The Impact of Intelligent Aiding for Multiple Unmanned Aerial Vehicle Schedule Management

M.L. Cummings
A.S. Brzezinski
J.D. Lee

Abstract

There is increasing interest in designing systems such that the current many-to-one ratio of operators to unmanned vehicles (UVs) can be inverted. Instead of lower-level tasks performed by today's UV teams, the sole operator would focus on high-level supervisory control tasks. A key challenge in the design of such single-operator systems will be the need to minimize periods of excessive workload that arise when critical tasks for several UVs occur simultaneously. Thus some kind of decision support is needed that facilitates an operator's ability to evaluate different action alternatives for managing a multiple UV mission schedule in real-time. This paper describes two decision support experiments that attempted to provide UAV operators with multivariate scheduling assistance, with mixed results. Those automated decision support tools that provided more local, as opposed to global, visual recommendations produced superior performance, suggesting that meta-information displays could saturate operators and reduce performance.

Keywords: I.6.9., Visualization, I.5.5.a, Interactive systems, Intelligent aiding

I. Introduction

Unmanned vehicles (UVs) are quickly becoming ubiquitous in almost every aspect of hostile environment operations, including air, both on and under ground and water, and even space. With reduced radar signatures, increased endurance and the removal of humans from immediate threat, unmanned aerial vehicles (UAVs) have become indispensable assets to militarized forces around the world. In addition to potential military applications, remotely guided underwater vehicles are regularly used by the oil and gas industry for maintenance purposes. Unmanned surface water vehicles are under development and testing for harbor patrol. The mining industry is increasingly looking towards remotely operated vehicles for solutions in very hostile climates. Transcending earthly bounds, unmanned ground vehicles are now exploring the surface of Mars by the remotely guided twin rovers, Spirit and Opportunity. However, despite the absence of a crew onboard any of these vehicles, human operators are still needed for supervisory control.

These UVs require human guidance to varying degrees and often through several operators. For example, the Predator requires a crew of three to be fully operational. However, with current military focus on streamlining operations and reducing staffing, there has been an increasing effort to design systems such that the current many-to-one ratio of operators to vehicles can be inverted. While this manpower inversion has received the most interest primarily in the air and ground unmanned vehicle domains, it can be seen in other domains such as mining. In order to replace multiple personnel currently required to operate a single UV with a single operator, the UVs will have to become more autonomous, and instead of lower-level tasks performed by today's UV teams, the sole operator will need to focus on high-level supervisory control tasks such as monitoring mission timelines and reacting to emergent mission events.

A key challenge in the design of these futuristic one-controlling-many systems will be the need to minimize periods of excessive operator workload that can arise when critical tasks for several UVs occur simultaneously. To a certain degree, it is possible to predict and mitigate such periods in advance. However, actions that mitigate a particular period of high workload in the short term may create long-term episodes of high workload that were previously non-existent. Thus some kind of decision support is needed that facilitates an operator's ability to evaluate different action alternatives for managing a mission schedule in real-time. To this end, this paper will present an iterative design cycle that attempts to leverage both intelligent, predictive aiding as well as human judgment and pattern recognition to maximize both human and system performance in the supervision of four UAVs.

II. The Experimental Test Bed

To study what types of decision support tools would help an operator with multiple UAV schedule management, including what kinds of intelligent aiding would be the most beneficial, a dual screen simulation test bed named the Multi-Aerial Unmanned Vehicle Experiment (MAUVE) interface was developed (Figure 1). This interface allows an operator to supervise four independent UAVs simultaneously, and intervene as the situation requires. In this simulation, users take on the role of an operator responsible for supervising four UAVs tasked with destroying a set of time-sensitive targets in a suppression of enemy air defenses (SEAD) mission. Because the simulated UAVs are highly autonomous, they only require that operators provide high level mission planning and execution actions as inputs. The operator’s job in the MAUVE simulation is to monitor each UAV’s progress, replan aspects of the mission in reaction to unexpected events, and in some cases manually execute mission critical actions such as arming and firing of payloads.

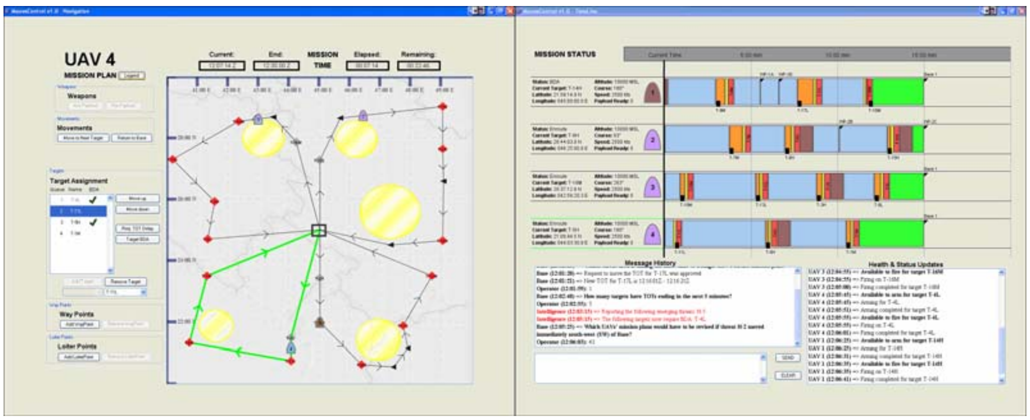


Figure 1: The MAUVE dual screen interface

The UAVs supervised by participants in MAUVE were capable of six high-level actions in the simulation: traveling enroute to targets, loitering at specific locations, arming payloads, firing payloads, performing battle damage assessment, and returning to base, generally in this order. Battle damage assessment (BDA, otherwise known as battle damage imagery or BDI) is the post-firing phase of weapons release where imagery is taken of the target. In this simulation, BDA is semi-automated in the sense that operators are responsible for scheduling BDA in advance, but the UAV performs it automatically after firing, if scheduled. Performing BDA was not required for every target, but was dependent on preplanning or in-flight contingencies.

