Intelligent Decision Technologies 6 (2012) 221–231
DOI 10.3233/IDT-2012-0138
IOS Press

Assessing operator strategies for real-time replanning of multiple unmanned vehicles

Andrew S. Clare[a,*], Pierre C.P. Maere[b] and M.L. Cummings[a]
[a]*Aeronautics and Astronautics Department, Massachusetts Institute of Technology, Cambridge, MA, USA*
[b]*Delft University of Technology, Delft, The Netherlands*

Abstract. Future unmanned vehicles systems will invert the operator-to-vehicle ratio so that one operator controls a decentralized network of heterogeneous unmanned vehicles. This study examines the impact of allowing an operator to adjust the rate of prompts to view automation-generated plans on system performance and operator workload. Results showed that the majority of operators chose to adjust the replan prompting rate. The initial replan prompting rate had a significant framing effect on the replan prompting rates chosen throughout a scenario. Higher initial replan prompting rates led to significantly lower system performance. Operators successfully self-regulated their task-switching behavior to moderate their workload.

Keywords: Unmanned vehicles, automation, decentralized algorithms, human-machine interface, human supervisory control, human-computer interaction, mental workload

1. Introduction

A future concept of operations for controlling Unmanned Vehicles of different types (UxVs) is one of a single, forward-deployed soldier supervising multiple, heterogeneous (air, sea, land) UxVs [26]. In order to achieve this concept of operations, significant collaborative autonomy will have to be embedded within and across these teams of vehicles, so that the vehicles can execute basic operational and navigational tasks autonomously [9]. Operators will supervise these vehicles by providing high level direction to achieve mission goals. They will need to comprehend a large amount of information while under time pressure to make effective decisions in a dynamic environment. They will be assisted by automated planners to reduce workload. As a result, human management of the automated planner is crucial, as auto-planners do not always generate accurate solutions. Though fast and able to handle complex computation 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 [33].

Effective decision-making in this complex system requires the right balance between human and automated decisions. Automated decisions are possible in situations where the decision-making agent can follow a set of predetermined rules, known as rule-based behavior [30]. For example, planning an optimal path for a UxV to take to its next task can be done by the automation, as long as the environment is known with moderate certainty [34]. Other complex decisions, such as prioritizing tasks or interpreting camera imagery, require a human decision-maker. Humans have the ability to conduct such "knowledge-based reasoning" [30] because of our superior improvisation, flexibility, and inductive reasoning skills as compared to computers. Striking the right balance between human and computers for complex decision-making has been explored by many researchers [8,17,20,32], and we build on that work here with a particular focus on multiple UxV control in a changing and uncertain environment.

In a previous experiment, the impact of increasing automation replan prompting rates on operator performance and workload was examined [6]. The operator

*Corresponding author: Andrew Clare, Massachusetts Institute of Technology, 77 Massachusetts Avenue 33-311, Cambridge, MA 02139, USA. Tel.: +1 617 253 0993; Fax: +1 617 253 4196; E-mail: aclare@mit.edu.

ISSN 1872-4981/12/$27.50

was prompted to replan at various intervals, but could choose to replan whenever he or she desired. When replanning, the operator was presented with plans created by the automated planner, which he or she could accept, reject, or attempt to modify manually. Results showed that the rate of replan prompting had a significant impact on workload and performance, with higher replan prompting rates resulting in both degraded human and system performance. Surveys conducted after each trial revealed that approximately 35% of the participants were frustrated by the automated planner. Results from this experiment also showed that operators who collaborated with the automated planner had significantly higher performance and lower workload than those who ignored automation requests for replan consideration [10].

The experiment described in this paper builds on the previous experiment by allowing operators to set the rate of replan prompting. The purpose of this research is to see if there is a replanning rate that human operators prefer, and whether there is an effect on performance. The rest of this paper is organized as follows. In Section 2, background literature in human-computer collaboration for path planning and task allocation is introduced, followed by the detailed description of the testbed created that allows for this investigation in Section 3. In Section 4, the methodology to investigate operator strategies for collaborative decision making in a dynamic environment is explained. Finally, in Section 5 the results are presented and conclusions are drawn.

2. Background

When controlling multiple highly automated UxVs, human supervision can be beneficial due to the uncertainty inherent in command and control operations, 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 [1,2,31], scheduling UxVs [5], or trade space exploration for large scale design optimization [3]. Events are often unanticipated and automated planners are typically unable to account for and respond to unforeseen problems [18,29].

Due to the dynamic and uncertain nature of such missions, real-time control of decentralized UxVs requires a planning algorithm which allows for rapid reactions to changes in the environment [35]. As opposed to a centralized algorithm, where all decisions are made by a single agent, the decentralized framework is robust to a single point of failure, since no single agent is globally planning for the fleet. Plans can be carried out even if the communication link with the human operator is intermittent or lost. The architecture is scalable, since adding additional agents also adds computational capability.

In a decentralized UxV system, the human operator provides high-level *goal-based* control, as opposed to more low-level *vehicle-based* control, by approving which tasks should be completed by the vehicles. The list of operator-approved tasks is referred to as a *strategic-level plan*. In such architectures, operators do not directly individually task a single vehicle. For the remainder of this paper, replanning means that the operator chooses to compare the algorithm's suggested new strategic-level plan with the current plan, regardless of whether the operator accepts the new plan. When appropriate, the decentralized task planner can modify the *tactical-level plan* without human intervention, which includes changing the task assignment without affecting the overall plan quality (i.e. agents switch tasks). The algorithm is able to make these local repairs faster through inter-agent communication than it could if it had to wait for the next update from the human operator. While modern automated planners are capable of generating new schedules on the order of seconds [4,35], human approval of these frequent schedule updates can cause high operator workload with potentially negative performance consequences [6].

