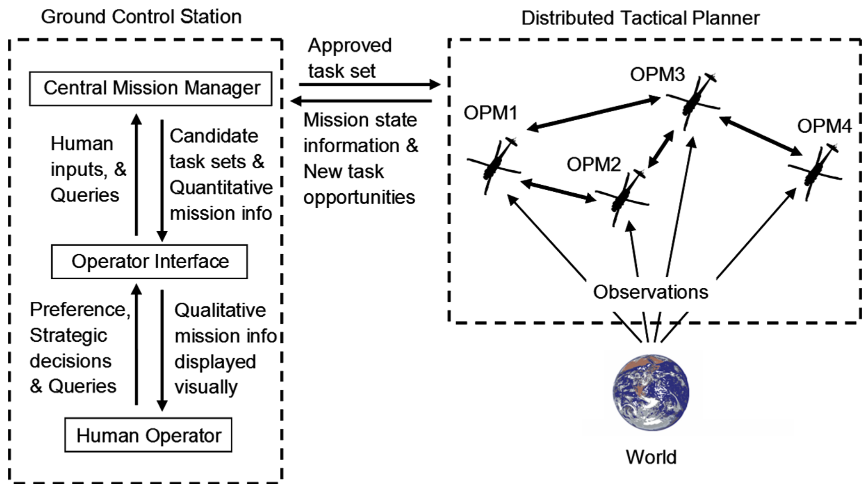


Fig. 1 OPS-USERS system architecture.

adding additional agents also adds computational capability, and the decentralized framework is robust to a single point of failure because no single agent is globally planning for the fleet.

The decentralized task planner used in OPS-USERS is the consensus-based bundle algorithm (CBBA), a decentralized, polynomial-time, market-based protocol [54]. CBBA consists of two phases that alternate until the assignment converges. In the first phase, task selection, agents select the set of tasks for which they get the highest reward. The agents place bids on the tasks they choose where the bids represent the marginal improvement in the score of their plan. In the second phase, conflict resolution, plan information is exchanged between neighbors and tasks go to the highest bidder. CBBA is guaranteed to reach a conflict-free assignment.

One key advantage of CBBA is that it is able to solve the multiple assignment problem where each agent is assigned a set of tasks (a plan), as opposed to solving the single assignment problem, where each agent is only assigned to their next task. Planning several tasks into the future improves effectiveness in complex missions [55,56].

Operators were shown the results of this bidding process through the display that showed unassigned tasks that could not be completed by one or more of the vehicles. However, if unhappy with the UV-determined search or track patterns, operators could create new tasks, in effect forcing the decentralized algorithms to reallocate the tasks across the UVs. This human–automation interaction scheme is one of high-level goal-based control, as opposed to more low-level vehicle-based control. Operators could never directly individually task a single vehicle. The operator interface is described in more detail in the next section.

B. Operator Interface

Participants interacted with the OPS-USERS simulation via two displays. The primary interface is a map display (Fig. 2). The map shows both geospatial 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. As in real-life scenarios, changing external conditions often require the human and the system to adapt, which are represented through rules of engagement (ROEs) received through the chat tool. Icons represent vehicles, targets of all types, and search tasks, and the symbology is consistent with MIL-STD 2525 [57]. In this interface, operators identify targets, approve weapon launches, and insert new search tasks as desired or dictated via the chat box. The performance plot in Fig. 2 gives operators insight into the automated planner performance, as the graph shows predicted plan score (red) versus current plan score (blue) of the system. When the predicted performance score is above the current score, the automated planner is effectively proposing that better performance could be achieved if the operator accepts the proposed plan (based on the planner’s prediction of how the vehicles will bid on the tasks).

When the automated planner generates a new plan that is at least 5% ‘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, the operator is taken to the schedule comparison tool (SCT) for conducting scheduling tasks in collaboration with the automation. Operators can elect to select the replan button at anytime. The SCT display then appears, showing three geometrical forms colored gray, blue, and green at the top of the display (Fig. 3), which are configural displays that enable quick comparison of 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 plan modification. The rectangular grid on the upper half of each shape represents the estimated area of the map that the UVs will search according to the proposed plan. The hierarchical priority ladders show the percentage of tasks assigned in high, medium, and low priority levels, respectively.

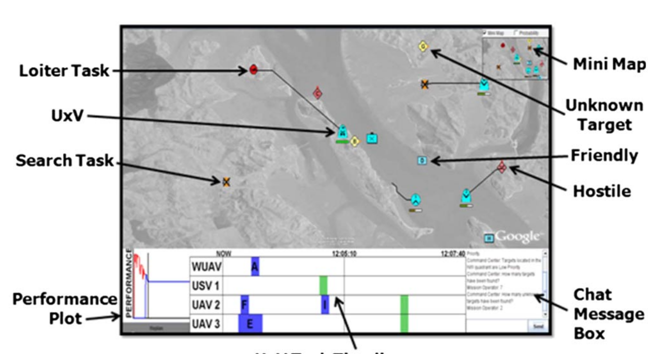


Fig. 2 Map display.

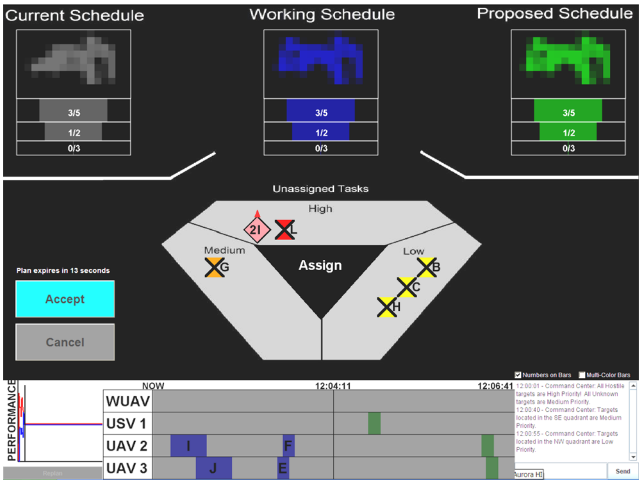


Fig. 3 Schedule comparison tool.