The left-hand side of the MAUVE interface is the navigation display, and it consists of a mission time window, map display, and a mission planning and execution bar (Figure 1, left side). A large mission time box showing both time elapsed and time remaining in absolute and relative terms is located on the top right of the display. The map display represents a two-dimensional spatial layout of the battlespace, updated in real-time. Threat or hazard areas, circular in shape, have a striped yellow coloring pattern, and can be dynamic throughout scenarios, changing size, locations, disappearing entirely, or emerging as time progresses. The UAVs, always held constant at four, independently change colors according to their current action (one of the six as discussed previously).

Targets are designated by a diamond-shaped icon, and are assigned a relative priority of high (H), medium (M), or low (L). UAV routes on the map display can be changed in minor ways by selecting a particular waypoint or loiter point and dragging it to the desired location. More major routing changes such as the addition or removal of waypoints, loiter points, or targets can be accomplished using the mission planning and execution bar to the left of the map. Routing changes were only required as a result of unexpected scenarios and represents real-time replanning.

Operators are provided with a “Request TOT Delay” button which allows them limited opportunities to manipulate the time-on-targets (TOTs) for those targets assigned. Operators can request a TOT delay for a given target for two reasons: 1) According to the current mission plan, they are predicted to arrive late to that target and therefore will miss their deadline, or 2) for workload purposes, i.e., if an operator feels they need to spread out a very high workload period to manage the UAVs more effectively. However, participants were warned that this function should be used with care because moving back one target's deadline likely affects the UAV's arrival time at all subsequent targets. A change of TOT is a request, not a command, and operators' requests can be approved or denied. The probability of approval increases as a function of how far in advance of the deadline the request is sent, as would likely be the case in true military situations. When a TOT deadline is immediately approaching, the chance of approval is zero, but nearly 1.0 when requested 15 minutes in advance (participants were told this). Users can request as many TOT delays as they wished for a given target, but with no guarantee of approval.

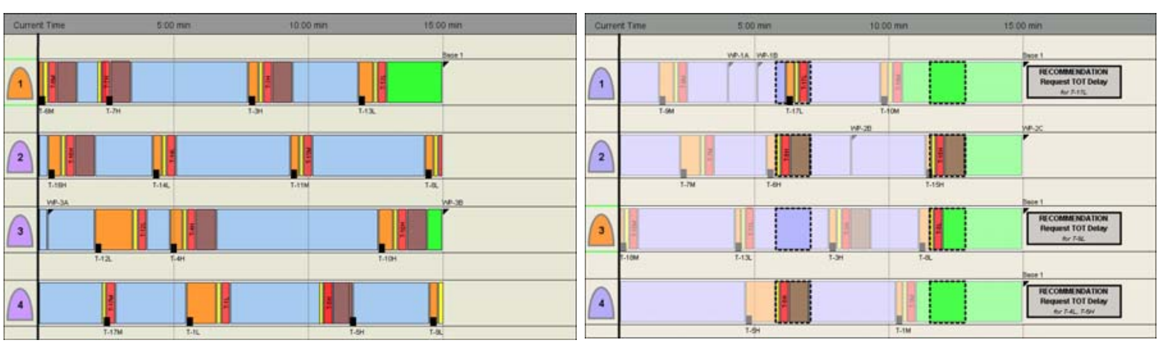
The right-hand side of the MAUVE simulation in Figure 2 provides operator decision support, and consists of a UAV status window, chat box, UAV health and status updates, and the decision support window. The status window at the top left of the decision support display gives operators low level, detailed information for each UAV such as current target, current action being performed, position in latitude and longitude, course, and weapons information. Speed and altitude are also shown in the status display, although they are not directly controllable by operators.

The bottom left of the decision support display (right side, Figure 1) has a text-based communication tool known as a chat box that contains a time history of all human communication interactions. The chat box is included because it is an established method of communications in current day military command and control scenarios, and is an embedded secondary workload tool [1]. The chat box window displays various notification messages that appear in response to scenario events or actions taken by users, as well as periodic task-relevant questions for operators to answer. One message that is particularly important to operators is notification that a TOT request is accepted or denied. The bottom right of the decision support display contains a UAV health and status notification window which separates human communications in the simulation from system communications, and only contains messages from individual UAVs.

