188 *Int. J. Applied Decision Sciences, Vol. 3, No. 3, 2010*

Global vs. local decision support for multiple independent UAV schedule management

Mary L. Cummings*

Aeronautics and Astronautics Department,
Massachusetts Institute of Technology,
Cambridge, MA 02139, USA
Fax: 617-253-4196
E-mail: missyc@mit.edu
*Corresponding author

Amy S. Brzezinski

Expedition Vehicle Division,
NASA Johnson Space Center,
Houston, TX 77058, USA
E-mail: amy.s.brzezinski@nasa.gov

Abstract: As unmanned aerial vehicles (UAVs) become increasingly autonomous, time-critical and complex single-operator systems will require advance prediction and mitigation of schedule conflicts. However, actions that mitigate a current schedule conflict may create future schedule problems. Decision support is needed allowing an operator to evaluate different mission schedule management options in real-time. This paper describes two decision support visualisations for single-operator supervisory control of four independent UAVs performing a time-critical targeting mission. A configural display common to both visualisations, called StarVis, graphically depicts current schedule problems, as well as projections of potential local and global schedule problems. Results from an experiment showed that subjects using the locally optimal StarVis implementation had better performance, higher situational awareness, and no significant increase in workload over a more globally optimal implementation of StarVis. This research effort highlights how the same decision support design applied at different abstraction levels can produce different performance results.

Keywords: multiple unmanned aerial vehicles; supervisory control; configural displays; decision support; schedule; visualisation.

Reference to this paper should be made as follows: Cummings, M.L. and Brzezinski, A.S. (2010) 'Global vs. local decision support for multiple independent UAV schedule management', *Int. J. Applied Decision Sciences*, Vol. 3, No. 3, pp.188–205.

Biographical notes: Mary L. Cummings received her BS in Mathematics from the US Naval Academy in 1988, her MS in Space Systems Engineering from the Naval Postgraduate School in 1994, and her PhD in Systems Engineering from the University of Virginia in 2004. A Naval Officer from 1988–1999, she was one of the Navy's first female fighter pilots. She is currently an Associate Professor in the Aeronautics & Astronautics Department at the Massachusetts Institute of Technology. Her research interests include human interaction with

autonomous vehicle systems, modelling human interaction with complex systems, decision support design, and the ethical and social impact of technology.

Amy Brzezinski received her SB in 2005 and SM in 2008 in Aeronautical & Astronautical Engineering from the Massachusetts Institute of Technology. She is currently a NASA Flight Controller for the International Space Station in the Command and Data Handling Group. Her research interests include display design, and space and ground-control human factors.

1 Introduction

As unmanned aerial vehicle (UAV) flight and navigation tasks become more automated, UAV missions will transition from teams of people operating one UAV to one person supervising multiple UAVs (Cummings et al., 2007). Increases in UAV autonomy will alter the human operator's role to one of supervisory control, in which the operator will be primarily responsible for high-level mission management as opposed to low-level tasking and manual flight control. Because of this reduction in tasks requiring direct human control, a UAV operator may be able to supervise and divide attention across multiple UAVs.

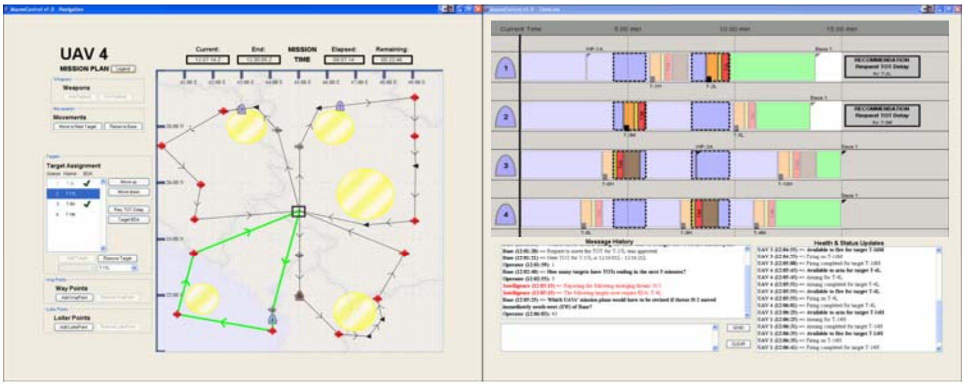
In this scenario, a critical human factors issue is one of mental workload, which is a function of attention allocation across numerous tasks and the ability to quickly and accurately switch between tasks (Crandall and Cummings, 2007). Additionally, the effect of increased automation and workload on an operator's situational awareness is an area of concern. While increased automation will be necessary to achieve the one operator-multiple UAV control paradigm, automation can increase an operator's mental workload and decrease situation awareness (SA) due to opacity, lack of feedback, and mode confusion (Billings, 1997; Parasuraman et al., 2000).

An experiment involving a multiple UAV simulation was conducted to explore how decision support tools could help an operator cope with high workload periods while supervising multiple UAVs. Specifically, the ability of configural decision support visualisations (DSVs) to help operators proactively manage their schedule and increase the overall system's performance was studied.

2 Background

In a single operator-multiple UAV mission, it is likely that more than one UAV will require the operator's attention in simultaneous mission-critical tasks, thus creating potential high workload periods. In order to investigate this and other related issues, a simulation test bed was created, called the multi-aerial unmanned vehicle experiment (MAUVE) test bed. Using MAUVE, one human operator supervises four UAVs in a time-critical targeting mission. MAUVE consists of a map and timeline-decision support display (Figure 1). In addition to the geo-spatial representation, the map display includes a UAV interaction panel allowing operators to send commands to the UAVs. The timeline-decision support display includes a colour-coded timeline representing all four UAV schedules for the next 15 minutes, as well as an instant messaging window for human-human communications and a UAV datalink window for human-UAV communication. Depending on the level of automation used in the simulation, mission management recommendations are provided to the right of the timelines.

Figure 1 The MAUVE interface with the (left) map display and the (right) decision support timeline (see online version for colours)



MAUVE presents pre-planned missions to the operator for real-time execution. Operator mission tasks during the simulation consist of arming and firing UAV payloads at scheduled times, replanning UAV paths in response to emergent threats, assigning emergent targets to the most appropriate UAV, and answering questions about the mission from an automated 'supervisor' through the instant messaging window. Targets are assigned priorities of low, medium and high. The most mission critical task is arming and firing UAV payloads at scheduled times, called times on targets (TOTs). On the timeline, each event is defined as yellow and red bars corresponding to the operator actions of arming and firing the payload, respectively. Occasionally, TOTs include post-firing target imaging tasks (also known as battle damage assessment), indicated on the timeline with a brown bar [Figure 1(right)].