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 [58] to quickly compare the three schedules. This 'what if' query, which essentially is a preview display [59], represents a collaborative effort between the human and automation [60]. Operators adjust team coordination metrics at the task level as opposed to the individual vehicle level, which has been shown to improve single-operator control of a small number of multiple, independent robots [61]. Details of the OPS-USERS interface design and usability testing can be found elsewhere [62].

The automated planner in the original test bed 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 used in this experiment. Five nondimensional quantities, detailed next, 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. 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 searching.

Search/loiter tasks: As opposed to allowing the automation to 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 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, burned fuel more slowly, and did not have to refuel as often.

For this experiment, only a binary choice of 'on' or 'off' was allowed for each variable. Tversky and Kahneman [63] 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. Selecting a quantity gave it a weighting of 1.0 in the objective function of the automated planner, while deselecting 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 deselected, 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 Fig. 4. 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 used a radio button interface shown in Fig. 5. 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 (Fig. 2), were the three possible types of SCT that operators could use in the experiment.

IV. Experiment

An experiment was conducted to evaluate the impact of the dynamic objective function on decentralized UV control system performance, as well as the impact on human cognitive workload and operator satisfaction. This experiment addresses one of the research gaps 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 operators to express their desires to an automated planner.

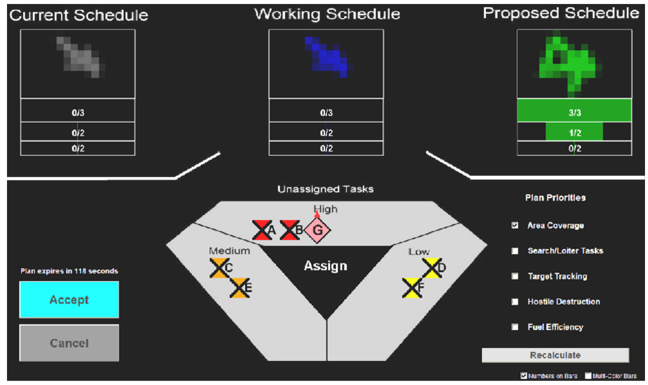


Fig. 4 Schedule comparison tool with checkbox interface.

A. Participants

Thirty undergraduate students, graduate students, and researchers were recruited for this experiment (21 men and 9 women). The age range of participants was 18–38 years old with an average age of 21.30 and a standard deviation of 3.98. Only one 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 [64]. Each participant filled out a demographic survey before the experiment that included age, gender, occupation, military experience, average hours of television viewing, video gaming experience, and perception of UAVs.

B. Apparatus

The experiment was conducted using two Dell 17 in. flat-panel monitors operated at 1280 × 1024 pixels and a 32-bit color resolution. The primary monitor displayed the test bed, and the secondary monitor showed a legend of the symbols used in the system. 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.

C. Experimental Design

Three scenarios, a practice scenario and two test scenarios, were designed for this experiment. Each scenario involved controlling four UVs (one of which was weaponized) in a mission to conduct surveillance of an area 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) to notify the operator when a target was detected so that the operator could view sensor information 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 by the automation.

Each scenario had 10 targets initially hidden to the operator. These targets always had a positive velocity and moved on preplanned 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. The operator was occasionally 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.

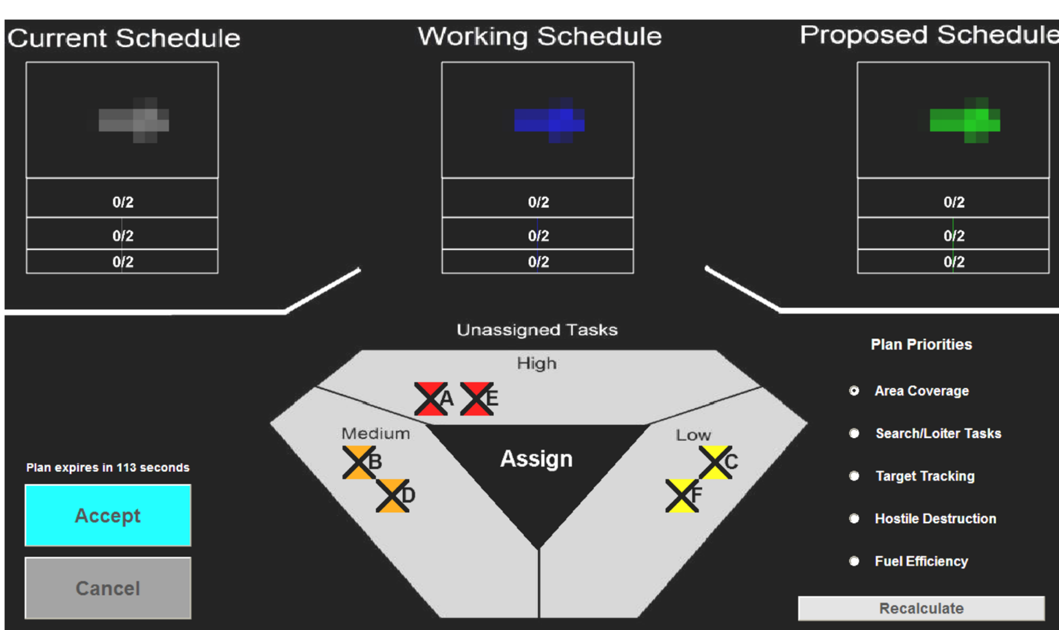


Fig. 5 Schedule comparison tool with radio button interface.

Table 1 Rules of engagement

	Time	Rules of Engagement
Standard mission	n/a	Track all found targets and destroy all hostile targets found.
Dynamic mission	0 min	Cover as much area as possible to find new targets. Tracking found targets is low priority. Do not destroy any hostiles.
	5 min	Conduct search tasks in southeast and southwest quadrants. Second priority: track all targets previously found. Do not destroy any hostiles.
	10 min	Track all targets closely; it is important not to lose any targets. Second priority: conserve fuel. Third priority: destroy hostile targets.
	15 min	All hostile targets are now high priority; destroy all hostiles.