The decision support always appears in the top right of the decision support display and the manipulation of the appearance and functionality of this window is the primary independent variable for two experiments that will be discussed in the next sections. The basic premise of the decision support is to simplify a priori mission planning information and provide a schedule of events and resource allocation for the pre-specified mission. The interface is detailed further in [2]

III. Experiment 1—Level of automation in multiple UAV management

The focus of the first experiment was to determine what level of automation would provide the best decision support to operators. While many levels of automation were studied in this experiment (from no automated support to high levels in which the UAVs decided and acted autonomously, with limited human veto authority), only two will be presented here that demonstrate potential problems with intelligent aiding (Figure 2). (For detailed experimental results, see [2].) The decision support tool in Figure 2a, termed Passive because it only presents information, assimilates pre-planned mission information and transforms it into a 15 minute horizontal timeline format, color-coded by action.



(a) Passive Aiding **(b) Active Aiding**

Figure 2: Passive vs. Active Aiding in MAUVE

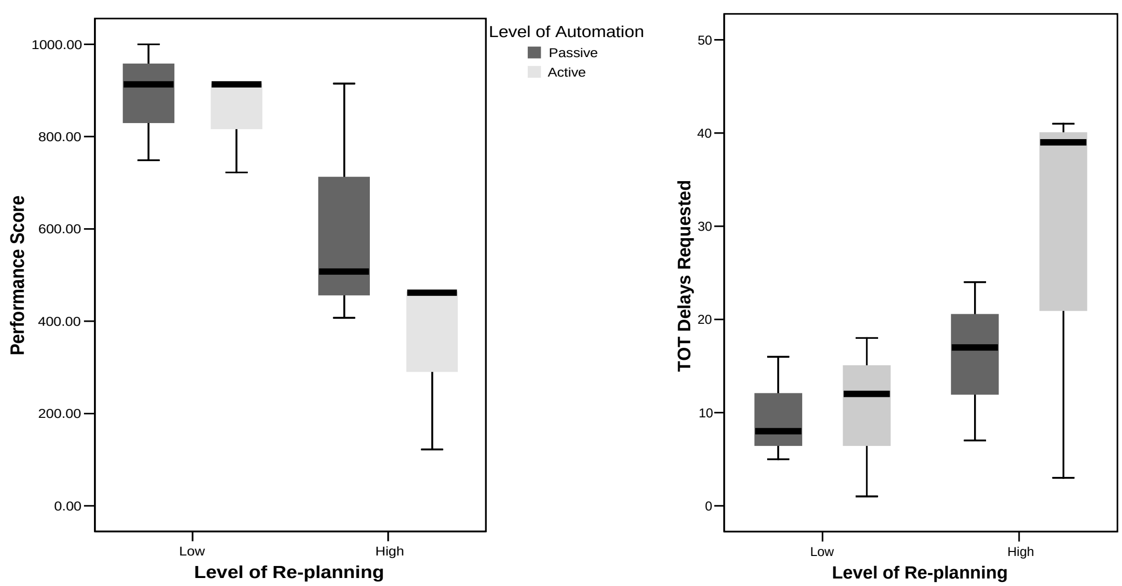
Figure 2b demonstrates the Active decision support, so named because it harnesses an algorithm which alerts the operator to possible periods of high workload and offers recommendations for resolution. The automation draws attention to these possible high workload areas through a reverse shading technique, in which potential “bottlenecks” are highlighted while the rest of the timeline’s colors are muted, but still visible. These bottlenecks are defined as predicted periods of time when two or more UAVs require human interaction. In this simulation, operator initiation of arming, firing, and BDA sequences was required, which meant that an operator had to follow a set of procedures to activate the event. The reverse shading technique was used to represent these bottlenecks since no information would be hidden, only made less salient, so that the operator’s attention could be directed to the appropriate areas of the schedule while allowing them to maintain situation awareness for the rest of the mission.

In addition to identifying areas of high workload, the computer also recommends a course of action to alleviate the high workload areas, such as moving a particular TOT. Tulga and Sheridan [3] demonstrated that even with preview in time-critical tasks, the time participants plan ahead typically decreases as workload increases, so we hypothesized that by providing both preview into the future as well automated recommendations, the need for planning time would be decreased and thus performance would improve. To this end, computer recommendations appeared in gray boxes to the right of each relevant UAV’s timeline. While the automation made locally optimal recommendations, the algorithm was not globally optimal. Following the computer’s recommendation to relieve a high workload area removed that particular schedule conflict, but sometimes created another in the process. Moreover, the algorithm’s predictions were more likely correct in the near-term (5 minutes), but more uncertain for events occurring 15 minutes into the future. See [2] for more details concerning the scheduling and recommendation algorithms.

A total of 12 participants took part in this experiment, 10 men and 2 women, and were recruited if they had UAV, military and/or pilot experience. The subject population consisted of a combination of students, both undergraduates and graduates, as well as those from the local reserve officer training corps (ROTC) and active duty military personnel. Participants had two main objectives in this experiment: 1) To guide each UAV’s actions so all UAVs under their supervision properly executed the required missions, which changed over time, and 2) To answer periodic questions from commanders.

Two independent variables were of interest in this experiment: level of decision support (Figure 2) and level of replanning. The replanning factor represented an objective workload factor, so low and high levels of schedule replanning were investigated. The low replanning condition contained 7 replanning events, while the high replanning condition contained 13. The level of decision support was a between-participants variable and the level of replanning was a within-participants repeated variable, so participants were randomly assigned to a LOA factor level but experienced both replanning conditions.