Operators engaged in these dynamic, high workload environments must both concentrate attention on the primary task (i.e., monitoring vehicle progress and identifying targets) and also be prepared for automation replan alerts. This need to concentrate on a task, yet maintain a level of attention for alerts requires both interrupt and task-driven processing. The allocation of attention between these two incurs cognitive costs that negatively impact overall system performance [25]. Poor attention allocation has been shown to be a significant contributor to poor operator performance in single operator control of multiple unmanned vehicles [7,15].

Researchers have attempted to address the problem of high operator workload and poor attention allocation strategies through interface design [14,22] and workload prediction through modeling [21,24,27]. Neither approach allows for real-time adjustments to the system to deal with the dynamic command and control environment. Other researchers have focused on changing the role of the human and computer through research on adjustable autonomy and adaptive automation. Such

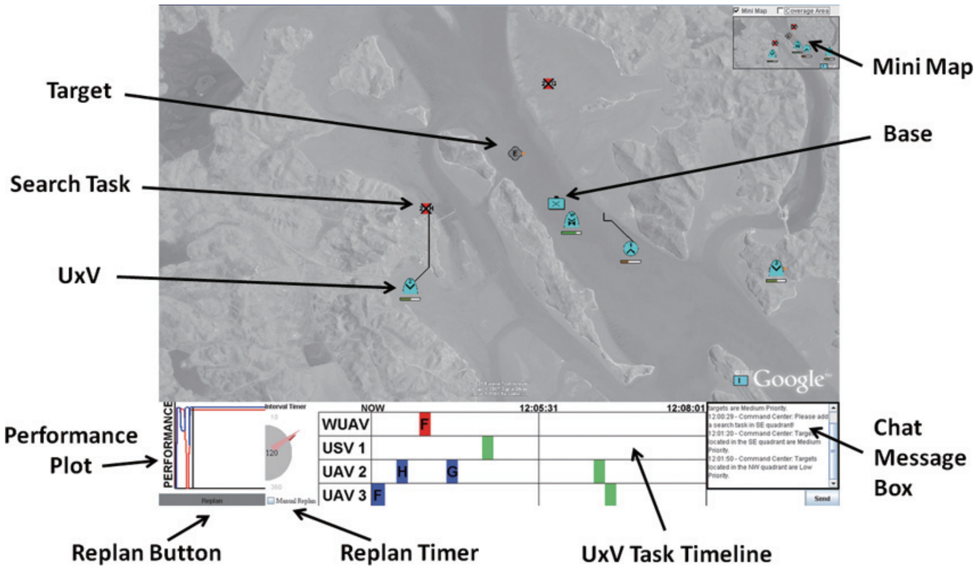


Fig. 1. The Map Display. (Colours are visible in the online version of the article; http://dx.doi.org/10.3233/IDT-2012-0138)

adjustments can be made during a mission, either with the human operator instigating the change through adjustable autonomy [5,16], or with the computer automatically deciding to adjust the level of automation through adaptive automation [23,28]. 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 allowing the human operator to adjust the rate of replan prompting. Neither the human operator nor the computer controls whether the vehicles are more or less autonomous. No matter what the rate of replan prompting is, the decentralized network of vehicles continues to execute the last plan that was approved and has the ability to modify the tactical-level plan by changing the assignment of tasks to individual UxVs [35]. Adjusting the replan prompting rate changes the rate at which the operator is notified to compare the current strategic-level plan with an updated plan. The operator still has the choice to modify the strategic-level plan at any time if the environment changes, new targets are detected, or the vehicles require attention. Allowing the operator to self-regulate his or her task switching behavior could enable operators to successfully perform their objectives. Thus, there is a need to conduct research to determine the impact of allowing operators to adjust the rate at which they are prompted to replan. Our attempt to address this research question is detailed in the next section.

3. Experimental test bed

This effort utilizes a collaborative, multiple UxV 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 [19]; all the decision support displays described here have operated actual small air and ground UxVs in real-time.

The mission of interest is to search, track, and destroy enemy targets. The objective of the operator is to command multiple, heterogeneous unmanned vehicles for the purpose of searching an area of responsibility for new targets, tracking targets, and approving weapons launch. All targets are initially hidden, but once a target is found, it is designated as hostile, unknown, or friendly, and given a priority level by the user. Hostile targets are tracked by one or more of the vehicles until they are destroyed by a weaponized Unmanned Aerial Vehicle (UAV). Operators must approve all weapon launches. Unknown targets are revisited as often as possible, tracking target movement.

Provided with intelligence via a text messaging "chat" box, the operator has the ability to re-designate unknown targets or create search tasks for emergent targets. The primary interface used by the operator is a Map Display, shown in Fig. 1. The operator is assisted by an automated planner in scheduling the search tasks

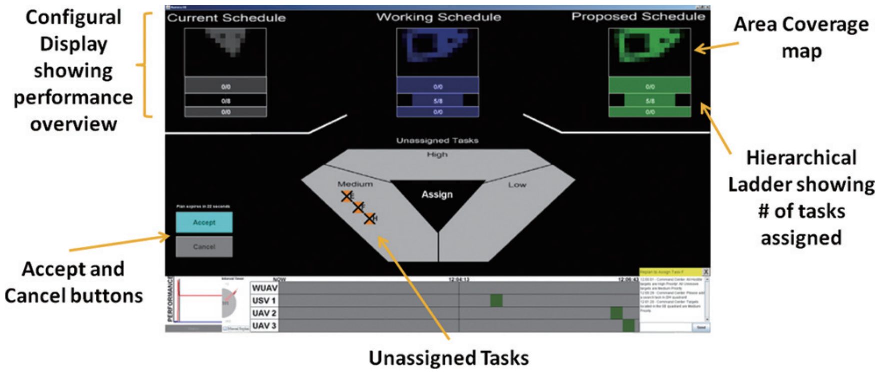


Fig. 2. The Schedule Comparison Tool (SCT). (Colours are visible in the online version of the article; http://dx.doi.org/10.3233/IDT-2012-0138)

and target tracking assignments to be completed by the UxVs.