Because of the time-critical, multiple task nature of the mission, it is possible that multiple UAVs could require the operator to arm and fire payloads at the same time, creating a potentially high workload period. This overlap of simultaneous targeting tasks is called a TOT conflict. In the MAUVE simulation, an operator can request a schedule change from a higher authority, called a TOT delay, in order to push a target's TOT into the future. TOT delay requests are not always granted. Through a probabilistic algorithm, MAUVE simulates a mission commander who can grant or deny TOT delay requests. The earlier an operator requests a TOT delay, the more likely the delay will be granted, so delay requests made for near-term TOTs are rarely granted. Additionally, the amount of time a target's TOT is delayed is unknown to the operator until the delay is granted. Operators are instructed that a granted TOT delay request could create other future TOT conflicts or could cause late arrivals of UAVs to their assigned targets (hereafter referred to as late arrivals).

Because management-by-consent (MBC) control of multiple unmanned vehicles has been previously shown to be effective, (i.e., operators modify and accept automation-generated plans) (Ruff et al., 2002), MBC was enabled in a version of MAUVE to alleviate operators' workload in schedule management (Cummings and Mitchell, 2006).

In this MAUVE version (Figure 1), TOT conflicts were highlighted on the timelines, and the automation recommended which TOT should be delayed. However, despite expectations that such interactive human-computer scheduling would promote overall performance, experimentally operators performed worse than those with no decision support (Cummings and Mitchell, 2006). This poor operator performance was traced to operator misuse of the TOT delay request function. The requests should have been used sparingly to mitigate high workload periods, which were the instructions during operator training. However, operators under the MBC control paradigm requested, on average, more than twice the number of TOT delays as compared to those with no automation support, causing a workload bottleneck.

The overuse of the TOT delay request capability by MBC operators indicates an inability to generate effective stopping rules when trying to achieve schedule changes. These operators focused more on globally optimising their future schedule and less on performing current mission critical tasks, which reduced their performance and SA. In addition, the MBC decision support notifications were overly salient and did not include uncertainty information, so operators likely did not understand the impact of TOT delay requests on their future mission schedules. Granted schedule changes could create other TOT conflicts later in the mission schedule, or even late arrivals of UAVs to targets. Because no information was provided about the uncertainty involved in requesting TOT delays, operators were unable to fully understand the effects of their schedule management decisions (Cummings and Mitchell, 2006).

The results from this previous research indicate that multiple UAV operators need to better understand the potential effects of schedule management decisions on both current and future schedules. We hypothesised that if operators understood these effects, they may generate better stopping rules for schedule optimisation, prompting them to request schedule changes that worked toward achieving mission objectives. Furthermore, with less schedule change requests (and hence, less workload), operators would be better able to manage the schedule in real-time, thus increasing operator performance and SA. In order to address this operator-automation scheduling interaction issue, a new decision support tool was designed to show current schedule problems and the potential effects of schedule changes. This decision support tool was embedded into a slightly redesigned MAUVE simulation for experimentation to study its effectiveness in improving operator performance, SA, and workload level, which is discussed in detail below.

3 StarVis

3.1 Redesign of the decision support timeline

The MBC decision support timeline in Figure 1 was redesigned due to its over-salient graphical portrayal of potential high workload periods by removing the reverse shading that indicates TOT conflicts, and instead shows conflicting TOTs by encircling the conflict targets involved with dashed lines. On-time and late arrivals of UAVs to targets are displayed as in the original timeline design, with small black boxes on the timeline indicating when a UAV will arrive to a specific target. The automated recommendation was replaced with a graphical decision support display to show operators the potential future effects of schedule management decisions, discussed in the next sections.

3.2 StarVis configural display

A configural display is a single geometrical form mapping multiple variables in order to integrate information in a graphical form. A hallmark feature of a configural display is that changes in the individual variables cause the overall shape to vary (Bennett and Flach, 1992). In addition, configural displays support the proximity compatibility principle (Wickens and Carswell, 1995) through integration of relevant variables. Configural displays leverage direct perception-action by allowing operators to utilise more efficient perceptual processes rather than cognitively demanding processes relying on memory, integration, and inference (Gibson, 1979). The use of direct perception-action in user display design has been shown to improve operator performance in complex tasks (Sanderson, 1989; Buttigieg and Sanderson, 1991).

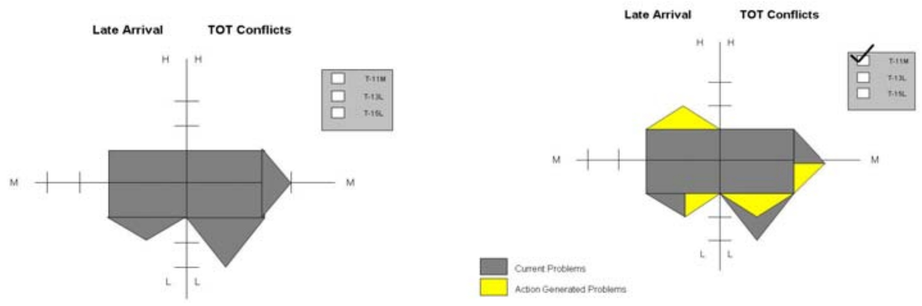
A configural display, named Star Visualization (StarVis, Figure 2) after its general shape, was designed to leverage these benefits to support multiple UAV schedule management. The StarVis display represents the number of current targets involved in two types of schedule problems, represented in the left and right halves of the display:

1 UAV late arrivals to targets (left side)

2 TOT conflicts (right side).

StarVis also includes information about the priorities of the targets (low, medium, or high) with the two schedule problems. In addition to showing current schedule problems and the targets involved, StarVis is a projective, ‘what if’ tool, allowing operators to see the effects of future schedule management decisions.