The results for the overall performance score, which measured how well participants achieved the numerous objectives for a test session (calculated as a product of the targets correctly destroyed, including



a) Overall Performance **b) TOT Delay Requests**

Figure 3: The influence of automation and re-planning on performance and delay requests.

their priority and difficulty level, and number of times BDA was correctly performed) were unexpected (Figure 3a). Under low workload, increasing automated decision support provided no additional benefit or negative consequence. While participants with either interface performed well under low workload (low replanning), particularly under high workload those subjects with the Active decision support performed significantly worse that those with the Passive (p=.027, all alpha = .05). This was unanticipated as it was originally thought that the more “intelligent” decision support would produce superior performance. Further investigation revealed that operators in the Active condition used the “Request TOT Delay” feature more than did those under the Passive Condition (Figure 3b). Use of this feature was indicative of ineffective strategies, which will be discussed below.

Situation awareness was captured through a subjective SA scale constructed from expert observer ratings. Situation awareness (SA) is generally defined as the perception of the elements in the environment within a volume of time and space, and the comprehension of their meaning and the projection of their status in the near future [4]. SA has three distinct levels, which are: 1) the perception of the elements in the environment, 2) the comprehension of the current situation, and 3) the projection of future status [4]. Subjective SA scales based upon expert observer ratings have been found to be a reliable and valid measure way to measure SA [5]. This expert rating scale was based upon threat area incursions, system wait time at targets, number of targets missed, and percentage of replanning events successfully completed. Although not statistically significant, the results showed a strong trend in which the operators with passive decision support had higher situation awareness than those with active decision support.

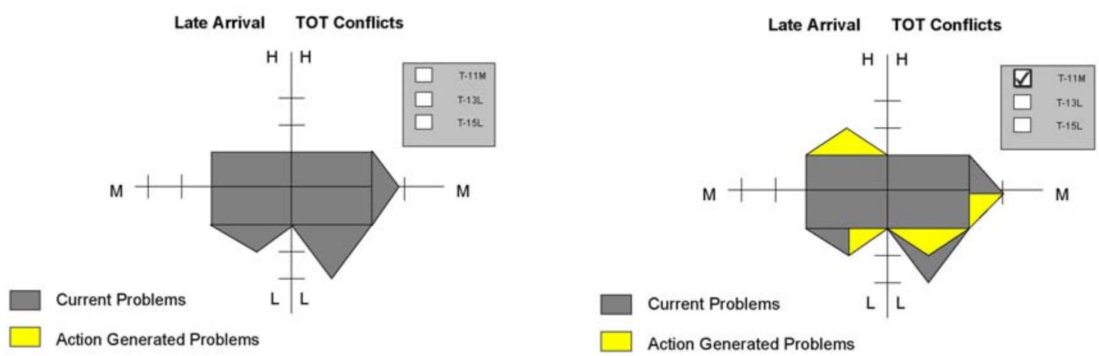
It was hypothesized that those participants in the Active condition performed poorly for two reasons. The first was the inability of the operators to generate appropriate stopping rules when trying to achieve a particular schedule change, as evidenced by their overuse of the Request TOT Delay function. Stopping rules are the criteria that individuals use to “satisfice” in uncertain situations, i.e. choosing the current best plan that is good enough [6]. While often humans can adapt effective heuristics for generating stopping rules [7], it is particularly difficult for them to do under dynamic, uncertain, and time-pressured conditions typical of command and control domains [8, 9]. To the detriment of other tasks and vehicles requiring their attention, participants often focused on obtaining a particular delay until they obtained it, often disregarding the probabilistic outcomes. In seeking to minimize TOT conflicts, participants narrowly focused on the timeline display, but on the future at the expense of the present. Because the operators focused on avoiding TOT conflicts for the entire schedule (which included events with a low probability of occurrence), they neglected more important elements of immediate UAV management and consequently performed poorly.

The second reason for poor performance was a tendency for operators to work towards eliminating TOT conflicts without considering the consequences of these changes. One possible contributor to this behavior was the reverse-shading visualization technique. The reverse shading highlighted potential high workload area. However, its salience relative to other information on the timeline may have been too high, especially in combination with the absence of more information about the uncertainty. Moreover, while the reverse shading notified operators of a possible problem, other than indicating which target should be delayed, the aid did not indicate what problems might be caused by delaying a target. Thus, the visualization made TOT conflicts very salient but did not reveal the consequence of TOT delays on subsequent TOT conflicts or target arrival times. Operators had no clear representation of how much worse the situation might be if the near-term problems were fixed. The results from this experiment motivated a redesign of the decision support, which will be discussed in the next section.

IV. The Decision Support Redesign

The first experiment shows that not only was the reverse shading too salient, but also that a significant variable was not explicitly represented, that of late target arrivals. A possible late arrival could be easily seen on the timeline (i.e., when the TOT window failed to align with the target icon), but this information was much less salient than the TOT conflicts and could easily go unnoticed. Although this was generally not a problem under low workload conditions, under high workload conditions, the more salient TOT conflicts drew attention away from more subtle display elements, such as the late arrival information. Moreover, because the salient TOT conflicts encouraged operators to request TOT delays under the Active condition, it likely initiated a prejudicial cycle in which the overly salient TOT conflicts led to the initiation of delays, which increased workload. This increased workload caused operators to rely more heavily on the salient elements of the display, which in turn led to the initiation of more delays. In this cycle that negatively impacted performance, operators were quickly saturated and did not see late arrivals in time to request a TOT delay.