In order to aid the operator in understanding the current state of the UxVs and their progress towards mission goals, a decision support interface, called the Schedule Comparison Tool (SCT), was developed, shown in Fig. 2. Details of the interface design and usability testing are provided in previous research [13].

Given previous results that showed that at least one-third of operators ignored preset automation replanning intervals [10], a new component was added to the map display, as shown in Fig. 3. The operator can adjust the replan prompting interval by using the Replan Interval Dial. The dial can be set to a preferred value, as long as it is between the two boundary values, 1 and 360 seconds. The maximum setting of 360 seconds was chosen to be three times longer than the maximum initial prompting interval of 120 seconds from the previous experiment [10]. In addition, automated replan prompting intervals could be disabled by selecting the manual replan option.

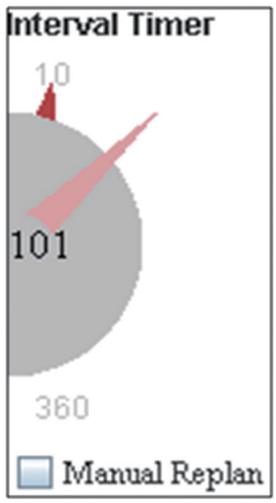


Fig. 3. Replan Interval Dial. (Colours are visible in the online version of the article; http://dx.doi.org/10.3233/IDT-2012-0138)

4. Methodology

In futuristic multiple unmanned vehicle settings, the rate at which a human operator must confirm or alter plans has been shown to have a significant impact on operator workload and system performance [10]. In this previous work, operators were prompted to replan at three different intervals: 30 seconds, 45 seconds, and 120 seconds. The replanning prompt was given through the green illumination of the replan button (Fig. 1) and an aural replan alert sounded when a schedule was available that the automation deemed better than the current schedule. Although the automation could generate plans on the order of seconds, to experimentally control for task load, operators would only be notified of a new plan at the intervals of 30 seconds, 45 seconds, and 120 seconds.

The experiment detailed here provides operators with the ability to adjust the rate of replan prompting, which was not possible in the previous experiment. The aim is to determine whether there is a replanning rate that human operators prefer given the ability to adjust it, and whether there is an effect on operator workload, system performance, or subjective operator assessment of the system. In this experiment, the independent variable was the initial replan prompting interval. The three levels for the independent variable were also 30, 45, and 120 seconds. The key difference is that operators could change this prompting interval to any interval between 1 and 360 seconds. Operators also had the option to turn off the replan prompt, so that there would be no notification of when the automation had a new plan for the operator to evaluate.

The dependent variables included objective workload metrics, mission performance metrics, and subjective operator ratings of confidence and workload. Workload was measured via a utilization metric (i.e., percent busy time) because utilization has proven to be sensitive to changes in workload in similar multiple tasking, time-pressured scenarios [11,12]. Operators were considered "busy" when performing one of the following tasks: creating search tasks to specify locations on the map where UxVs must search for targets; identifying targets by looking at the imagery and designating a target type and priority level; approving weapons launches on hostile targets; chat messaging with the virtual, remote command center; and replanning in the SCT. The mission performance dependent variables included the percentage of area covered, the percentage of targets found, the percentage of time that targets were tracked, and the percentage of hostile targets destroyed. Finally, a survey was presented to the operator after each of the three scenarios to gather subjective self-ratings of confidence and workload on a 1–5 Likert scale.

This experiment was conducted using a Dell 17" flat panel monitor operated at 1280 $\times$ 1024 pixels and 32-bit color resolution. 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 worn by each participant during the experiment. All data regarding each participant's interactions with the system for controlling the simulated UxVs was recorded automatically by the system.

In order to familiarize each subject with the interfaces in Figs 1 and 2, a self-paced, slide-based tutorial was provided, which typically took subjects approximately twenty minutes to complete. Then, subjects had a ten-minute practice session during which the experimenter walked the subject through all the necessary functions to use the interface and to develop schedules before accepting them. Each subject was given the opportunity to ask the experimenter questions regarding the interface and mission during the tutorial and practice session.

The actual experiment for each subject consisted of three 15 minute test sessions. The three possible initial replan prompting intervals were 30 seconds, 45 seconds and 120 seconds. Each subject experienced each of these initial rates in a counterbalanced and randomized order to control for learning effects. Subjects were able to change the intervals between prompting using the Replan Interval Dial. Each scenario was different, but similar in difficulty. The interface recorded all operator actions.

Subjects were selected from a sample population similar to that which the military is interested in for the types of operations simulated by this interface. The subject population consisted of twenty-nine subjects: 20 men and 9 women. Ages ranged from 18 to 31 years with a mean of 23.6 years and standard deviation of 3.8 years. All subjects had previous experience with this simulation testbed without the Replan Interval Dial. Sixteen subjects had participated in a previous experiment with this testbed at least a month prior, and were labeled "Experienced", while 13 subjects received equivalent practice time of 30 minutes with the interface (without the Replan Interval Dial) immediately prior to this experiment, and were labeled "Inexperienced". About a third of subjects had military experience (ROTC, Air Force Academy, or Active Duty). Each subject filled out a demographic survey prior to the experiment that included age, gender, occupation, military experience, comfort level with computers, and video gaming experience.