D. Independent Variables

The experimental design was a 3×2 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 test bed objective function as described in Fig. 3, which was set a priori and which the operator did not have the opportunity to modify. 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, the automated planner optimized the vehicles to cover 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 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 mission priorities to guide their high-level decision making. The ROEs also specified when hostile target destruction was permitted. For the dynamic mission, every 5 min during the 20 min 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 shown in Table 1.

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.

E. Dependent Variables

The dependent variables for the experiment were mission performance, primary workload, secondary workload, SA, and subjective ratings of performance, workload, and confidence. Overall mission performance was measured by 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. Adherence to the ROEs presented to the operator during the dynamic mission 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 min of the mission, when covering area to find new targets was the highest priority, 3) percentage of targets found during the first 5 min of the mission, and 4) percent of time that targets were tracked between 10 and 15 min, 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 workload metric 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 [65].

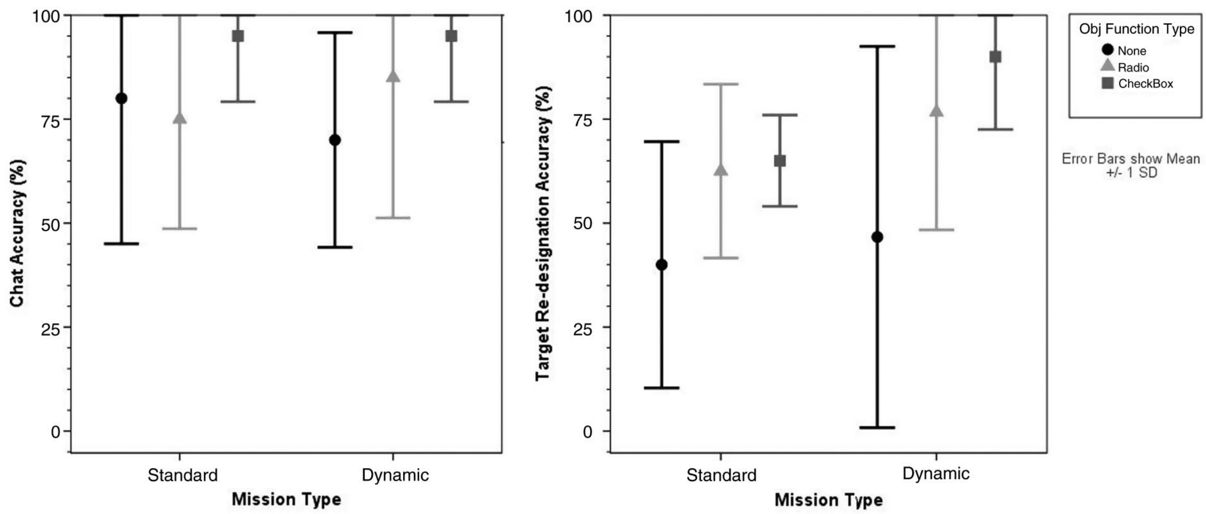
SA was measured through the accuracy percentage of responses to periodic chat box messages querying the participant about aspects of the mission. Additionally, four of the targets were originally designated as unknown. Chat messages provided intelligence information to the operator about whether these targets were actually hostile or friendly (based on their location on the map). It was up to the operator to redesignate these targets based on this information. Therefore, a second measure of SA was the ratio of correct redesignations 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 (where 1 is low and 5 is high). Subjective ratings provide an additional measure of workload and evaluate 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 [66].

F. Procedure

To familiarize each subject with the interface, a self-paced, slide-based tutorial was provided. Subjects then conducted a 15 min 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 five-question slide-based test. 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 attempt.

The actual experiment for each subject consisted of two 20 min 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



a) b)

Fig. 6 Chat accuracy and target redesignation comparison.

used to record the operator's actions on the screen. Subjects were paid \$10 per hour for the experiment and were told that a performance bonus of a \$100 gift card would be given to the individual who obtained the highest mission performance metrics (to encourage maximum effort).

V. Results and Discussion

A 3×2 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, nonparametric analyses were used.

A. Mission Performance and Situational Awareness

The results did not indicate any statistically significant differences in the overall mission performance metrics among the different types of objective function. Thus, regardless of which objective function type operators had, they generally achieved the same level of performance in terms of area searched, targets found and tracked, and hostiles destroyed. The second independent variable in the experiment, mission type, was a significant factor in mission performance. Regardless of the objective function used, operators found significantly more targets ($Z = -2.795$, $p = 0.005$) in the dynamic mission as compared to the standard mission. Given the inherent experimental confound that a static objective function could not be changed given dynamic external changes, this direct comparison is somewhat inherently biased, but it is noteworthy that, in the dynamic mission, 11% more targets were found.

Operator performance was then analyzed as an indicator of SA, which has been shown to be an important attribute in system performance [59,67]. The results showed that the omnibus test on the accuracy of responses to chat box questions was significant for objective function type, $\chi^2(2, N = 60) = 6.167$, $p = 0.046$, as was the omnibus test on the accuracy of redesignations of unknown targets, $\chi^2(2, N = 60) = 10.392$, $p = 0.006$.

Further, Mann–Whitney independent pairwise comparisons of chat accuracy showed that operators with the checkbox objective function had higher chat accuracy than the none and radio objective function users ($p = 0.057$ and $p = 0.013$, respectively) and marginally significantly different from the radio objective function. There was no significant difference between the radio and none objective functions ($p = 0.551$), indicating a similar level of SA, but lower than those participants with the checkbox objective function.

Mann–Whitney independent pairwise comparisons of redesignation accuracy showed that the none objective function was lower than checkbox and radio objective function accuracies ($p = 0.003$ and $p = 0.019$, respectively), but the checkbox and radio objective functions were not statistically different ($p = 0.342$). The box plots in Fig. 6 illustrate the results for chat accuracy and redesignation accuracy.

Operators using the checkbox objective function had significantly higher target redesignation and chat accuracies than the operators using the none objective function. Although the addition of the capability to modify the objective function did not significantly increase system performance, it may have enhanced SA. It is likely that the use of the checkbox interface, which supports multi-objective optimization and operator engagement, was the cause of this enhanced SA. Thus, operators who could manipulate the system in some way were more actively involved with goal management, which also led to improved secondary task performance.

