Decision Support Visualizations for Schedule Management of Multiple Unmanned Aerial Vehicles

Amy S. Brzezinski¹, Amanda L. Seybold², and M.L. Cummings³
Massachusetts Institute of Technology, Cambridge, MA, 02139

Unmanned aerial vehicles (UAVs) are quickly becoming indispensable in military operations, particularly in time-critical missions. Although UAV systems are currently controlled by a team of people, in the future increased automation could allow one person to supervise multiple UAVs. These time-critical, complex single-operator systems will require advance prediction and mitigation of mission schedule problems. One challenge in designing an interface for the human/multi-UAV system is informing the operator of long-term consequences of potential mission schedule changes he or she may make. This paper presents two different decision support visualizations, StarVis and BarVis, designed to show the operator current mission schedule problems as well as the consequences of requesting schedule changes. An experiment tested these two visualizations against a no visualization control in a multiple UAV simulation. Results showed that StarVis produced the best performance and lowest subjective workload across different operational tempos, while BarVis supported lower but consistent performance and perceived workload under different operation tempos. This research effort highlights how different information provided in decision support can have different effects on performance and workload in a multiple UAV human supervisory control task.

I. Introduction

Unmanned vehicles (UVs) will play an increasingly important role in military and scientific endeavors in the future because they offer many benefits over traditional manned vehicles. Because there is no need to accommodate a human in the vehicle, UVs will be less complex, cheaper, and be able to carry out longer, more dangerous missions without risk of casualties. Currently the US military uses unmanned aerial vehicles (UAVs) for reconnaissance, surveillance, and attack missions, and there is also opportunity for UAV applications toward scientific endeavors, such as for land surveying and also for civil applications.

Current unmanned aerial systems (UASs) require support from a multiple personnel ground-based team, who perform flight, navigational, and high-level mission planning tasks. However, as automation technology advances, future UAVs will be more autonomous and will not have to rely on constant human supervision. With a decrease in UAV control tasks requiring human execution, current UAV systems could transition to one operator supervising a team of semi-autonomous UAVs, as opposed to the converse. The operator in a future UAS will perform high-level mission planning tasks and react to emergent, unexpected mission events.

A major challenge in this complex single-human/multiple-UAS is how to prevent high operator workload caused by the need to attend to multiple UAVs at once. In such situations, because of divided attention, the operator may forget or accidentally miss performing other time- and mission-critical tasks. Increased workload, combined with increased system automation, could decrease an operator’s awareness of the mission situation due to opacity, lack of feedback, and mode confusion. ¹, ² High workload periods can be predicted and prevented in advance, but at the cost of changing the schedule. Schedule changes could create additional high workload periods and other problems in the future. Thus, there is a need for decision support visualizations to inform the operator of the long-term effects of short-term schedule management choices. This paper presents two different decision support visualization designs which, through configural displays, represent different current and future schedule problems. Using a multiple UAV

¹ Graduate Research Assistant, Department of Aeronautics & Astronautics, 77 Massachusetts Ave. Room 35-220, Graduate Student Member
² Undergraduate Researcher, Department of Electrical Engineering & Computer Science, 77 Massachusetts Ave. Room 35-220, Undergraduate Student Member
³ Assistant Professor, Department of Aeronautics & Astronautics, 77 Massachusetts Ave. Room 33-305, Senior Member.

American Institute of Aeronautics and Astronautics
simulation called the Multi-Aerial Unmanned Vehicle Experiment (MAUVE), the two visualization designs were tested to determine if operators could better manage their schedule in a time-critical targeting mission, thus improving their task performance and reducing their perceived workload.

II. The Multi-Aerial Unmanned Vehicle Experiment (MAUVE)

The MAUVE simulation was designed to investigate the human supervisory control challenges presented by one operator controlling multiple semi-autonomous UAVs. In MAUVE, an operator simultaneously oversees 4 UAVs as they execute a pre-planned mission involving tightly scheduled targeting tasks. The operator’s supervision tasks include arming and firing UAV payloads, and occasional intervening when the mission situation changes due to emergent and unexpected events. The operator’s mission objectives are to destroy as many targets as possible, to comply with changing mission requirements, to avoid enemy fire from threat areas; to return all UAVs to base within the mission time limit, and to answer occasional questions from simulated superiors about the mission via a chat interface.

Figure 1 depicts the two interfaces used in MAUVE. The map display (Fig.1a) contains a mission clock, a geospatial map, and mission planning and execution functions. The map shows the area where the mission is occurring, and each UAV, represented by a bullet-shaped icon, travels on its own pre-planned path, starting and ending at a base (the center black rectangle in Fig 1a). Red diamonds and black triangles on the map display represent targets and waypoints, respectively. Yellow circles on the maps are threat areas where UAVs could potentially receive damage from enemy fire. As a UAV travels along its path, its icon changes color to represent its current action.