5. Results and discussion

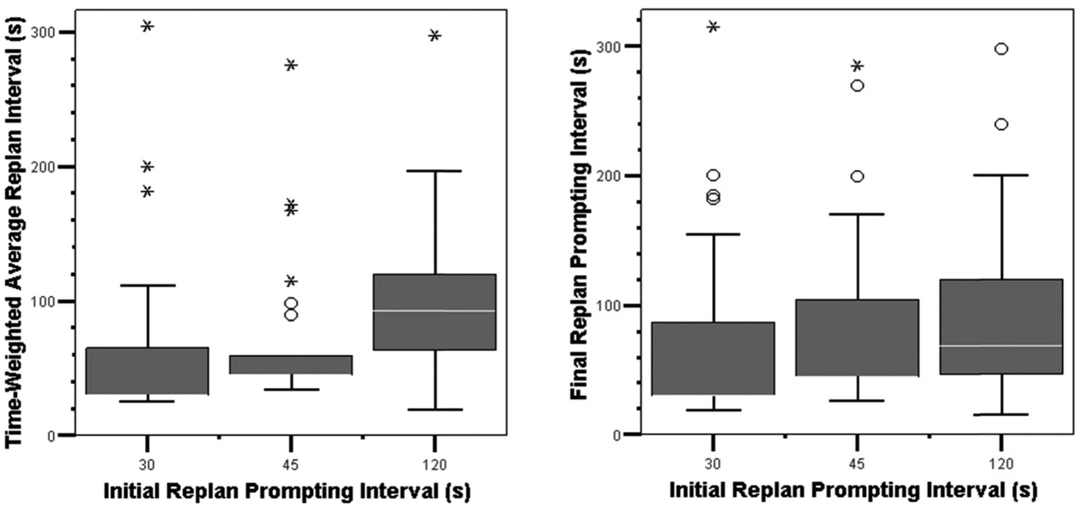
Non-parametric analyses were used for all dependent variables ($\alpha = 0.05$) due to non-normal distributions and heteroscedasticity.

5.1. Replan Interval Dial strategy

An analysis of the operators' strategies was conducted based on experimental data. First, an analysis was conducted of how much operators used the Replan Interval Dial to modify the replan prompting rate. Of the 87 test trials, the Replan Interval Dial was utilized in 54 of them, or 62% of the trials. Two of the 29 operators never changed the prompting interval in any of the 3 scenarios. Another 9 operators only made changes to the prompting interval in 1 of the 3 scenarios. Of those 9 operators, 7 made a change to the prompting interval when the initial prompting interval was 120 seconds. For the 30 second initial replan prompting rate, 12 of the 29 operators never changed the prompting interval. For the 45 second initial replan prompting interval, 13 operators never changed the prompting interval. For the 120 second initial replan prompting interval, however, only 8 operators never changed the prompting in-

Table 1
Replan prompting strategy of operators across the initial replan prompting intervals

	30s Initial replan interval	45s Initial replan interval	120s Initial replan interval
Number of changes to replan prompting interval (Mean and St. Dev)	2.41 ± 1.8	2.75 ± 2.8	2.24 ± 2.1
Time-Weighted average replan prompting interval	Mean: 65.46s Median: 30.00s St. Dev: 68.4s	Mean: 70.22s Median: 45.00s St. Dev: 56.1s	Mean: 102.24s Median: 92.81s St. Dev: 57.9s
Final replan prompting interval	Mean: 71.48s Median: 30.00s St. Dev: 73.9s	Mean: 80.75s Median: 45.00s St. Dev: 72.5s	Mean: 93.64s Median: 68.50s St. Dev: 67.6s
Number of operators who ended the scenario with replan prompting off	2	1	1



(a) Time-weighted average replan promping interval (b) Final replan prompting interval

Fig. 4. Replan prompting intervals chosen by operators across initial intervals. (a) Time-weighted average replan promping interval; (b) Final replan prompting interval.

terval. These results reflect a finding from the previous experiment [10], where operators generally could not wait the full 120 seconds between replans.

For operators who did use the Replan Interval Dial, the average number of Replan Interval Dial changes was evaluated. For operators who made at least one change, the average number of changes to the Replan Interval Dial was 2.44 over the course of the 15 minute session. The breakdown for the average number of changes for each of the initial replan intervals is shown in Table 1.

Upon further analysis, it appears that the initial replan interval had an impact on the Replan Interval Dial settings chosen by each operator. A time-weighted average of the replan prompting intervals chosen by the operator was calculated using Eq. (1).

$$Average = \sum_i (p_i * t_i) \Big/ T \quad (1)$$

Where p_i is the prompting interval set by the operator, t_i is the time that the prompting interval was in effect for, and T is the total simulation time of 15 minutes. If, during a particular trial, the operator ever turned off the replan prompt, that trial was excluded from the time-weighted average replan interval data set. This occurred in only 10 out of the 87 trials. The Kruskal-Wallis omnibus test showed a significant difference in the time-weighted average replan prompting interval based on initial replan prompting interval, $\chi 2$ $(2, N = 77) = 15.261, p < 0.001$, with the lowest time-weighted average replan prompting interval at the 30 second initial interval, as shown in Fig. 4(a). As might be expected, the initial replan prompting interval had a "framing" effect on the operator, leading the operator to choose replan intervals close to the initial replan prompting interval, as shown by the median values in Table 1.