Additionally, operators had significantly higher accuracy in the redesignation of unknown targets in the dynamic mission ($Z = -2.482$, $p = 0.013$) as compared to the standard mission. Operators had both higher mission performance and enhanced SA during the dynamic mission, where the ROEs changed every 5 min. It is possible that more frequent reminders of mission goals, through the changing ROEs, could have played a role in this increase in performance and SA.

B. Rules of Engagement Adherence and Violations

At the beginning of each scenario, participants would receive the initial ROE via a chat message. As shown previously in Table 1, the initial ROE in the dynamic mission specifically instructed participants that their highest priority was searching the area for new targets, while the standard mission had a more general ROE. The two missions were similar but with slightly different target locations to prevent learning effects. The results showed that operators found significantly more targets in the first 5 min of the dynamic mission as compared to the standard mission, $F(1, 27) = 25.357$, $p < 0.001$, regardless of the type of objective function used.

Further analysis of the dynamic mission results showed that the omnibus test on targets found in the first 5 min 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

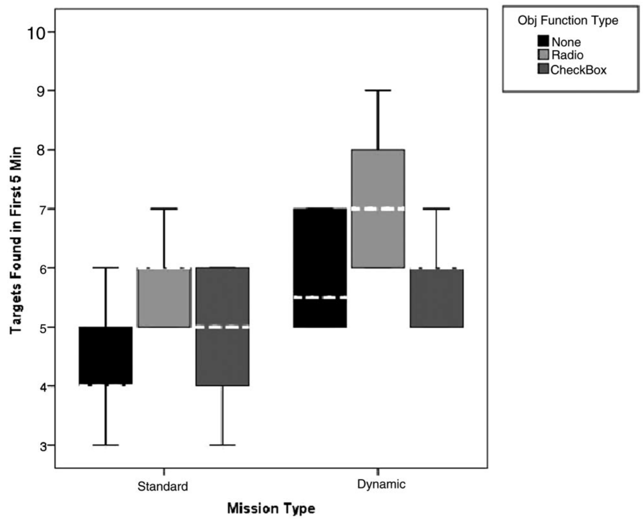


Fig. 7 Targets found in the first 5 min.

different ($p = 0.823$). Operators who used the radio objective function found more targets in the first 5 min of the dynamic mission. The box plots in Fig. 7 illustrate the results for number of targets found in the first 5 min.

ROEs 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 who received specific instructions at the start of the dynamic mission to search for new targets were able to find more targets during the first part of the mission. Additionally, operators using the radio objective function found more targets in the first 5 min of the dynamic mission. These results support the claim that providing the operator with a dynamic objective function could 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 toward 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 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 three operators who violated the ROEs by destroying a hostile target during the first 10 min 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, or inadequate system design.

C. Workload

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. In addition, it was found that there was no significant difference in average time spent in the SCT across the three objective function types. As can be expected, operators conducting the more complicated dynamic mission had significantly higher utilization, $F(1,27) = 5.216$, $p = 0.030$, and spent significantly more time in the SCT on average, $F(1,27) = 20.786$, $p < 0.001$, as compared to the standard mission.

Mental workload was also measured through embedded secondary task reaction times. 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 s into the mission, a prompt to create a search task at 300 s, another prompt to create a search task at 725 s, and finally a chat message question requiring a response at 1104 s.

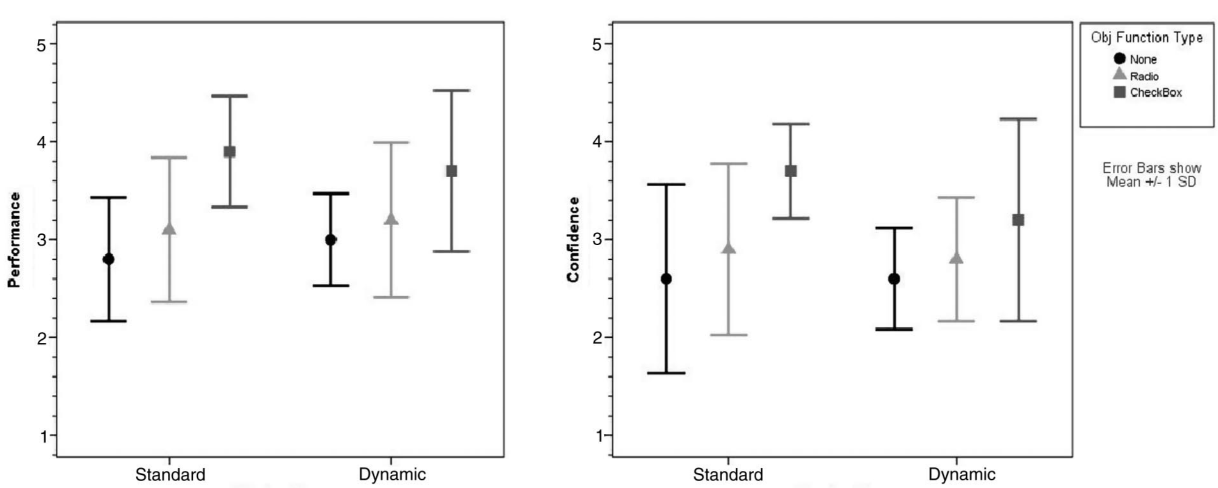


Fig. 8 Performance and confidence self-ratings comparison.

The omnibus test for the reaction time to the chat question at 235 s was significant for objective function type, $F(2, 26) = 8.839$, $p = 0.001$. Tukey pairwise comparisons showed that the none objective function produced slower reaction times as compared to the checkbox and radio conditions ($p = 0.001$ and $p = 0.002$, respectively). The checkbox and radio objective functions chat reactions times were not statistically different ($p = 0.703$).

The omnibus test for the reaction time to the chat question at 1104 s was significant for objective function type, $F(2, 26) = 3.411$, $p = 0.048$. Tukey pairwise comparisons showed that the only significant pairwise comparison was between the checkbox objective function and the none objective function ($p = 0.022$). Generally, operators using the checkbox objective function had faster chat reaction times.