Figure 2 The StarVis display for multiple UAV schedule management, (a) current problems mode (b) projected problems mode (see online version for colours)



(a) (b)

The StarVis display operates in two modes: current and projected problems. Figure 2(a) shows the default, where StarVis indicates schedule problems that currently exist on a single UAV’s timeline for the next 15 minutes. The left side of StarVis represents targets with late arrivals, while the right side represents the UAV’s targets involved in TOT conflicts. If no problems exist in the next 15 minutes of a UAV’s schedule, the StarVis simply displays a grey rectangle. Grey triangles grow off the StarVis rectangle when schedule problems are detected by an automated search algorithm. High priority targets with a schedule problem (late arrival and/or TOT conflict) are represented by triangles that emerge from the top of the rectangle. Targets of medium and low priority are represented by triangles on the sides and bottom of the rectangle, respectively. The height of the triangles represents the number of targets of a specific priority involved in a particular schedule problem. In Figure 2(a), the StarVis shows that for a UAV's current schedule, there is one low priority target with an expected late arrival, and one medium and two low priority targets involved in separate TOT conflicts which could possible cause high operator workload.

Next to the StarVis configural display is a list of targets represented by the triangles on the StarVis. By selecting one of the checkboxes, the operator puts the StarVis into a projective mode, as shown in Figure 2(b). By selecting a checkbox, the operator virtually queries "if I request a TOT delay for this target and it is granted, what will happen to this UAV's schedule?" Selecting a checkbox potentially causes yellow triangles to appear, representing new problems that could arise if the selected target was delayed. Split grey and yellow triangles indicate that the current timeline problem could continue to exist if the selected target is delayed. Grey triangles continue to indicate current timeline problems. For the example in Figure 2(b), if a TOT delay request for the selected target is granted, the UAV could still have a low priority target late arrival and a medium priority target with a TOT conflict. These problems also exist on the current timeline. Additionally, there will be one less low priority target involved in a TOT conflict, and a new high priority target late arrival.

StarVis supports preattentive processing and facilitates direct perception of a UAV's timeline state, allowing the operator to quickly discern if a UAV has any schedule problems (indicated by the number of triangles on the rectangle and their heights) or not (indicated by a rectangle with no triangles). As a mission schedule begins to experience problems (either late arrivals and/or TOT conflicts), visual representations of these problems 'emerge' on the StarVis as triangles grow on the rectangle. By comparing the surface areas of each UAV's StarVis [which is an established way to promote preattentive processing (Ware, 2000)], an operator can quickly determine which UAV is experiencing the most problems and specifically what kind. Not only does StarVis provide a high-level overview through emergent features and preattentive processing, but it also gives low-level details when the operator wants to focus on particular variables of interest.

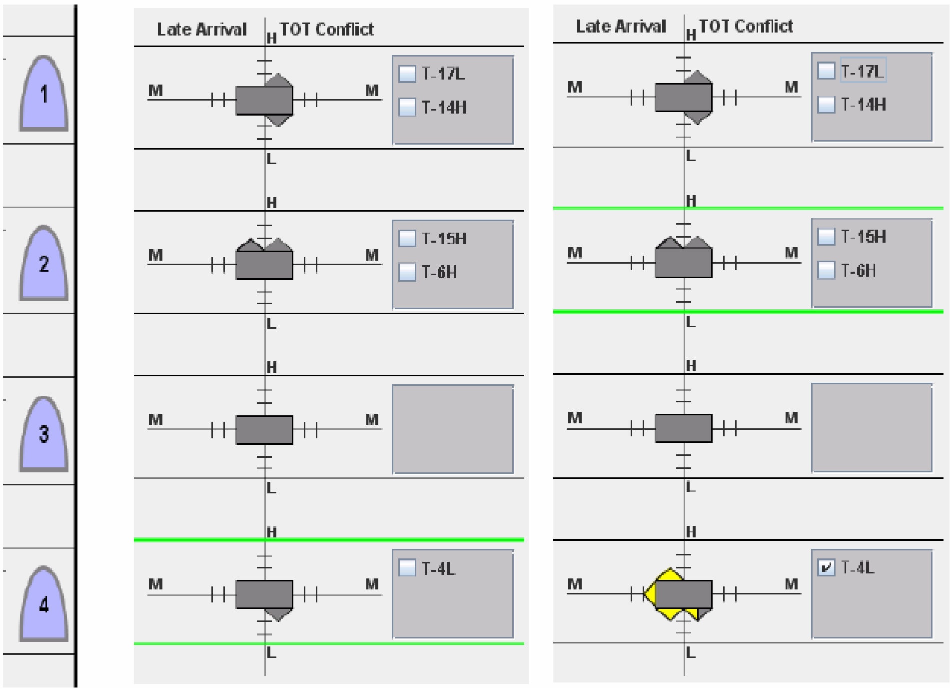
3.3 StarVis implementation

One issue that arose in pilot testing was whether the what-if mode should highlight local problems (i.e., just for the UAV in question), or whether a more global use would promote overall better system performance (i.e., a TOT delay request should be considered across all UAVs instead of a single). From a decision-theoretic perspective, it would seem obvious that the more global approach would be superior since in many cases, locally optimal solutions provide significantly suboptimal results (Winston, 1993). However, the more local approach may map more directly to a user's mental model, both in manipulating a single schedule, and also in achieving satisficing results (Simon et al., 1986), in that given the dynamic nature of command and control, a good-enough solution that can be executed quickly and understood by all parties may be preferred.

To this end, the StarVis configural display was implemented into MAUVE in two different DSVs designs: A Local StarVis DSV and a Quasi-Global (Q-Global) StarVis DSV. The two StarVis DSVs indicate current timeline problems identically; however, the designs differ in the way the projective 'what if' modes show the consequences of TOT delays, discussed in detail below.

The Local StarVis DSV implementation is shown in Figure 3. Next to each UAV's StarVis display is a list of targets with schedule problems for that UAV's current schedule, represented on the StarVis with grey triangles. In the Local mode, the operator can select only one target checkbox for one UAV's StarVis in order to activate the projective 'what if' mode, which shows the effects of delaying the selected target on only that single UAV's schedule. Thus, in the Local StarVis, yellow 'what if' triangles only appear on an individual UAV's StarVis if a target checkbox is selected. Notice that in Figure 3(b), UAV 4 is the only StarVis with a selected checkbox and there are only yellow triangles on its StarVis. Although each UAV StarVis may have only one target checkbox selected at a time, more than one StarVis can have a target checkbox selected. With multiple StarVis's in the 'what if' mode, operators may compare decision alternatives between UAVs when resolving schedule problems.