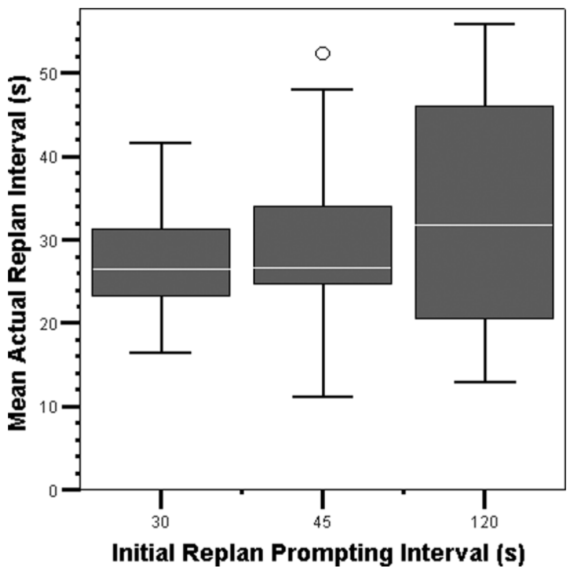


Fig. 5. Mean actual replanning intervals across initial prompting intervals.

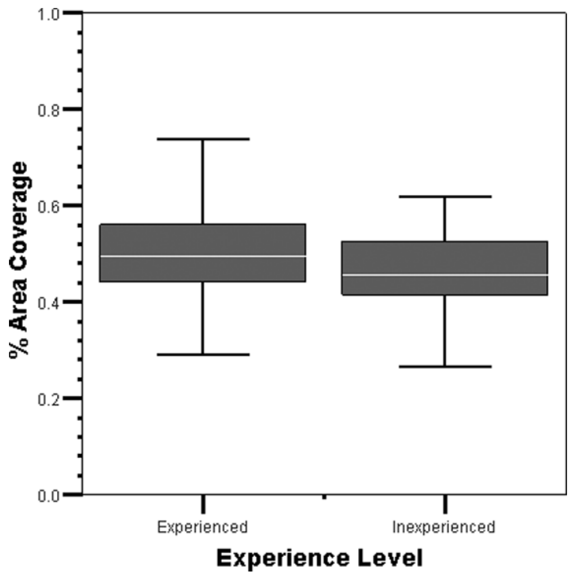


Fig. 6. Area coverage performance across experience levels.

In addition, the initial replan prompting interval had a significant effect on the final prompting interval ($p =$ 0.041). The average final replan prompting intervals are shown in Table 1. This data excludes subjects who ended the experiment with the replan prompt set to "off". One subject chose to switch the replan prompt to "off" as the last Replan Interval Dial setting for all 3 of the trials, while one other subject switched the prompt to "off" during the 30 second initial replan prompting interval scenario. Although this framing effect occurred, the Replan Interval Dial settings were typically higher than the initial value for the 30 and 45 second intervals, and lower than the initial value for the 120 second interval scenario, as shown in Fig. 4 and Table 1.

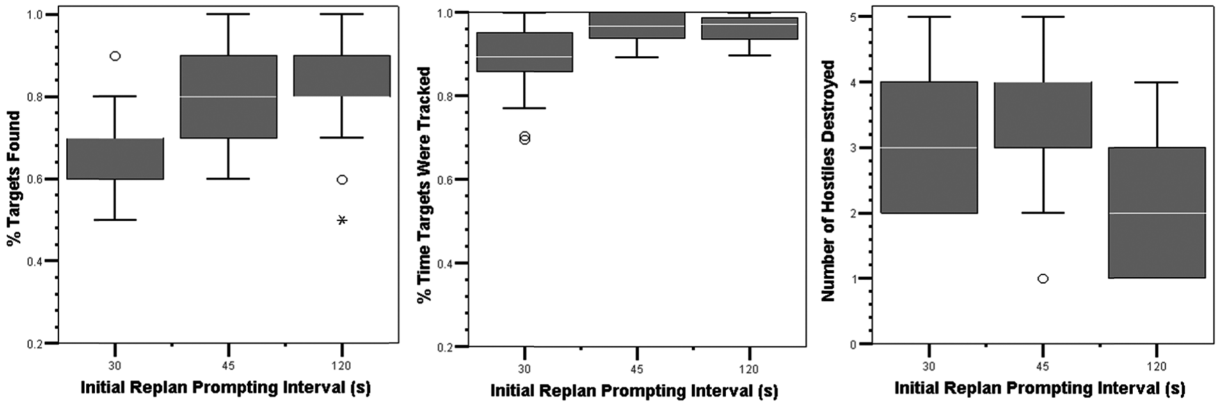
Despite this framing effect on the prompting interval, operators still had independent preferences for the *actual* frequency with which they replanned, regardless of prompting. A distinction must be made between a prompt to replan, which is an automated notification to view a new schedule, versus actually choosing to replan by pressing the replan button, which launches the Schedule Comparison Tool. When analyzing the rate at which operators *actually pressed the replan button* in order to view new plans, it was found that the initial replan prompting interval had no significant effect on the mean interval between actual replans ($p =$ 0.466), as shown in Fig. 5. The initial replan prompting interval did not influence the average rate at which operators actually replanned.

It should be noted that the initial replan prompting interval did have a significant effect on the variance of the mean actual replanning interval ($p <$ 0.001), with standard deviations of 6.5, 9.5, and 13.0 seconds for the 30, 45, and 120 initial replan intervals respectively, which can also be seen in Fig. 5. This indicates that an initially longer prompting interval led to a variety in replanning strategies, ranging from rapid to infrequent replanning, whereas a shorter initial prompting interval led operators to take rapid replanning strategies with less variation.