These results show that, during the dynamic mission, operators using the checkbox objective function had significantly faster reaction times to a secondary task (chat message response) than operators using the none objective function. At one of those points, the operators using the radio objective function were also significantly faster. These results indicate that, at certain points during the mission, operators with access to a dynamic objective function were able to respond more quickly than operators using a static objective function, suggesting there was some spare mental capacity with these tools. However, overall utilization and subjective workload measures show that there were no differences across the three objective function types.

These are encouraging results, as they mean that the addition of another interface tool and associated task of manipulating the objective functions did not add any appreciable workload to operators and, in some cases, allowed operators to respond even more quickly. This is important because a significant concern in single-operator control of multiple UVs is high levels of workload [7], so it is critical that any additional decision-support tools do not overload operators.

D. Operator Strategies

Investigating the number of objective function modifications made by operators using the dynamic objective functions is indicative of operator strategies, and there was a significant difference between the strategies adopted by checkbox and radio objective function users. 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 twice the amount of checkbox operators, with an average of 28.3 modifications over the 20 min simulation as compared to 12.4 modifications for the checkbox operators.

Of all the 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.

Overall, radio objective function operators had a higher percentage of SCT sessions where they modified the objective function at least once, made twice 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 by switching between the variables under consideration, although this workload difference was not reflected in the overall time spent replanning, nor did subjective measures indicate any difference in workload as compared to the other conditions. However, radio objective function operators were the only group that violated any ROEs. This is a significant negative performance indicator in and of itself and suggests that although overall performance in terms of mission objectives was the same despite the presence of a decision-support tool, participants using the radio objective function committed more serious errors that those with either the checkbox function or no ability to modify the objective function.

E. 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 (where 1 is low and 5 is high). Participants were also given open-ended questions to prompt them to give general feedback.

The Kruskal–Wallis omnibus test on subjective performance rating was 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 Kruskal–Wallis omnibus test on the confidence rating was 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 ($p = 0.430$). The plots in Fig. 8 illustrate the self-rating results.

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 the expectation that use of a dynamic objective function would 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 one of the objective functions. 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.

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. In contrast, the operators using the limited radio objective function only had single objective optimization capabilities and were forced to perform numerous ‘what ifs’ 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. Eighty-seven percent of participants indicated that the automated planner was fast enough for this dynamic, time-pressured mission. Four of the 10 participants who used the radio objective function gave written complaints about the restriction to select only one variable as their top priority, and more complained verbally during training. This feeling of restriction in objective function choice is also reflected in the lower subjective ratings of the radio objective function.

A few participants also reported that they were frustrated because of perceived suboptimal automation performance. For example, one ‘none’ participant wrote, “the automated planner is fast but doesn’t generate an optimal plan,” and another ‘radio” operator wrote, “I did not always understand decisions made by the automated planner. . .namely it would not assign tasks. . .while some vehicles were seemingly idle.” Also, one ‘checkbox’ 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.” These perceptions, despite the optimized solutions produced by the algorithm, are crucial to understand because developing an appropriate level of operator trust in the automation is necessary for effective performance of this dynamic, high-risk mission [68].

In this particular experiment, the interface, tutorial, and practice session were designed to enable a novice user to achieve proficient use of the system in 45 min, and thus simplicity was emphasized. As a result, these participants had little knowledge of the inner workings of the task-allocation and path-planning algorithm, and thus it is likely that they were not aware of all of the variables and constraints that the algorithm took into account when creating plans. This is likely representative of future real-world operations, where human controllers will have limited knowledge of exactly how the ‘black box’ automation works. We attempted to increase automation transparency via the SCT, which gave operators the opportunity for sensitivity analysis, to change the algorithm’s objective function, and to attempt to directly modify the plan. Despite these attempts at greater transparency, when the final plans did not seem ‘logical’ to the operator (regardless of the actual plan quality), trust in the automated planner decreased. Further understanding of how and why operators perceive algorithms to be suboptimal and how we, both human factors engineers and controls engineers, can work together to address this gap is an important area of future research.

VI. Conclusions

To meet the increasing demand for unmanned vehicles (UVs) across a variety of civilian and military purposes, reduce operating expenses, and enhance UV capabilities, human operators will need to supervise multiple UVs simultaneously. To successfully conduct this form of supervisory control, operators will need the support of significant embedded collaborative autonomy. Although 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 lower system performance or increase operator 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.

One way to promote more human–automation collaboration to achieve superior multi-UV system performance is to provide operators with the ability to change the planner’s objective function in real time. A dynamic objective function increases automation transparency and reduces automated planner brittleness, which enhances the ability of a human operator to work successfully with the automation. To this end, a test bed was designed that included the ability for operators to dictate single variable objective functions (radio), or multivariate objective functions (checkbox). An experiment was conducted to assess this collaboration scheme under dynamic and static mission goal environments.

The results of this experiment established that, although allowing operators the ability to change either a single or multiple variables did not significantly improve mission performance metrics, operator situational awareness was improved as was adherence to changing mission priorities. Moreover, operators generally preferred such interactions. However, in the case of the single-variable objective function manipulation, rules of engagement (ROE) violations occurred, which was not the case for any other condition. Because this method required extensive interaction to achieve an acceptable plan, the chance of error was increased, likely leading to these violations.

Given that operator workload is a major concern, it is interesting to note that, despite the fact that operators were working harder during the mission with changing ROEs, they also performed better, suggesting that they were nowhere near their maximum cognitive capacity. These conclusions are further evidenced by both subjective and objective workload metrics, which demonstrate that, even with the added decision-support tool to change the objective functions, operators were not working any harder than those without the tool (in both the standard and dynamic missions) and, in some cases, had more spare mental capacity.

Related to workload, operators using a dynamic objective function with multi-objective capabilities needed fewer modifications to achieve an acceptable plan. One of the most revealing results of the experiment was 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, and providing the capability to modify the objective function for multi-objective optimization can aid in developing this trust.