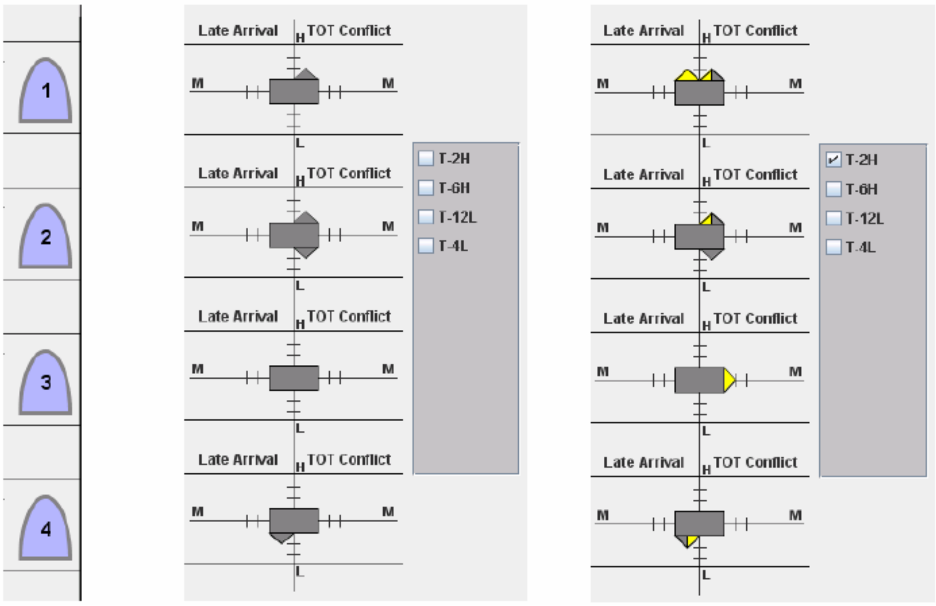
Figure 3 Local StarVis DSV (a) current schedule problems (b) current schedule problems with a 'what if' projection (see online version for colours)



(a) (b)

The Q-Global StarVis DSV implementation, shown in Figure 4, is termed 'quasi-global' because operators can explore the consequences of solutions to multiple schedule problems across all UAV schedules, but only for targets operators elect to consider. In the Q-Global StarVis, all targets with schedule problems are listed together, and when one or more checkboxes are selected, projected problems potentially appear on all UAV StarVis configural displays. Notice that in Figure 4(b), one target checkbox is selected and yellow triangles, representing projected problems, appear on multiple StarVis displays. In contrast to the Local StarVis DSV, the Q-Global StarVis design shows the effects of delaying target TOTs on the entire mission schedule, instead of only on the UAV schedule the target is assigned. However, Q-Global StarVis allows operators to layer solutions to multiple schedule management decisions, and the order of decision implementation matters. Moreover, Q-Global StarVis does not provide the operator recommendations on how to fix all schedule problems at once, which is why it is not a purely global solution. Q-.

Figure 4 Q-Global StarVis DSV (a) current schedule problems (b) current schedule problems with a 'what if' projection (see online version for colours)



(a) (b)

In summary, the Local StarVis DSV shows operators the effects of schedule changes on one UAV's timeline, while the Q-Global StarVis DSV shows operators the aggregated effects of multiple schedule changes on all UAVs. However, it is critical to remember that the two DSV designs only differ when checkboxes are selected and the 'what if' mode is engaged, and this difference is merely a functional one and not a visualisation difference.

4 Experimental assessment

An experiment was conducted to determine if and to what degree the StarVis DSVs helped operators mitigate schedule problems and improve their performance in the multi-UAV supervisory task. The experimental test bed contained the map as seen in Figure 1, but used three different schedule management DSVs: a control visualisation that consisted of only a timeline, one that combined the timeline and the Local StarVis DSV as in Figure 3, and lastly, a visualisation that utilised the timeline and the Q-Global StarVis DSV (Figure 4).

The Local and Q-Global StarVis DSV implementations each had design benefits and drawbacks. Because of the Local StarVis's ability to compare multiple alternatives of one decision, it was predicted that subjects using this visualisation would be able to better mitigate schedule problems for individual UAVs. However, because the Local StarVis only displayed the effects of a decision solution for one UAV, it was predicted that operators might not take into account effects of decision solutions across the whole mission schedule. This lack of operator consideration for cascading effects could degrade overall mission schedule management.

Because the Q-Global StarVis displays possible effects of decision solutions across all UAVs, it was hypothesised that operators using this DSV could have a better system-wide mission understanding. Additionally, the Q-Global StarVis allowed for the combination of solutions to multiple schedule problems, as it displayed the resultant effects of more than one decision across all UAV schedules. However, because of the Q-Global StarVis' inability to directly compare multiple decision solutions to one schedule management problem, it was hypothesised that operators could potentially spend excessive time optimising one decision, leading to possible decreases in performance and increased workload. Because the advantages and disadvantages of both StarVis DSV designs did not convey a clear prediction as to which implementation would be superior, the three DSVs were treated as the primary independent factor in the experiment.

4.1 Participants and procedures

The experiment, including training and testing, was performed on a four screen workstation (Osga et al., 2002). A total of 15 participants, 11 males and 4 females, took part in the experiment. The subject population consisted of students, both undergraduates and graduates, and young professionals in technical fields. All subjects were paid \$10 per hour for their participation, and a \$50 gift certificate was offered as an incentive price to the best performer in the experiment. The age of subjects ranged from 20 to 31 years, with a mean of 24 years.

A subject's main task was to supervise four UAVs in a time-critical targeting mission from start to end, which included replanning UAVs based upon emergent events. The operator's secondary objective was to answer questions about the mission status through the instant messaging tool (Figure 1, left). All subjects received between 90 and 120 minutes of training over three to four practice scenarios until they demonstrated basic competency. If subject demonstrated proficiency in training by correctly arming and firing upon a certain number of targets on time, as well as successfully completing a specific number of replanning tasks, he or she was then tested on two consecutive 30-minute mission scenarios of low and high mission replanning levels (representing low and high operational tempos). After each mission scenario, subjects completed a NASA Task Load Index (TLX) survey (Hart and Staveland, 1988), and provided feedback on a post-experiment questionnaire.