As a result of these problems that surfaced in Experiment 1, two major changes were made. The first was the addition of a configural display to each UAV timeline, called the StarVis (Star Visualization) (Figure 4). A configural display is a single geometrical form that maps multiple variables onto it and changes in the individual variables cause the form to vary [10]. The variables integrated in the StarVis include the type of schedule problem (late arrival or TOT conflict), the number of targets involved in a specific problem type, and their relative priorities (low, medium, or high). Additionally, the StarVis is a projective "what if" tool, allowing operators to see the effects of requesting a TOT delay, prior to taking any action.



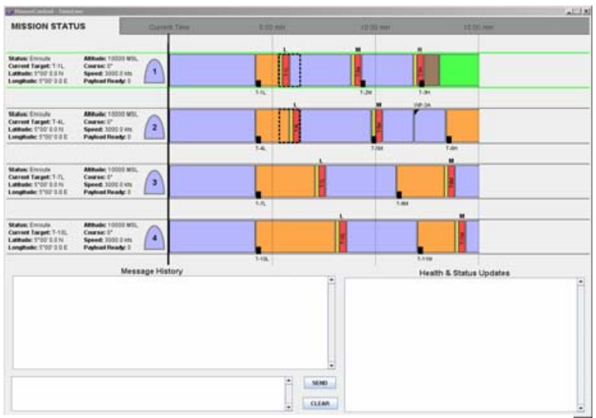
(a) Current problem **(b) Projected "what if" problem**

Figure 4: The StarVis decision support tool

The StarVis operates in two modes: current and projected problems. Figure 4a shows the default mode, which displays the current projected late target arrivals (left side) and TOT conflicts (right side) for a UAV in the next fifteen minutes. If no problem is predicted to exist in the next fifteen minutes of a UAV's schedule, the StarVis contains a gray rectangle, which represents the ideal state. As problems are detected by the automation, gray triangles appear on the StarVis for one or more target problems in one of three locations (top for high priority (H), side for medium priority (M), and bottom for low priority (L).) These triangles represent a problem either for late target arrivals (left) or TOT conflicts (right). The height of each triangle represents the number of targets involved in a particular problem. In Figure 4a, the UAV in question has a single projected late arrival for a low priority target (left) and one medium priority and two low priority time-on-target conflicts.

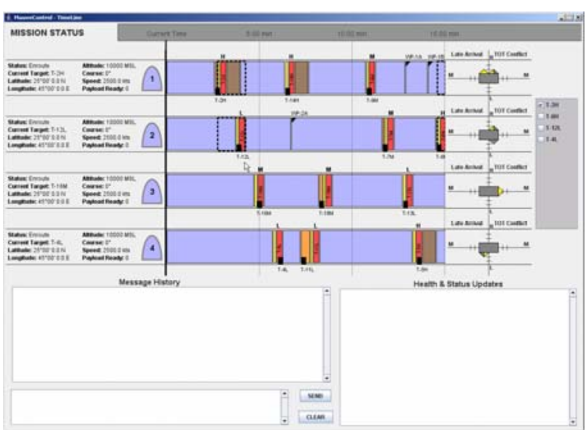
Next to the StarVis is the list of targets with these schedule problems, so they map to the triangles on the StarVis. By selecting one of the checkboxes (Figure 4b), the operator puts the StarVis into the projective "what-if" mode. By selecting a checkbox, the operator is virtually querying "if I request a TOT delay on this target, and it is granted, what will happen to this UAV's schedule?" Selecting a checkbox may cause yellow triangles to appear, which represent how the schedule is likely to change if the TOT request is granted. Split gray and yellow triangles indicate that the same problem that exists on the current timeline would continue to exist if the selected target was delayed. In Figure 4b, if a TOT delay request is made for T-11M (target 11 of medium priority), the low priority late target arrival would remain, but an additional high priority target would also be reached late. For the TOT conflicts, the medium priority target would still be in conflict, but only with one low priority target. For the case in Figure 4, requesting the TOT delay for T-11M would be inadvisable because additional higher level problems are created by this request.

The StarVis was designed to leverage direct perception-action which allows operators the ability to utilize more efficient perceptual processes rather than cognitively demanding processes that rely on memory, integration, and inference [11]. One important feature of configural displays that exploit the benefits of direct perception is the concept of an emergent feature. Emergent features are produced by the interaction between display elements, and provide a higher-level aggregate view of a system [10]. Visual representations of late arrivals or TOT conflicts "emerge" as the triangles grow from the base rectangle. In a quick glance, operators can immediately discern for not just one UAV, but for all of them, whether or not any problems exist (no triangles = no problems). The surface area provides a direct indicator as to which UAV is experiencing the most problems, and specifically what kind. Thus the StarVis provides a high-level overview through emergent features, but also provides low level details should an operator decide to focus on a particular variable of interest.