Future work could include introducing more options for manipulating the values of the weightings in the objective function, for example, rating each value as high, medium, or low or ranking the values in priority order. In this experiment, all the weightings were always equal. Also, it is unclear from this work 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. In addition, providing more training and information about the automation before the experiment could influence the operator’s interactions with the automation. An important avenue of future research could be to quantify the impact of the degree of automation transparency as well as other methods of enhancing human–computer collaboration, including training.

Finally, 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. The five quantities were chosen because of their direct relationship to the system performance metrics that were measured in this experiment and upon which operators were told they would be judged. These specific objective function quantities may not be the best possible selections. Further investigation is necessary and underway to derive the full set of objective functions that could be used in various application scenarios. Monte Carlo techniques have already been employed with this test bed to explore the impact of communication delays and different search strategies and such techniques will be used to explore different objective functions and weighting concepts. This research highlights a fundamental issue in such complex command and control scenarios, which is that with, changing mission priorities and dynamic constraints and variables, the definition of ‘optimal’ will always be difficult to define.

Acknowledgments

This research is sponsored by Aurora Flight Sciences and the U.S. Office of Naval Research. Karl Kulling of Aurora Flight Sciences provided extensive algorithm support. Cameron Fraser of the Massachusetts Institute of Technology (MIT) Aerospace Controls Laboratory provided test bed support. Han-Lim Choi of the MIT Aerospace Controls Laboratory provided consensus-based bundle algorithm support.

References

[1] Westwood, J., “Global Prospects for AUVs,” *Offshore Technology Conference*, Houston, TX, 2001.

[2] Sweeten, B. C., Royer, D., and Keshmiri, S., “Meridian UAV Flight Performance Analysis Using Analytical and Experimental Data,” AIAA Infotech@Aerospace Conf., Seattle, WA, AIAA Paper 2009-1899, April 2009.