4.2 Experimental design

Operational tempo via replanning level was included in the experiment as a factor because previous research (Cummings and Mitchell, 2006) demonstrated that under low workload conditions, operator performance was relatively robust across different timeline designs. However, under high workload conditions, there was a difference in performance across the designs. A secondary goal of this experiment was to confirm this result. The low replanning scenario contained seven replanning events, while the high replanning scenario contained 13. Operational tempo was a within subjects factor while display was a between subjects factor. The order in which the subject was exposed to the factors was randomised and counter-balanced.

Dependent experimental variables included performance score, number of TOT delay requests, secondary workload, subjective workload, SA, and the net number of different schedule problems (late arrivals and TOT conflicts) mitigated by operators. As in previous work (Cummings and Mitchell, 2006), operator performance score was based upon the number of targets correctly destroyed weighted by their priority and difficulty level, as well as if certain targets were correctly assessed for battle damage. The performance score incurred penalties for incorrectly firing at erroneous targets (i.e., friendly forces), for damage taken by UAVs while traversing threat areas, and for UAVs returning to base beyond the mission time limit. This performance score also penalised operators for excessive TOT delay requests, as abuse of this capability would have tangible consequences for the individual and organisation in actual time-critical military operations.

Secondary workload measures were relevant to this experiment as the StarVis DSVs were intended to help operators mitigate potential high workload periods in mission scheduling tasks. Secondary workload was measured by the average response time to online chat questions that appeared at predetermined times in each experimental mission scenario. Previous research has shown that the use of chat question responses is an effective technique for measuring spare mental capacity, and thus workload in command and control settings (Cummings and Guerlain, 2004). In addition to reducing actual workload, adding a DSV should not increase perceived operator workload. Therefore, operator subjective workload was measured using the NASA TLX; however, since the mission task involved no physical demand, subjects were instructed to purposefully rank physical demand as a low contributor.

Four indicators within the mission were used to measure SA: UAV threat area incursion, excessive UAV loitering at missed targets, missed targets, and the percentage of emergent replanning events successfully and correctly completed. These events represented both levels 2 (comprehension) and 3 (future projection) of SA (Endsley, 1995). Different ranges of possible values for each of the SA indicators were grouped and then ranked on a 1–5 scale to generate an overall measure of SA.

Since the purpose of the StarVis DSVs was to aid operators in schedule management, specific scheduling performance was assessed by measuring the different schedule problems (late arrivals and TOT conflicts) mitigated by operators. This was calculated by subtracting the number of problems the operator encountered or created in a scenario from the number of problems the operator fixed.

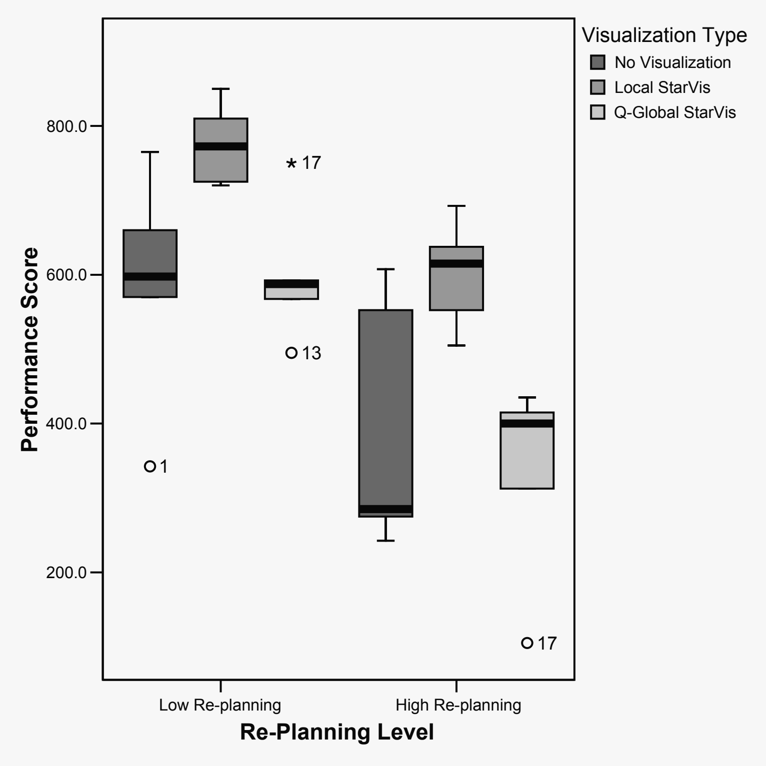
For statistical analysis, a 2 × 3 repeated measures MANOVA was used, which considered both replanning level (low vs. high) and DSV type (no visualisation, Local StarVis, and Q-Global StarVis). Pearson correlations were also found in order to determine relevant relationships between specified variables. All variables met normality and homogeneity assumptions, $\alpha = 0.05$.

5 Results

5.1 Performance score

Figure 5 shows the box plots for the performance scores for the three visualisation conditions under the two different levels of replanning. Both the level of replanning ($F(1, 12) = 22.5$, $p < 0.001$) and visualisation type ($F(2, 12) = 9.9$, $p = 0.003$) were statistically significant. There was no significant interaction between the independent variables. Subjects using the Local StarVis DSV performed better than subjects in the no visualisation and Q-Global StarVis conditions. Tukey posthoc comparisons determined that subjects with the Q-Global StarVis performed statistically no different than those subjects with no visualisation ($p = 0.903$), while those with the Local StarVis outperformed operators with the Q-Global StarVis ($p = 0.004$) and no visualisation ($p = 0.009$).