The Map Display’s mission planning and execution bar contains operator controls for arming and firing UAV payloads, and also controls for re-planning the UAV paths and schedules. Paths can be altered by adding and removing targets and waypoints. Waypoints can also be moved on the map display through dragging. The mission planning and execution bar also contains the Time on Target (TOT) delay request button, an important schedule management tool in the MAUVE interface.

A TOT is defined as a pre-designated target’s arming and firing time window. Some targets will be imaged after weapons release, so a TOT includes a battle damage assessment (BDA) time window as well. TOT windows are short in duration relative to the amount of time it takes UAVs to travel between targets. The Request TOT Delay button allows an operator to request that a TOT be delayed. An operator might do this for two reasons: first, when the current mission plan predicts a UAV will arrive late to the target (after its TOT window has passed), and thus will be unable to destroy it within the prescribed time window; second, when the mission schedule predicts a potential high workload period, defined as when multiple UAVs have simultaneous TOTs, which could cause the operator to not have enough time to execute all the arming and firing steps for multiple vehicles. Thus, by spreading out the TOTs across the mission schedule, the operator could reduce his workload. TOT delay requests are not automatically approved; the probability of request approval decreases the closer the TOT is to the present. Thus, the probability of approval for a TOT delay request is almost 1 if the request is made 15 minutes in advance, and

![Figure 1. The MAUVE dual screen interface.](image-url)
decreases to nearly 0 if the TOT is imminent. In the MAUVE simulation, an operator cannot request a TOT to be moved to a specific place in the schedule; if a TOT delay request is granted, the TOT is moved into the future as determined by a simulated mission commander who updates the operator’s schedule automatically. The operator has no way of knowing how far into the future a TOT could be moved until a TOT delay request is granted.

Figure 1b shows the MAUVE timeline-decision support display, which contains UAV status, a graphical timeline for each UAV, a chat interface for military communications, and status updates from the UAVs. To the right of the graphical timelines is space for a schedule management decision support visualization (DSV), which will be expanded upon later in the paper. The UAV status information shows a colored bullet-shaped icon representing the UAV and its current action, as well as text data about the UAV’s current target and position. The graphical timeline shows each of the UAV’s schedules, with each color representing a different UAV action. The current time is indicated to the far left of the timeline. The timeline also represents TOTs, which are indicated by yellow arming and red firing windows and are labeled with the target’s designation (a number) and priority (low, medium, and high), such as the target T-4H for target 4 with high priority. The UAV time of arrival to each target is shown on the timeline by a small black box labeled with the target’s designation.

The chat interface contains a time-stamped military communications history, including approvals and denials for TOT delays, and mission re-planning orders, such as the removal of a target from the mission plan or the emergence of a new threat area. The chat interface is also used for questioning operators about the mission to experimentally measure their secondary workload. The UAV status update box informs the operator when the UAV is available to arm and fire, which requires a UAV to be at a target during the arming and firing time window, and when these actions are completed. This box also informs operators when a UAV is being fired upon because it is traversing a threat area.

III. Schedule Management Decision Support Visualizations

As previously discussed, two types of schedule problems could occur in a multiple UAV time-critical targeting mission: first, when a UAV was scheduled to arrive at a target after its TOT scheduled window had passed, making it too late for the UAV to arm and fire upon it (hereafter called a late arrival); second, when the TOTs of two or more targets in different UAVs’ schedules occurred at the same time (hereafter called a TOT conflict). Operators could try to fix these problems by requesting a TOT delay. However, because operators did not know how far into the future a TOT would move if the delay was granted, operators were unable to predict if a schedule change would actually fix the schedule problem and/or if it would create more problems farther into the future schedule. Thus, some sort decision support was needed to help operators understand the potential effects of requesting TOT delays to fix schedule problems seen on their current timeline. The purpose of the decision support visualizations was to inform operators about current schedule problems, and to predict consequences of TOT delay requests for problem targets so as to assist operators in making decisions on how to manage their mission schedule. One visualization, StarVis, was designed to provide information about both types of schedule problems. The second visualization, BarVis, was constructed to inform operators only about late target arrivals. These will be discussed in more detail in the following sections.