(a) Timeline Only

(b) Local StarVis



(c) Q-Global StarVis

Figure 5: Experiment #2 displays

V. Experiment 2 - Representation of local and global schedule conflicts

To assess this redesign, a second experiment was conducted. It was hypothesized that the presence of the StarVis would improve performance over a display without such a tool, as seen in Figure 5a. One important design consideration was exactly how to implement the StarVis display in terms of should it display local problems (per UAV) or should it display global issues (across all UAVs). In the local implementation, each UAV's StarVis only showed how TOT delay requests affected that UAV's schedule (called the Local StarVis, Figure 5b.) However, because TOT conflicts involve multiple UAVs' near-simultaneous TOT occurrence, it was also hypothesized that perhaps a better implementation would be a more global one that showed for one target, how a TOT change would affect all the UAVs. This implementation was called Quasi-Global (shown as Q-Global in Figure 5c), because it showed how a TOT delay request for a single UAV propagated across the other UAVs. The Q-Global StarVis, however, did not show the operator how to fix all timeline problems at once, thus was not a fully global recommendation. Figure 6 demonstrates the difference between the Local and Q-Global StarVis representations.

In the second experiment, 15 subjects took part (11 males and 4 females) and were given the same basic training, objectives, and test scenarios as those participants in Experiment 1. They also experienced the same low and high levels of replanning as in Experiment 1. The only difference was the type of visualization support, as seen in Figure 5. Because of the results of Experiment 1, the overly salient reverse shading was replaced with a simple dashed box outline for possible conflicts (Figure 5).

Figure 7 shows the results of the overall performance scores, which were calculated in the same manner as in Experiment 1. Two important trends can be seen in the data. First, the Local StarVis produced significantly better performance results under both low and high replanning than those participants with no decision support and the Q-Global StarVis (p=.004). Moreover, operators with Q-Global performed statistically *no differently* than those operators with no visualization support. Situation awareness measures were taken in exactly the same way as in Experiment 1 and they showed the same trend as the performance score, i.e., those subjects with Local StarVis had the highest SA and Q-Global subjects had lower SA (p=.014), which was not statistically different from those with no visualization (p=.730). Another significant finding in this study in terms of SA was that those subjects with the Local StarVis did not experience any significant decrease in SA across levels of replanning, i.e., they had high SA in both the high and low replanning conditions (p=.001). However, operators with the other two visualizations experienced reduced SA when operational tempo increased.

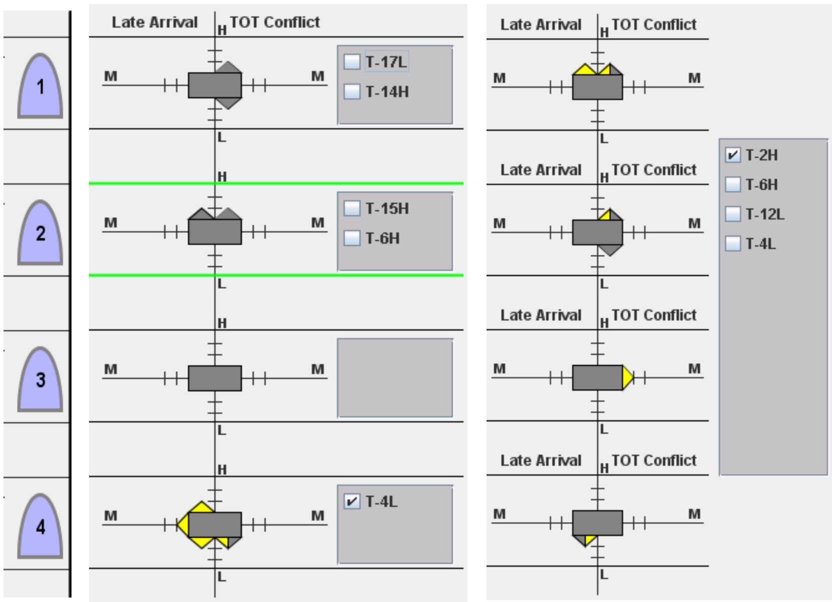


Figure 6: Local (left) vs. Quasi-Global (right) StarVis

The differences in performance scores across the visualizations (Figure 7) demonstrate that the same decision support tool, with a slightly different mapping of information, can produce dramatically different results. We hypothesize that this disparity in performance and situation awareness between the two StarVis conditions occurred because the Q-Global StarVis provided information that was not critical for the decision. Moreover, it was difficult to use in selecting a course of action, especially if the projective "what if" tool was used.

With the Q-Global StarVis, selecting a target checkbox often caused many split triangles (showing current and projective problems) and yellow triangles (showing projective problems) to appear on one or more StarVis configural displays across the UAVs. This property had negative consequences for operators. Similar to the reverse shading in Experiment 1, the Q-Global StarVis made global information salient, which was difficult for operators to understand because they had to look at all the UAV StarVis displays to understand the impact of a single decision. With the Local StarVis, however, selecting one target checkbox only affected the StarVis display of only one UAV. Thus, operators with the Local StarVis needed to only look at StarVis displays that corresponded to the checkboxes they had selected. This resulted in operators having less information to analyze in the "what if" condition. Although the Local StarVis was limited in its ability to show how a single decision would affect more than one UAV as compared to the information provided by the Q-Global StarVis across all UAVs, the projective "what if" information given in the Local StarVis was enough to help operators make effective