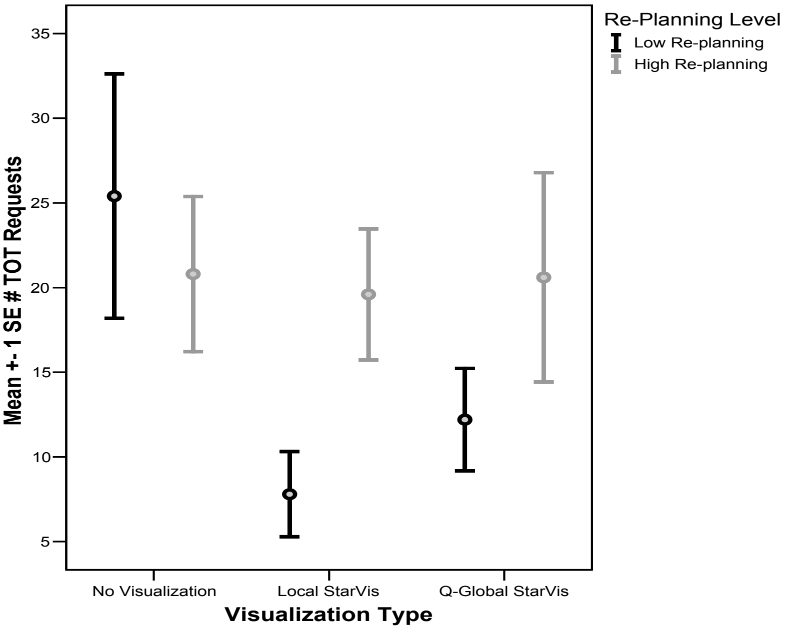
Figure 5 Performance scores



5.2 Number of TOT delay requests

Figure 6 graphically portrays the impact of the replanning and visualisation factors on TOT delay requests. While there was a trend for a low number of requests for the StarVis DSVs under low replanning conditions, there were no statistically significant differences for either replanning level or visualisation. However, previous research found that performance was lower when the number of TOT delay requests was high (Cummings and Mitchell, 2006), and an overall correlation in this experiment, controlling for the visualisation factor, confirmed this earlier finding ($r = -0.581$, $p = 0.001$). Thus, the more a subject requested TOT delays, the more likely a lower performance score resulted.

Figure 6 TOT delay requests



Given this relationship, when the 2 × 3 repeated measures model for the performance score included the number of TOT delay requests as a covariate, the results essentially remained the same in that no visualisation and Q-Global operators performed no differently than those with no decision support ($p = 0.454$), and Local StarVis performance was higher than the no visualisation and Q-Global conditions ($p = 0.039$ and $p = 0.001$ respectively).

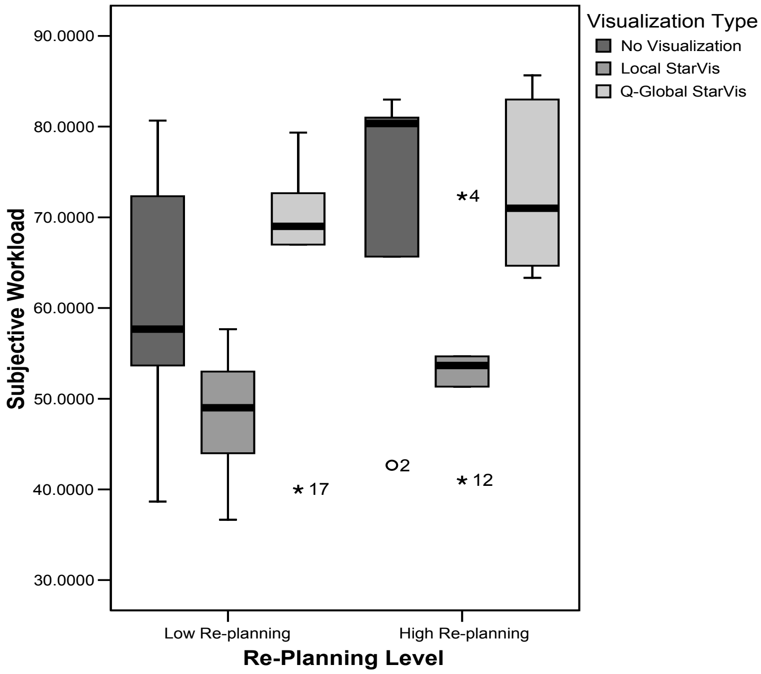
5.3 Secondary workload

Secondary workload was statistically significant across level of replanning ($F(1, 12) = 7.1$, $p = 0.02$), but was not significant across visualisation type (($F(2, 12) = 0.478$, $p = 0.631$). There was no significant interaction between the independent variables. It was expected that this metric would be statistically significant across replanning level, as the levels represented different operational tempos, which would affect operator workload. However, secondary workload did not statistically increase with the addition of any StarVis DSV, which is a promising result indicating that the StarVis configural display in either implementation did not add any additional operator workload.

5.4 Subjective workload

Subjective workload was marginally significant for both replanning ($F(1, 12) = 4.1$, $p = 0.065$) and visualisation type ($F(2, 12) = 3.8$, $p = 0.051$). There was no significant interaction. Tukey posthoc comparisons for this metric were only significant between Local and Q-Global StarVis subjects, at a marginal level ($p = 0.052$). Figure 7 shows a box plot of subjective workload scores for the different visualisations under the two replanning levels. Corresponding to the Tukey test results, operators using Local StarVis reported less perceived workload than Q-Global StarVis subjects. Statistically, the subjective workload for Local StarVis subjects was no different across low and high replanning.

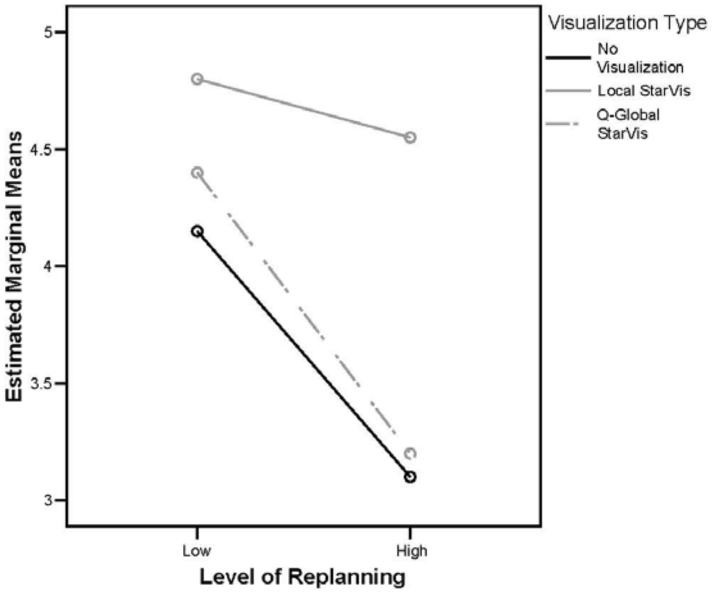
Figure 7 Subjective workload box plots



5.5 Situation awareness

Operator SA was statistically significant across levels of replanning ($F(1, 12) = 45.9$, $p < 0.001$) and visualisation type ($F(2, 12) = 9.8$, $p = 0.003$). There was a significant interaction between replanning level and visualisation type ($F(2, 1\ 2) = 5.7$, $p = 0.018$). Tukey post-hoc comparisons demonstrated that subjects SA with the Q-Global StarVis was no different from those subjects with no visualisation ($p = 0.775$), while those with the Local StarVis had superior SA as compared to operators using the Q-Global StarVis ($p = 0.013$) and no visualisation ($p = 0.004$). As shown in Figure 8, the significant interaction is due to the relatively small difference in SA under low replanning (i.e., low workload) across the three visualisations, as compared to the significantly greater SA scores for Local StarVis under the high replanning condition.