Finally, as might be expected, there was a significant correlation between the time-weighted average replan prompting interval and the mean actual replan interval, $\rho =$ 0.325, $p =$ 0.004. The prompting interval set by the operator aligned closely with the actual interval at which operators replanned. This again reflects a result from the previous experiment [10], where some operators consented to replan at the rate at which they were prompted to replan. In this experiment, by allowing operators to set their own replan prompting interval, operators were more likely to follow the automation's prompts to replan.

5.2. *Performance metrics*

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. In terms of the effect of experience on system performance, a trend was



(a) Targets found (b) Time targets were tracked (c) Hostiles destroyed

Fig. 7. Performance metrics across initial replan prompting intervals. (a) Targets found (b) Time targets were tracked (c) Hostiles destroyed.

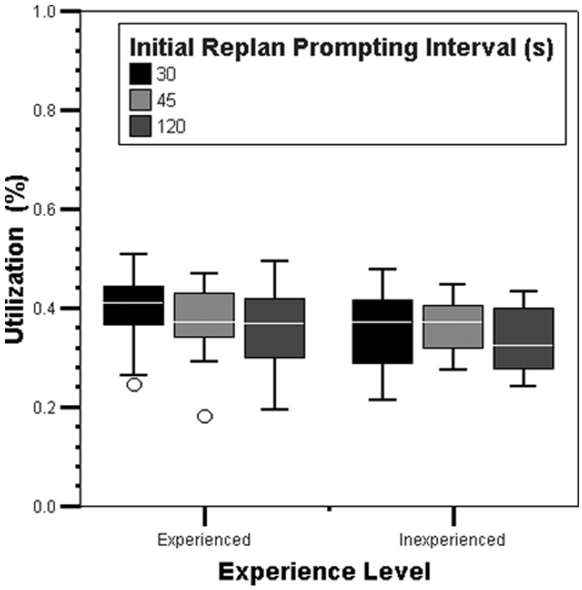


Fig. 8. Utilization across experience levels and initial replan interval.

found towards more experienced users having higher area coverage. A Mann-Whitney Independent test confirmed that there was a marginally significant difference in area coverage between experienced and inexperienced operators, $Z = -1.939$, $p = 0.052$. These results are shown in Fig. 6.

Significant differences were found in mission performance based on initial replan prompting interval. The Kruskal-Wallis omnibus test showed a significant difference in targets found based on initial replan prompting interval, $\chi 2(2, \mathrm{N} = 87) = 18.368, p < 0.001$, with the lowest number of targets found at the 30 second initial interval, as shown in Fig. 7a. A similar trend was found for the percentage of time that targets were tracked, $\chi 2(2, \mathrm{N} = 87) = 17.947, p < 0.001$, with the lowest tracking times at the 30 second initial interval, as shown in Fig. 7b. Finally, there was a significant difference in the number of hostiles destroyed based on initial interval, $\chi 2(2, \mathrm{N} = 87) = 26.936$, $p < 0.001$, with the most number of hostiles destroyed at the 45 second interval, as shown in Fig. 7(c). These findings support previous experimental results [6] showing that even starting operators at too rapid of a rate of replan prompting, at the 30 second interval, can lead to lower system performance. There were no significant correlations between the time-weighted average prompting interval and any of the performance metrics.

5.3. *Workload metrics*

Descriptive statistics were calculated for the workload metric, utilization, which measured the percent busy time of operators during their missions. A trend was found for experienced operators to have higher utilization than inexperienced operators, as shown in Fig. 8. A Mann-Whitney Independent test confirmed that there was a marginally significant difference in utilization between experienced and inexperienced operators, $Z = -1.870, p = 0.061$. Although both sets of operators received the same total training time, operators who had used the simulation testbed in a previous experiment instead of immediately prior to this experiment were potentially working harder. This trend held true for both operators who modified the replan prompting interval and for operators who did not use the Replan Interval Dial. It was also shown above that these experienced operators had marginally higher performance in terms of area coverage, which may be related to their higher workload. The experienced opera-

Table 2
Subjective ratings across initial replan prompt intervals

Metric	Initial replan prompt interval (s)	Mean	Std Dev	Mode	Median
Confidence self-rating	30	3.00	0.60	3	3
	45	3.00	0.76	3	3
	120	2.97	0.73	3	3
Workload self-rating	30	2.83	0.66	3	3
	45	2.69	0.76	3	3
	120	2.76	0.64	3	3

Table 3
Summary of experimental findings

Category	Metric	Initial replan prompting interval	Experience level
Replan interval dial strategy	Time-weighted average replan prompting interval	*Average interval aligned with initial interval* (*$p < 0.001$*)	Indistinguishable ($p = 0.694$)
	Final prompting interval	*Final interval aligned with initial interval* (*$p = 0.041$*)	Experienced had longer final interval ($p = 0.061$)
	Mean interval between actual replans	Indistinguishable ($p = 0.466$)	Indistinguishable ($p = 0.705$)
	Variance of the mean actual replanning interval	*Larger initial interval led to wider variation in actual replanning interval* (*$p < 0.001$*)	Indistinguishable ($p = 0.101$)
System performance	% Area coverage	Indistinguishable ($p = 0.214$)	*Experienced* (*$p = 0.052$*)
	% Targets found	*45 and 120 second intervals* (*$p < 0.001$*)	Indistinguishable ($p = 0.381$)
	% Time targets tracked	*45 and 120 second intervals* (*$p < 0.001$*)	Indistinguishable ($p = 0.534$)
	Hostiles destroyed	*45 second interval* (*$p < 0.001$*)	Indistinguishable ($p = 0.676$)
Primary workload	Utilization	Indistinguishable ($p = 0.318$)	Experienced ($p = 0.061$)
Subjective ratings	Workload	Indistinguishable ($p = 0.931$)	Indistinguishable ($p = 0.141$)
	Confidence	Indistinguishable ($p = 0.838$)	Indistinguishable ($p = 0.821$)

tors may have spent more time guiding the automation to search a greater amount of the area of interest.