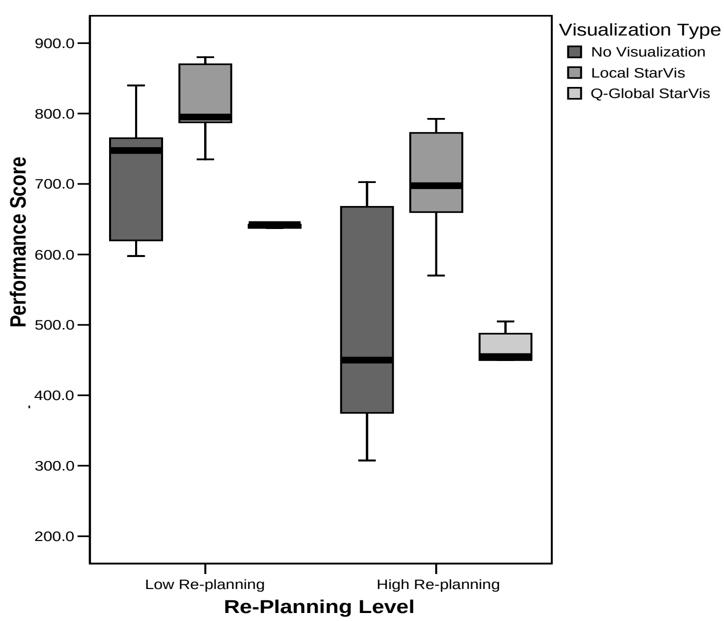


Figure 7: The effect of visualization on performance.

decisions, even though the information was not globally optimal. Thus the Local SarVis display supported a “fast and frugal” heuristic [7] which allowed operators to quickly gather just enough information to make a satisficing decision. Such decision support tools are particularly useful in dynamic, military command and control environments where time pressure and uncertainty are high. However, it is important to note that decision support tools that support such heuristics must be evaluated to assess their robustness because poor performance could result if applied to situations that violate the assumptions underlying the heuristic.

Additionally, the Local StarVis design enabled strategies not seen in the first experiment. StarVis enabled operators to assess multiple options simultaneously. This was particularly true for TOT conflicts, where users could select the projective checkboxes for the targets involved and compare the effects directly. Operators with the Q-Global StarVis tended to have more difficultly in assessing the effects on the schedule of delaying one target versus taking no action. Toggling behavior, where users selected and deselected one target checkbox multiple times, was a strategy primarily used by Q-Global StarVis subjects to try to understand the difference in the current schedule and the what-if schedule for a possible TOT delay request. This toggling behavior was not as prevalent with Local StarVis subjects, who tended to spend less time using the StarVis than Q-Global subjects (Local Toggle Mean = 8.7, Q-Global Toggle Mean = 18.2.)

Yet another strategy exhibited was the use of the TOT Delay Request option. When examining the number of TOT delay requests, the results showed that subjects with either StarVis decision support visualization requested more TOT delays under high than low re-planning. The opposite was true for subjects that had no StarVis decision support. This is likely due to the fact that those subjects without a predictive visualization did not recognize potential conflicts in their existing schedule. In comparison with the TOT delay request results from Experiment 1, operators with both StarVis implementations performed better than those subjects using the Passive display. This suggests that neither StarVis display caused the same fixation on future TOT conflicts seen with the Active level of automation in Experiment 1. Thus, StarVis configural display, regardless of implementation, helped operators balance TOT conflicts with target delays to better manage their schedule. However, while the presence of either StarVis display helped subjects *see* possible future problems, only those with the Local StarVis were able to effectively *solve* the problems.

In addition, correlations between performance scores and TOT delays ($r = -.550$) and number of late arrival mitigations ($r = .553$) demonstrated that the best performing subjects did not request many TOT delays, and focused on fixing late arrivals instead of TOT conflicts, which was a more efficient strategy. These subjects realized early on that a late arrival was a more significant schedule problem than a TOT conflict, as a late arrival guaranteed that a target would be missed. However, TOT conflicts did not guarantee that targets would be missed, only that a future period of high workload *could* lead to a missed target.

Although the results suggest a strong benefit of the Local StarVis, the results call into question some of the components of the StarVis configural display. First, because subjects favored mitigating late target arrivals over TOT conflicts, which improved performance, it may be advantageous to eliminate TOT conflicts from the StarVis configural display. More generally, these results point to the need to match the content and salience of the display elements to their task importance. In addition, these experiments demonstrate the power of a visual representation in guiding operator behavior. An alternate solution to such a visualization might be to off-load the mitigation of TOT conflicts to automation, where an algorithm could search for TOT conflicts, use a cost function to determine how severe the conflict is, alert the operator when some threshold might be exceeded, and offer a recommendation for mitigation. This will be the next focus for work within this multiple UAV supervisory control simulation.

VI. Conclusion

In the quest to design systems such that a single operator can control multiple UAVs, decision support tools will be critical, but their effect on operator performance and situation awareness cannot always be predicted. The two studies presented here demonstrate that intelligent decision support design is not straightforward and that the most well-intentioned decision aids can actually undermine performance.

In the first experiment, which investigated how levels of automation affect an operator’s schedule management strategies in the supervision of multiple independent homogeneous UAVs, a predictive intelligent aid produced unexpectedly poor results. When provided with a visualization of potentially high workload periods, as well as automated recommendations for workload mitigation, operators attempted to globally optimize their schedules, and did not adequately weigh uncertainty in their decisions. This fixation on the future, potentially driven by visualization salience, prevented participants from generating effective stopping rules and significantly degraded their performance to the point that operators without any decision support performed better than those with intelligent aiding.