Figure 8 SA results



In addition, while there was a large drop in SA in the no visualisation and Q-Global StarVis conditions as the level of re-planning increased, the Local StarVis produced statistically no different performance across the increase in operational tempo and workload. Thus, the Local StarVis helped subjects maintain high levels of SA, even while their workload doubled.

5.6 *Schedule problem mitigation performance*

A Kruskal-Wallis test, marginally significant, showed operators in the Local StarVis condition mitigated more late target arrivals as compared to the other two other DSV conditions ($\chi2(2) = 0.102$, $\alpha = 0.1$). When controlling for the visualisation factor, late arrival mitigation was significant associated with higher performance ($r = 0.560$, $p = 0.002$), revealing that one of the ways the best performers achieved higher scores was through mitigating late arrivals.

In terms of TOT conflict mitigation, the higher number of conflicts addressed negatively correlated with performance ($r = -0.372$, $p = 0.047$). While seemingly counterintuitive, further investigation revealed this relationship was driven primarily by the performance in the no visualisation condition, where operators experienced significantly more, and thus mitigated more, TOT conflict events than either Local or Q-Global operators ($p < 0.001$). This higher rate was likely due to the fact that operators in the no visualisation condition had difficulty seeing the long-term impacts of TOT delay requests, and thus generated more problems than they fixed. However, there was no statistical difference between Local and Q-Global either for TOT conflicts generated or mitigated.

Thus, when considering both late arrivals and TOT conflict mitigation, the best performers prioritised late arrivals as the more important schedule problem, and were able to minimise the number of TOT conflicts generated through schedule changes. This optimal performance strategy was exhibited more by operators in the Local StarVis condition than in the Q-Global.

6 Discussion

The key experimental finding in this research effort was that subjects using the Local StarVis DSV achieved higher performance scores, requested fewer TOT delays, reported lower subjective workload, and had higher situational awareness than subjects with the Q-Global StarVis or no DSV. It is especially interesting that Local StarVis subjects performed better on average than the Q-Global StarVis subjects, even though both groups used minor variations of the same configural display.

It was expected that Local StarVis subjects would outperform subjects with no visualisation, but the large gap in performance and other metrics was not expected between the Local and Q-Global StarVis subjects. Across almost every metric, subjects using the Q-Global StarVis performed at the same low level as subjects with no visualisation. Thus, one configural decision support tool, applied in two slightly different contexts, contributed to either very good or poor performance. This disparity in performance between the two StarVis DSV conditions could have been due to the inability of Q-Global StarVis operators to acquire the information they needed, especially when they used the projective ‘what if’ tool.

For example, in the Q-Global StarVis DSV design, selecting a target checkbox often caused many grey-yellow split triangles (showing current and projected problems) and yellow triangles (showing projected problems) to appear on multiple StarVis configural displays. Thus, operators had difficulty quickly understanding if delaying the selected target(s) was a good decision because they had to look at potentially all the StarVis displays and synthesise the provided information. This difficulty in quickly synthesising the provided information may have contributed to the reported higher subjective workload ratings for Q-Global.

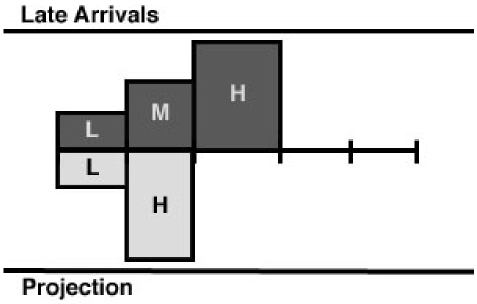
In the Local StarVis design, however, selecting one target checkbox only affected the StarVis display for one UAV. This resulted in Local StarVis operators having less information to analyse in the ‘what if’ condition. Additionally, the Local StarVis allowed operators to tailor the decision support for comparison of possible alternative solutions to one decision. Although the Local StarVis did not provide global information about the potential effects of schedule changes across all UAVs (as the Q-Global StarVis did), the projective ‘what if’ information provided by Local StarVis was enough to help operators make ‘good enough’ decisions. The Local StarVis design had less information overhead than the Q-Global design, and thus, it supported a ‘fast and frugal’ heuristic (Todd and Gigerenzer, 2000), allowing subjects to quickly gather just enough information to make a satisficing decision. Such ‘just-in-time’ decision tools are useful in dynamic military command and control environments where time pressure and uncertainties are high.

In addition, the Local StarVis DSV tended to be more intuitive to users as it allowed for the concurrent consideration of multiple decision options. This was particularly true for mitigating TOT conflicts. Operators could select StarVis checkboxes for the targets involved in a conflict and compare the effects directly on the individual StarVis displays. Subjects using Q-Global StarVis tended to have difficultly in comparing the consequences of delaying one target, versus taking no action. Toggling behaviour, where users selected and deselected one target checkbox multiple times, was a strategy used exclusively by Q-Global StarVis subjects to understand the difference between the current and 'what-if' schedule for a possible TOT delay. This toggling behaviour was costly in terms of time and cognitive capacity.

In addition, the best performing subjects, generally those in the Local StarVis condition, prioritised late arrivals and generated fewer TOT conflicts, which generally led to higher performance. These subjects understood that late arrivals were more important scheduling problems than TOT conflicts, as late arrivals preordained that targets would definitely be missed unless action was taken. However, TOT conflicts did not guarantee missed targets, only that a future period of high workload was predicted which *could* lead to missed targets. Thus, those operators in the Local StarVis condition appropriately allocated their TOT delay requests to fix late arrival problems, which was generally not the case in the other two visualisation conditions. This highlights a strength of the Local StarVis DSV, which is its ability to increase operator performance while not increasing secondary or subjective workload, even under high operational tempos.