[3] Girard, A. R. and Hedrick, J. K., "Border Patrol and Surveillance Missions Using Multiple Unmanned Air Vehicles," *43rd IEEE Conf. on Decision and Control*, IEEE Publ., Piscataway, NJ, 2004.
[4] Casbeer, D. W., Kingston, D. B., Beard, R. W., McLain, T. W., Li, S. M., and Mehra, R., "Cooperative Forest Fire Surveillance Using a Team of Small Unmanned Air Vehicles," *International Journal of Systems Science*, Vol. 37, No. 6, 2006, pp. 351–360.
doi:10.1080/00207720500438480
[5] Haddal, C. C., and Gertler, J., "Homeland Security: Unmanned Aerial Vehicles and Border Surveillance," Congressional Research Service, Rept. RS21698, Washington, D.C., July 2010.
[6] "United States Air Force Unmanned Aircraft Systems Flight Plan 2009-2047," United States Air Force, Washington, D.C., 2009.
[7] Cummings, M. L., Bruni, S., Mercier, S., and Mitchell, P. J., "Automation Architecture for Single Operator, Multiple UAV Command and Control," *International Command and Control Journal*, Vol. 1, No. 2, 2007, pp. 1–24.
[8] Alighanbari, M., and How, J. P., "Robust Decentralized Task Assignment for Cooperative UAVs," AIAA Guidance, Navigation, and Control Conf. and Exhibit, Keystone, CO, AIAA Paper 2006-6454, Aug. 2006.
[9] Bertuccelli, L. F., Choi, H.-L., Cho, P., and How, J. P., "Real-Time Multi-UAV Task Assignment in Dynamic and Uncertain Environments," AIAA Guidance, Navigation, and Control Conf., Chicago, AIAA Paper 2009-5776, Aug. 2009.
[10] Office of the Secretary of Defense, "Unmanned Aircraft Systems (UAS) Roadmap, 2005-2030," U.S. Dept. of Defense, Washington, D.C., 2005.
[11] Naval Studies Board, "Autonomous Vehicles in Support of Naval Operations," National Research Council, Washington, D.C., 2005.
[12] Clare, A. S., Hart, C. S., and Cummings, M. L., "Assessing Operator Workload and Performance in Expeditionary Multiple Unmanned Vehicle Control," 48th AIAA Aerospace Sciences Meeting, Orlando, FL, AIAA Paper 2010-763, Jan. 2010.
[13] Cummings, M. L., and Mitchell, P. J., "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, 2008, pp. 451–460.
doi:10.1109/TSMCA.2007.914757
[14] Schmidt, D. K., "A Queuing Analysis of the Air Traffic Controller's Workload," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 8, No. 6, 1978, pp. 492–498.
doi:10.1109/TSMC.1978.4310003
[15] Rouse, W. B., *Systems Engineering Models of Human-Machine Interaction*, Elsevier North Holland, New York, NY, 1983, pp. 79–86.
[16] J. Van Den Berg, Ferguson, D., and Kuffner, J., "Anytime Path Planning and Replanning in Dynamic Environments," *IEEE International Conference on Robotics and Automation (ICRA)*, IEEE Publ., Piscataway, NJ, 2006, pp. 2366–2371.
[17] How, J. P., Fraser, C., Kulling, K. C., Bertuccelli, L. F., Toupet, O., Brunet, L., et al., "Increasing Autonomy of UAVs," *IEEE Robotics & Automation Magazine*, Vol. 16, No. 2, 2009, pp. 43–51.
doi:10.1109/MRA.2009.932530
[18] Marquez, J. J., "Human–Automation Collaboration: Decision Support for Lunar and Planetary Exploration," Ph.D. Thesis, Dept. of Aeronautics and Astronautics, Massachusetts Inst. of Technology, Cambridge, MA, 2007.
[19] Silverman, B. G., "Human–Computer Collaboration," *Human–Computer Interaction*, Vol. 7, No. 2, 1992, pp. 165–196.
doi:10.1207/s15327051hci0702_2
[20] Smith, P., McCoy, E., and Layton, C., "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, 1997, pp. 360–371.
doi:10.1109/3468.568744
[21] Guerlain, S. A., "Using the Critiquing Approach to Cope With Brittle Expert Systems," *Human Factors and Ergonomics Society 39th Annual Meeting*, San Diego, CA, 1995, pp. 233–237.
[22] Polson, P., and Smith, N., "The Cockpit Cognitive Walkthrough," *10th Symposium on Aviation Psychology*, Ohio State University, Columbus, OH, 1999, pp. 427–432.
[23] Cummings, M. L., and Bruni, S., "Collaborative Human–Automation Decision Making," *Handbook of Automation LXXVI* (ed.), S. Y. Nof, Springer, New York, 2009, pp. 437–448.
[24] Forest, L. M., Kahn, A., Thomer, J., and Shapiro, M., "The Design and Evaluation of Human-Guided Algorithms for Mission Planning," *Human Systems Integration Symposium*, Annapolis, MD, 2007.
[25] Anderson, D., Anderson, E., Lesh, N., Marks, J., Mirtich, B., Ratajczak, D., et al., "Human-Guided Simple Search," *Proceedings of the 17th Conference on Artificial Intelligence (AAAI-00)*, AAAI Press, Menlo Park, CA, 2000, pp. 209–216.
[26] Scott, S. D., Lesh, N., and Klau, G. W., "Investigating Human–Computer Optimization," *Conference on Human Factors in Computer Systems (CHI)*, ACM, New York, NY, 2002, pp. 155–162.
[27] Bitton, E., and Goldberg, K., "Hydra: A Framework and Algorithms for Mixed-Initiative UAV-Assisted Search and Rescue," in *IEEE International Conf. on Automation Science*, IEEE Publ., Piscataway, NJ, 2008.
[28] Carlsen, D., Malone, M., Kollat, J., and Simpson, T. W., "Evaluating the Performance of Visual Steering Commands for User-Guided Pareto Frontier Sampling During Trade Space Exploration," *ASME 2008 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Brooklyn, NY, 2008, pp. 499–509.
[29] Cummings, M. L., "Automation Bias in Intelligent Time Critical Decision Support Systems," AIAA 1st Intelligent Systems Technical Conf., Chicago, AIAA Paper 2004-6313, Sept. 2004.
[30] Ruff, H. A., Narayanan, S., and Draper, M. H., "Human Interaction with Levels of Automation and Decision-Aid Fidelity in the Supervisory Control of Multiple Simulated Unmanned Air Vehicles," *Presence : Teleoperators and Virtual Environments*, Vol. 11, No. 4, 2002, pp. 335–351.
doi:10.1162/105474602760204264
[31] Burstein, M. H., and McDermott, D. V., "Issues in the Development of Human–Computer Mixed-Initiative Planning Systems," *Cognitive Technology: In Search of a Humane Interface*, edited by B. Gorayska, and J. Mey, Elsever Science, Amsterdam, 1996, pp. 285–303.
[32] Kramer, L. A., and Smith, S. F., "Optimizing for Change: Mixed-Initiative Resource Allocation with the AMC Barrel Allocator," *3rd International NASA Workshop on Planning and Scheduling for Space*, Houston, TX, 2002.
[33] Howe, A. E., Whitley, L. D., Barbulescu, L., and Watson, J. P., "Mixed Initiative Scheduling for the Air Force Satellite Control Network," *2nd International NASA Workshop on Planning and Scheduling for Space*, San Francisco, CA, 2000.
[34] Tillerson, M., Breger, L., and How, J. P., "Distributed Coordination and Control of Formation Flying Spacecraft," *American Control Conference*, Denver, CO, 2003, pp. 1740–1745.
[35] Bauer, F. H., Hartman, K., How, J. P., Bristow, J., Weidow, D., and Busse, F., "Enabling Spacecraft Formation Flying through Spaceborne GPS and Enhanced Autonomy Technologies," *Space Technology*, Vol. 20, No. 4, 2001, pp. 175–185.
[36] McCamish, S. B., "Distributed Autonomous Control of Multiple Spacecraft During Close Proximity Operations," Ph.D. Thesis, Dept. of Electrical and Computer Engineering, Naval Postgraduate School, Monterey, CA, 2007.
[37] Curtis, S. A., Rilee, M. L., Clark, P. E., and Marr, G. C., "Use of Swarm Intelligence in Spacecraft Constellations for the Resource Exploration of the Asteroid Belt," *3rd International Workshop on Satelllite Constellations and Formation Flying*, Pisa, Italy, 2003.
[38] Truszkowski, W., Hinchey, M., Rash, J., and Rouff, C., "NASA's Swarm Missions: The Challenge of Building Autonomous Software," *IT Professional*, Vol. 6, No. 5, 2004, pp. 47–52.
doi:10.1109/MITP.2004.66
[39] Sim, L., Cummings, M. L., and Smith, C. A., "Past, Present, and Future Implications of Human Supervisory Control in Space Missions," *Acta Astronautica*, Vol. 62, 2008, pp. 648–655.
doi:10.1016/j.actaastro.2008.01.029

[40] Fox, J., Breed, J., Moe, K., Pfister, R., Truszkowski, W., Uehling, D. L., et al., "User-Centered Design of Spacecraft Ground Data Systems at NASA's Goddard Space Flight System," in *2nd International Symposium on Spacecraft Ground Control & Data Systems SCDII*, 1999.

[41] Sierhuis, M., Bradshaw, J. M., Acquisti, A., van Hoof, R., Jeffers, R., and Uszok, A., "Human-Agent Teamwork and Adjustable Autonomy in Practice," *Proceedings of the 7th International Symposium on Artificial Intelligence, Robotics, and Automation in Space*, Nara, Japan, 2003.

[42] Chu, R. W., Mitchell, C. M., and Jones, P. M., "Using the Operator Function Model and OFMspert as the Basis for an Intelligent Tutoring System: Towards a Tutor/Aid Paradigm for Operators of Supervisory Control Systems," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, Vol. 26, No. 1*, Vol. 25, 1995, pp. 1054–1075.

[43] Mitchell, C. M., "GT-MSOCC: A Domain for Research on Human–Computer Interaction and Decision Aiding in Supervisory Control Systems," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, Vol. 26, No. 1*, Vol. 17, 1987, pp. 553–572.

[44] Kortenkamp, D., Schreckenghost, D., and Martin, C., "User Interaction with Multi-Robot Systems," *Multi-Robot Systems: From Swarms to Intelligent Automata*, edited by A. C. Schultz, and L. E. Parker, Kluwer Academic, Dordrecht, The Netherlands, 2002, pp. 177–184.