The second experiment, which tested a redesign of the original decision support that included a configural display (the StarVis) that allows for perceptual-based integration of multiple variables, demonstrated that the context of this display was critical in determining its effectiveness. In the "local" context in which the StarVis only showed future problems for a single UAV, the configural display promoted the best operator performance. However, when this same display was used to represent changes across all UAVs instead of one UAV for a single TOT change, performance significantly dropped and was no different from the performance of those participants who had no decision support.

One common trend that can be seen across both experiments is the impact of increasingly "global" displays that visually represent aggregated information. In the first experiment, a decision support tool that leveraged a predictive workload algorithm in concert with a visualization across all UAVs caused operators to focus more on globally optimizing the entire schedule. This fixation led to an inability to attend to more pressing, local events and performance was significantly degraded. In the second experiment, a decision support tool that demonstrated how a single TOT change would affect all UAVs led to significantly degraded performance as compared to decision support that only displayed how a TOT change would affect a single UAV. These experiments demonstrate two important considerations in how visualizations mediate the effect of intelligent aiding. First, operators attend to visually salient representations even if the information they contain is not the most critical for the task. Second, displays must be crafted so that they not only help operators *notice* a problem and *identify* the nature of the problem, but also to *solve* the problem. Thus, future decision support designs, whether visualization or recommendation-based, should take into account the importance of the information relative to its representation salience and how operators can use the display to solve emerging problems.

Identifying prediction visualization techniques and better automated recommendation schemes for time-sensitive operations will benefit not only futuristic multiple UAV operations, but also the entire concept of networked unmanned vehicle command and control operations. To this end, this research motivates the need to develop robust scheduling decision aids that convey possible options and uncertainty, while also effectively bounding the operator such that information overload is prevented and satisficing occurs in a time-constrained environment. The goal for multiple UV decision support tool design should be to develop a tool that allows humans use of their judgment, experience, and pattern recognition strengths but also constrains them so they do not revert to biased and potentially catastrophic heuristics, as well as suffer from information overload. While configural displays and aggregated information decision support tools have been suggested to promote effective decision-making, particularly in command and control applications (e.g., [12, 13]), this research clearly shows that integrated information displays are contextual and can undermine operator performance if not designed carefully.

Acknowledgements

This research was sponsored by Boeing Phantom Work and Mitre, Inc. Special thanks to Dave Pitman, Albert Leung, and Amanda Seybold who programmed the interface.

References

[1] M. L. Cummings and S. Guerlain, "Using a Chat Interface as an Embedded Secondary Tasking Tool," presented at 2nd Annual Human Performance, Situation Awareness, and Automation Conference, Daytona Beach, 2004.

[2] P. J. Mitchell, M. L. Cummings, and T. B. Sheridan, "Management of Multiple Dynamic Human Supervisory Control Tasks," MIT Humans and Automation Laboratory, Cambridge, MA HAL2005-02, May 2005.

[3] M. K. Tulga and T. B. Sheridan, "Dynamic Decisions and Work Load in Multitask Supervisory Control," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 10, pp. 217-232, 1980.

[4] M. R. Endsley, "Design and Evaluation for Situation Awareness Enhancement," presented at Proceedings of the Human Factors Society 32nd Annual Meeting, Santa Monica, CA, 1988.

[5] H. H. Bell and W. L. Waag, "Using observer ratings to assess situational awareness in tactical air environments" in *Experimental analysis and measurement of situation awareness*, D. J. Garland and M. R. Endsley, Eds., 1995.
[6] H. A. Simon, R. Hogarth, C. R. Piott, H. Raiffa, K. A. Schelling, R. Thaier, A. Tversky, and S. Winter, "Decision Making and Problem Solving," presented at Research Briefings 1986: Report of the Research Briefing Panel on Decision Making and Problem Solving, Washington D.C., 1986.
[7] P. M. Todd and G. Gigerenzer, "Precis of simple heuristics that make us smart," *Behavioral and Brain Sciences*, vol. 23, pp. 727-780, 2000.
[8] R. S. Goonetilleke and C. G. Drury, "Human optimization with moving optima," *Ergonomics*, vol. 32, pp. 1207-1226, 1989.
[9] G. Klein, "The Fiction of Optimization," in *Bounded Rationality: The Adaptive Toolbox*, G. Gigerenzer and R. Selton, Eds. Cambridge, MA: MIT Press, 2001, pp. 103-121.
[10] K. B. Bennett, "Graphical Displays: Implications for Divided Attention, Focused Attention, and Problem Solving," *Human Factors*, vol. 34, pp. 513-533, 1992.
[11] J. J. Gibson, *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin, 1979.
[12] J. Pfautz, E. Roth, A. M. Bisantz, G. Thomas-Myers, J. Llinas, and A. Fouse, "The Role of Meta-Information in C2 Decision-Support Systems," presented at Command and Control Research and Technology Symposium., San Diego, CA, 2006.
[13] C. M. Burns, "Putting It All Together: Improving Display Integration in Ecological Displays," *Human Factors*, vol. 42, pp. 226-241, 2000.