No significant differences were found in utilization between the three initial replan intervals. This result is distinct from the results of the previous experiment, where the initial replan prompting interval had a significant effect on utilization and a shorter fixed prompting interval caused higher operator workload [10]. By allowing operators to adjust the rate of replan prompting via the Replan Interval Dial, operators were able to successfully moderate their workload such that there was no difference in workload across the three scenarios.

5.4. Subjective self-rating metrics

A survey was provided at the end of each mission asking the participant for a subjective rating of his or her workload and confidence on a Likert scale from 1–5. There were no significant correlations among the final replan prompting interval and the time-weighted average replan prompting interval with either of the subjective ratings. Also, there were no significant differences in subjective ratings based on experience level or initial replan interval. Statistics are shown in Table 2. The results for the subjective self-ratings of workload are similar to the utilization results, an objective measure of workload, showing that there were no differences in workload across the three scenarios. Both sets of data show that operators were able to moderate their workload despite starting at different initial replan prompting intervals.

6. Conclusions

An experiment was conducted to examine the impact of allowing an operator to adjust the rate at which he or she is prompted to view new plans generated by an automated planner when supervising a decentralized network of multiple heterogeneous unmanned vehicles. This capability was used in 62% of all trials and was used heavily when the initial replan interval was long (120 seconds between replan prompts).

Results showed that the initial replan prompting interval had a "framing" effect on the operator in three key ways. First, the operators typically chose prompting intervals that were close to the initial replan prompting interval. Second, if the initial replan prompting interval was short, operators tended to consistently replan rapidly. By starting at an initially longer prompting interval, however, operators used a wider variety of actual replanning frequencies, ranging from rapid to infrequent replanning. Third, there was a difference in system performance based on the initial replan prompting interval. The results of this experiment confirmed previous results showing that rapid initial rates of replan prompting can cause lower overall system performance, while the highest level of system performance was achieved with a 45 second initial replan prompting interval. All of these results demonstrate the importance of determining the appropriate rate for automation alerts to consider schedule changes, as different initial rates significantly changed operator strategy and system performance. All of the results are summarized in Table 3, where the conditions with superior results are shown in bold.

These results also indicate that providing operators with the ability to adjust the rate at which they were prompted to view automation-generated plans affected workload. The prompting interval set by the operator aligned closely with the actual interval at which operators replanned. Operators, on average, chose to decrease their workload when initially prompted at 30 or 45 second intervals, but chose to increase their workload when initially prompted at 120 second intervals. This data confirmed previous results showing that operators generally could not wait the full 120 seconds between replans, potentially due to boredom. Finally, by allowing operators to adjust the rate of replan prompting via the Replan Interval Dial, operators were able to successfully moderate their workload such that there was no difference in subjective and objective workload across the scenarios despite different initial replan prompting intervals.

While all test subjects had the same total training time, there were some marginally significant differences in performance and workload between those who were trained immediately before this experiment and those who participated in a previous experiment. Experienced participants worked slightly harder with slightly better performance than their less experienced counterparts.

Although the allocation of attention between the mission itself and determining the appropriate replan prompting rate may have hidden costs, the benefits to the operator in terms of self-regulating his or her task-switching behavior and moderating workload levels are important for future unmanned systems operations in dynamic and uncertain environments. Given the framing bias and performance ramifications for high replanning rates seen in this effort, further research is necessary to more extensively investigate the impact of allowing operators to adjust the rate of replan prompting. While providing this capability did not lead to an overall performance improvement, incorporating the flexibility to allow operators to adjust the frequency at which the automation generates new plans for approval into future unmanned vehicles systems designs could help operators avoid high workload situations that could lead to costly or deadly mistakes.

Acknowledgments

This research is funded by the Office of Naval Research (ONR) and Aurora Flight Sciences. Olivier Toupet provided extensive support to generate the algorithms that run the simulation. Professor Jonathan How and the members of the MIT Aerospace Controls Laboratory provided the experimentation testbed support.

References

[1] D. Anderson, E. Anderson, N. Lesh, J. Marks, B. Mirtich, D. Ratajczak and K. Ryall, Human-Guided Simple Search, in: *Proceedings of Seventeenth Conference on Artificial Intelligence* (*AAAI-00*), Austin, TX, 2000.

[2] E. Bitton and K. Goldberg, Hydra: A Framework and Algorithms for Mixed-Initiative UAV-Assisted Search and Rescue, in: *Proceedings of IEEE International Conference on Automation Science*, Washington D.C., 2008.

[3] 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, in: *Proceedings of ASME 2008 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference*, Brooklyn, NY, 2008.

[4] H.-L. Choi, L. Brunet and J.P. How, Consensus-Based Decentralized Auctions for Robust Task Allocation, *IEEE Transactions on Robotics* **25**(4) (2009), 912–926.

[5] A.S. Clare, *Dynamic Human-Computer Collaboration in Real-time Unmanned Vehicle Scheduling*, Masters of Science Dissertation, Massachusetts Institute of Technology, 2010.