[45] Hanson, M. L., Roth, E., Hopkins, C. M., and Mancuso, V., "Developing Mixed-Initiative Interaction with Intelligent Systems: Lessons Learned from Supervising Multiple UAVs," AIAA 1st Intelligent Systems Technical Conf., Chicago, AIAA Paper 2004-6310, 2004.

[46] Miller, C., Funk, H., Wu, P., Goldman, R., Meisner, J., and Chapman, M., "The Playbook Approach to Adaptive Automation," *Human Factors and Ergonomics Society 49th Annual Meeting*, Orlando, FL, 2005, pp. 15–19.

[47] Bruni, S., and Cummings, M., "Human Interaction with Mission Planning Search Algorithms," *Human Systems Integration Symposium*, Arlington, VA, 2005.

[48] Cummings, M. L., and Bruni, S., "Human–Automation Collaboration in Complex Multivariate Resource Allocation Decision Support Systems," *Intelligent Decision Technologies*, Vol. 4, No. 2, 2010, pp. 101–114.
doi:10.3233/IDT-2010-0072

[49] Thorner, J. L., "Trust-Based Design of Human-Guided Algorithms," M.S. Thesis, Massachusetts Inst. of Technology, Cambridge, MA, 2007.

[50] Rieskamp, J., and Hoffrage, U., "Inferences Under Time Pressure: How Opportunity Costs Affect Strategy Selection," *Acta Psychologica*, Vol. 127, No. 2, 2008, pp. 258–276.
doi:10.1016/j.actpsy.2007.05.004

[51] Boussemart, Y., Donmez, B., Cummings, M. L., and Las Fargeas, J., "Effects of Time Pressure on the Use of an Automated Decision Support System for Strike Planning," *15th International Symposium on Aviation Psychology (ISAP '09)*, Dayton, OH, 2009.

[52] Whitten, A. K., "Decentralized Planning for Autonomous Agents Cooperating in Complex Missions," M.S. Thesis, Massachusetts Inst. of Technology, Cambridge, MA, 2010.

[53] Southern, D. N., "Human-Guided Management of Collaborating Unmanned Vehicles in Degraded Communication Environments," M.E. Thesis, Massachusetts Inst. of Technology, Cambridge, MA, 2010.

[54] Choi, H. L., Brunet, L., and How, J. P., "Consensus-Based Decentralized Auctions for Robust Task Allocation," *IEEE Transactions on Robotics and Automation*, Vol. 25, No. 4, No. 4, 2009, pp. 912–926.

[55] Choi, H.-L., Whitten, A. K., and How, J. P., "Decentralized Task Allocation for Heterogeneous Teams with Cooperation Constraints," *American Control Conf. (ACC)*, Baltimore, MD, 2010, pp. 3057–3062.

[56] Whitten, A. K., Choi, H.-L., Johnson, L. B., and How, J. P., "Decentralized Task Allocation With Coupled Constraints in Complex Missions," *American Control Conf. (ACC)*, San Francisco, CA, 2011, pp. 1642–1649.

[57] "Interface Standard, Common Warfighting Symbology," U.S. Dept. of Defense Rept. MIL-STD-2525B, Jan. 1999.

[58] Gibson, J. J., *The Ecological Approach to Visual Perception.* Houghton Mifflin, Boston, MA, 1979.pp. 147–263.

[59] Wickens, C. D., and Hollands, J. G., *Engineering Psychology and Human Performance*, 3rd ed., Prentice–Hall, Upper Saddle River, N.J., 2000. pp. 260–261, 523–525.

[60] Layton, C., Smith, P. J., and McCoy, C. E., "Design of a Cooperative Problem-Solving System for En-Route Flight Planning—An Empirical Evaluation," *Human Factors*, Vol. 36, No. 1, pp. 94–116, March 1994.
doi:10.1177/001872089403600106

[61] Goodrich, M. A., Johansen, J., McLain, T. W., Anderson, J., Sun, J., and Crandall, J. W., "Managing Autonomy in Robot Teams: Observations from Four Experiments," *2nd ACM/IEEE International Conference on Human Robot Interaction*, ACM, New York, NY, 2007, pp. 25–32.

[62] Fisher, S., "Replan Understanding for Heterogenous Unmanned Vehicle Teams," M.E. Thesis, Dept. of Electrical Engineering and Computer Science, Massachusetts Inst. of Technology, Cambridge, MA, 2008.

[63] Tversky, A., and Kahneman, D., "Judgment Under Uncertainty: Heuristics and Biases," *Science*, Vol. 185, No. 4157, 1974, pp. 1124–1131.
doi:10.1126/science.185.4157.1124

[64] Cummings, M. L., Clare, A. S., and Hart, C. S., "The Role of Human–Automation Consensus in Multiple Unmanned Vehicle Scheduling," *Human Factors*, Vol. 52, No. 1, 2010, pp. 17–27.
doi:10.1177/0018720810368674

[65] Cummings, M. L., and Guerlain, S., "Using a Chat Interface as an Embedded Secondary Tasking Tool," *Human Performance, Situation Awareness and Automation: Current Research and Trends*, Vol. 1, edited by D. A. Vincenzi, M. Mouloua, and P. A. Hancock, Routledge, London, 2004, pp. 240–248.

[66] Parasuraman, R., and Riley, V., "Humans and Automation: Use, Misuse, Disuse, Abuse," *Human Factors*, Vol. 39, No. 2, 1997, pp. 230–253.
doi:10.1518/001872097778543886

[67] Endsley, M. R., "Toward a Theory of Situation Awareness in Dynamic Systems," *Human Factors*, Vol. 37, No. 1, 1995, pp. 32–64.
doi:10.1518/001872095779049543

[68] Lee, J. D., and See, K. A., "Trust in Automation: Designing for Appropriate Reliance," *Human Factors*, Vol. 46, No. 1, 2003, pp. 50–80.
doi:10.1518/hfes.46.1.50_30392

C. Rouff
Associate Editor