Given that more attention to late arrivals as opposed to TOT conflicts produced superior results, and that assimilating the two types of information in StarVis could require unnecessary additional cognitive effort, a simplified DSV was proposed called BarVis (Figure 9). BarVis just shows late arrivals in a familiar bar chart format, with current late arrivals via priority on the upper half of the display, and projected later arrivals on the lower half, given selected targets. An experiment testing BarVis and StarVis, both using the local search paradigm described in detail in Brzezinski et al. (2007) was conducted, which showed that Local StarVis still produced superior results across the same set of metrics described in this paper. Thus, both current and projected late arrivals and TOT conflicts are pieces of information that are of value in multiple UAV scheduling.

Figure 9 BarVis DSV



Source: Brzezinski et al. (2007)

7 Conclusions

This research extends previous work by providing a graphical schedule management decision support tool that leverages preattentive processing and direct-perception interaction for operators supervising multiple UAVs in a time-critical targeting mission. This configural display, StarVis was designed to show problems for a current mission schedule and project potential problems that could occur if schedule changes were made. This display was implemented into a Local design, showing the effects of schedule management decisions on one UAV, and a Quasi-Global design, showing decision consequences across all UAV schedules.

An experiment utilising these two designs and a no decision support control condition showed that Local StarVis subjects performed better at multiple UAV mission management than subjects under the other conditions. Additionally, Local StarVis subjects made fewer schedule change requests and chose to manage late target arrivals over TOT conflicts, which were effective strategies. Local StarVis subjects were able to maintain high SA and understanding of mission events across varying operational tempos, suggesting the Local StarVis DSV is fairly robust to environmental changes. Lastly, the StarVis display itself did not add any additional workload overhead to subjects.

The relative high performance of the Local StarVis DSV over the other experimental conditions indicates that future decision support designs for time-critical tasks should potentially take a more local, rather than global, scope in order to promote quick, yet effective decisions. Experimental results made clear that a Quasi-Global design provided too much information that was not easily processed in a time-critical situation. Thus, future decision support designs, whether visualisation or recommendation-based, should take into account the scope and amount of information provided to human operators.

One future area of research that remains to be explored is how intelligent algorithms could be harnessed to provide explicit recommendations, perhaps truly global ones, in conjunction with the StarVis or similar configural displays. In addition, how these types of configural decision support displays interact with higher levels of vehicle autonomy, particularly for collaborative vehicles, remains an open question.

Acknowledgements

This research was sponsored by Mitre, Inc. We would also like to acknowledge the critical support of our superb undergraduate research team: Dave Pitman, Albert Leung, and Amanda Seybold.

References

Bennett, K.B. and Flach, J.M. (1992) 'Graphical displays: implications for divided attention, focused attention, and problem solving', *Human Factors*, Vol. 34, No. 5, pp.513–533.

Billings, C.E. (1997) *Aviation Automation: The Search for a Human-Centred Approach*, Lawrence Erlbaum Associates, Mahwah, NJ.

Brzezinski, A., Seybold, A. and Cummings, M.L. (2007) 'Decision support visualizations for schedule management of multiple unmanned aerial vehicles', *AIAA Infotech@Aerospace Conference*, Sonoma, CA.

Buttigieg, M.A. and Sanderson, P.M. (1991) 'Emergent features in visual display design for two types of failure detection tasks', *Human Factors*, Vol. 33, No. 6, pp.631–651.

Crandall, J.W. and Cummings, M.L. (2007) 'Attention allocation efficiency in human-UV teams', *AIAA Infotech@Aerospace Conference*, Rohnert Park, CA.

Cummings, M.L., Bruni, S. et al. (2007) 'Automation architecture for single operator-multiple UAV command and control', *The International Command and Control Journal*, Vol. 1, No. 2, pp.1–24.

Cummings, M.L. and Guerlain, S. (2004) 'Using a chat interface as an embedded secondary tasking tool', *Human Performance, Situation Awareness and Automation Technology II Conference*, Daytona Beach, FL, Lawrence Erlbaum Associates.

Cummings, M.L. and Mitchell, P.J. (2006) 'Automated scheduling decision support for supervisory control of multiple UAVs', *AIAA Journal of Aerospace Computing, Information, and Communication*, Vol. 3, No. 6, pp.294–308.

Endsley, M.R. (1995) 'Toward a theory of situation awareness in dynamic systems', *Human Factors*, Vol. 37, No. 1, pp.32–64.

Gibson, J.J. (1979) *The Ecological Approach to Visual Perception*, Houghton Mifflin, Boston.

Hart, S. and Staveland, L. (1988) 'Development of the NASA-TLX: results of empirical and theoretical research', in P.A. Hancock and N. Meshkati (Eds.): *Human Mental Workload*,. Amsterdam, North Holland, pp.139–183.

Osga, G., Van Orden, K. et al. (2002) *Design and Evaluation of Warfighter Task Support Methods in a Multi-Modal Watchstation*, SPAWAR, San Diego.

Parasuraman, R., Sheridan, T.B. et al. (2000) 'A model for types and levels of human interaction with automation', *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, Vol. 30, No. 3, pp.286–297.

Ruff, H.A., Narayanan, S. et al. (2002) 'Human interaction with levels of automation and decision-aid fidelity in the supervisory control of multiple simulated unmanned air vehicles', *Presence*, Vol. 11, No. 4, pp.335–351.

Sanderson, P. (1989) 'The human planning and scheduling role in advanced manufacturing systems: an emerging human factors domain', *Human Factors*, Vol. 31, No. 6, pp.635–666.

Simon, H.A., Hogarth, R. et al. (1986) 'Decision making and problem solving', Report of the Research Briefing Panel on Decision Making and Problem Solving, National Academy of Sciences, National Academy Press, Washington D.C. Vol. 16.

Todd, P.M. and Gigerenzer, G. (2000) 'Precis of simple heuristics that make us smart', *Behavioral and Brain Sciences*, Vol. 23, No. 5, pp.727–780.

Ware, C. (2000) *Information Visualization: Perception for Design*, Morgan Kaufmann Publishers Inc.

Wickens, C.D. and Carswell, C.M. (1995) 'The proximity compatibility principle: its psychological foundation and relevance to display design', *Human Factors*, Vol. 37, No. 3, pp.473–494.

Winston, P.H. (1993) *Artificial Intelligence*, Addison-Wesley Publishing Company.