[6] A.S. Clare, C.S. Hart and M.L. Cummings, Assessing Operator Workload and Performance in Expeditionary Multiple Unmanned Vehicle Control, in: *Proceedings of 48th AIAA Aerospace Sciences Meeting*, Orlando, FL, 2010.

[7] J.W. Crandall and M.L. Cummings, Attention Allocation Efficiency in Human-UV Teams, in: *Proceedings of AIAA Infotech@Aerospace Conference*, Rohnert Park, CA, 2007.

[8] M.L. Cummings and S. Bruni, Human-Automation Collaboration in Complex Multivariate Resource Allocation Decision Support Systems, *Intelligent Decision Technologies* **4**(2) (2010), 101–114.

[9] 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* **1**(2) (2007), 1–24.

[10] M.L. Cummings, A.S. Clare and C.S. Hart, The Role of Human-Automation Consensus in Multiple Unmanned Vehicle Scheduling, *Human Factors* **52**(1) (2010), 17–27.

[11] M.L. Cummings and S. Guerlain, Developing Operator Capacity Estimates for Supervisory Control of Autonomous Vehicles, *Human Factors* **49**(1) (2007), 1–15.

[12] M.L. Cummings and C.E. Nehme, Modeling the Impact of Workload in Network Centric Supervisory Control Settings, in: *Proceedings of 2nd Annual Sustaining Performance Under Stress Symposium*, College Park, MD, 2009.

[13] S. Fisher, *Replan Understanding for Heterogenous Unmanned Vehicle Teams*, Masters of Engineering Dissertation, Massachusetts Institute of Technology, 2008.

[14] L.M. Forest, A. Kahn, J. Thomer and M. Shapiro, The Design and Evaluation of Human-Guided Algorithms for Mission Planning, in: *Proceedings of HSIS'07 – Human Systems Integration Symposium*, Annapolis, MD, 2007.

[15] M.A. Goodrich, M. Quigley and K. Cosenzo, Task Switching and Multi-Robot Teams, in: *Proceedings of 3rd International Workshop on Multi-Robot Systems*, Washington D.C., 2005.

[16] M.A. Goodrich and A.C. Schultz, Human-Robot Interaction: A Survey, *Foundations and Trends in Human-Computer Interaction* **1**(3) (2007), 203–275.

[17] V. Gorodetsky, O. Karsaev, V. Samoylov and S. Serebryakov, Agent-based distributed decision-making in dynamic operational environments, *Intelligent Decision Technologies* **3**(1) (2009), 35–57.

[18] S.A. Guerlain, Using the critiquing approach to cope with brittle expert systems, in: *Proceedings of Human Factors and Ergonomics Society 39th Annual Meeting*, San Diego, CA, 1995.

[19] 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* **16**(2) (2009), 43–51.

[20] D. Lambert and J. Scholz, Ubiquitous command and control, *Intelligent Decision Technologies* **1**(3) (2007), 157–173.

[21] A. Majumdar and W.Y. Ochieng, The factors affecting air traffic controller workload: A multivariate analysis based upon simulation modeling of controller workload, in: *Proceedings of 81st Annual Meeting of the Transportation Research Board*, Washington, D.C., 2002.

[22] J.J. Marquez, *Human-Automation Collaboration: Decision Support for Lunar and Planetary Exploration*, Ph.D. Dissertation, Massachusetts Institute of Technology, 2007.

[23] C. Miller, H. Funk, P. Wu, R. Goldman, J. Meisner and M. Chapman, The playbook approach to adaptive automation, in: *Proceedings of Human Factors and Ergonomics Society 49th Annual Meeting*, Orlando, FL, 2005.

[24] C.A. Miller, Modeling human workload limitations on multiple UAV control, in: *Proceedings of 48th Annual Meeting of the Human Factors and Ergonomics Society*, New Orleans, LA, 2004.

[25] Y. Miyata and D.A. Norman, Psychological issues in support of multiple activities, in: *User Centered System Design: New Perspectives on Human Computer Interaction*, D.A. Norman and S.W. Draper, eds, Lawrence Erlbaum, Hillsdale, NJ, 1986, pp. 265–284.

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

[27] C. Nehme, *Modeling Human Supervisory Control in Heterogeneous Unmanned Vehicle Systems*, Ph.D. Dissertation, Massachusetts Institute of Technology, 2009.

[28] 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, 1992.

[29] P. Polson and N. Smith, The cockpit cognitive walkthrough, in: *Proceedings of 10th Symposium on Aviation Psychology*, Columbus, OH, 1999.

[30] J. Rasmussen, Skills, rules, and knowledge: Signals, signs, and symbols, and other distinctions in human performance models, *IEEE Transactions on Systems, Man, and Cybernetics* **13**(3) (1983), 257–266.

[31] S.D. Scott, N. Lesh and G.W. Klau, Investigating human-computer optimization, in: *Proceedings of Conference on Human Factors in Computer Systems* (*CHI*), Minneapolis, MN, 2002.

[32] T.B. Sheridan, Function allocation: Algorithm, Alchemy, or Apostasy? *International Journal of Human-Computer Studies* **52** (2000), 203–216.

[33] 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* **27**(3) (1997), 360–371.

[34] A. Stentz, Optimal and Efficient path planning for partially-known environments, in: *Proceedings of IEEE International Conference on Robotics and Automation*, San Diego, CA, 1994.

[35] A.K. Whitten, *Decentralized Planning for Autonomous Agents Cooperating in Complex Missions*, Masters of Science Dissertation, Massachusetts Institute of Technology, 2010.