A. StarVis

Figure 2 shows the first decision support visualization (DSV) design, termed StarVis, because triangles representing timeline problems grow from the center gray box, forming a star-like shape. StarVis was designed in the form of a configural display, which is a single geometrical shape that maps multiple variables onto its form. Changes in the individual variables cause the shape to vary, providing graphically dynamic information about how the system and its individual variables are changing. Configural displays support the proximity compatibility principle through the integration of variables needed by operators for comparison and computation for decision-making. StarVis, as a configural display, allows operators to exploit direct perception-action, freeing operators of cognitive workload by utilizing more efficient
perceptual processes instead of relying on the more cognitively demanding processes involving memory, integration, and inference. A configural display’s emergent features support direct perception-action by providing a higher-level aggregate view of a system state. User display designs using direct perception-action have shown improved operator performance in complex tasks.

The StarVis DSV was designed to provide operators information about both types of schedule problems, late arrivals and TOT conflicts. In the StarVis DSV, when there are no schedule problems, only the gray rectangle appears on the display. When there are late arrivals or TOT conflicts on a UAV’s current schedule, however, gray triangles emerge from the rectangle. Late arrival triangles grow from the rectangle’s left side, while TOT conflict triangles grow on the right side. The priority of the target involved in a specific schedule problem is indicated by its position on the rectangle: high priority targets have triangles on top, medium priority targets grow from the side, and low priority targets are represented by triangles on the bottom. The height of a triangle indicates the number of targets of a specific priority involved in a specific problem.

As Fig. 3 shows, each UAV has its own StarVis with a collection of checkboxes which correspond to that UAV’s targets involved in current schedule problems, as represented on the StarVis with gray triangles. When an operator checks one of these checkboxes, the StarVis will simultaneously display the gray triangles (current schedule problems), as well as yellow triangles representing the potential schedule problems that could occur if the selected target were granted a TOT delay. Based on a predictive algorithm, these yellow triangles indicate probable future schedule problems if the selected target is delayed. Because of the predictive nature of the algorithm, the yellow triangles were associated with some degree of uncertainty, meaning that just because the visualization showed a yellow triangle, it did not mean the related late target arrival or TOT conflict was guaranteed to occur. Split gray/yellow triangles indicate that a particular current schedule problem would still exist if the selected target were delayed. Selecting a checkbox for a UAV only causes yellow “what if” triangles to appear on that UAV’s StarVis. Multiple StarVis displays (one per UAV) can each have a checkbox selected, but only one checkbox can be selected for each individual StarVis.

It was believed that StarVis would help operators in schedule management decisions by visually calling to their attention late arrivals and TOT conflicts through the gray triangles, and then allowing them to test their potential decisions on how to fix those problems by visually showing them the possible effects of different decisions on a UAV’s schedule. Although a granted TOT delay request could fix a schedule problem for a UAV, it could create future problems by causing a UAV to be late to a different target later in the schedule, or by causing the delayed target’s TOT to conflict with a target belonging to another UAV. Thus, it was predicted that the StarVis DSV would better support operator schedule management decision-making over operators with no visualization, increasing StarVis operators’ performance, decreasing their workload, and increasing their situation awareness of the overall mission.

B. BarVis

Experiments with StarVis found that subjects often found the StarVis confusing, and that mitigating TOT conflicts in the time-critical targeting mission was not as critical as dealing with late target arrivals. BarVis, named for its resemblance to a bar graph and shown in Fig. 4, was designed to only display information about late arrival problems in UAV schedules in a simple manner, thus decreasing the complexity found in StarVis for schedule management. If no problems exist on a UAV’s timeline, BarVis contains the words “No late arrivals,” as shown in Fig. 5. Late arrivals on the current mission schedule are represented by

Figure 3. The StarVis for each UAV timeline. Note the checkboxes for each problem target, and the effect of checking one.

Figure 4. BarVis decision support visualization. Late arrivals in the current timeline are indicated by the darker bars above the center line; late arrivals in the projected timeline are indicated by the lighter gray bars below the middle line.
dark gray bars on the upper half of the display, with target priority indicated by bar height and number of targets indicated by bar width.

Similar to StarVis, BarVis shows the potential results of delaying a currently late target’s TOT. As seen in Fig. 5, each BarVis has a collection of checkboxes, one for each late arrival. Selecting a checkbox will show the potential late arrivals that could occur (with light gray bars) in the bottom half of the display if the selected target’s TOT is delayed. As with StarVis, selecting a checkbox for a UAV only causes the light gray bars to appear on that UAV’s BarVis. Multiple BarVis displays can each have a checkbox selected, but only one checkbox can be selected for each individual BarVis.

It was believed that due to the simpler, more straightforward nature of the BarVis DSV, operators using this visualization would perform better than operators using StarVis. Because BarVis encapsulated the same configural design principles as StarVis, as well as similar usage as a schedule management tool, it was expected that both StarVis and BarVis would facilitate higher performance, lower workload, and higher situation awareness from operators than operators with no decision support.

IV. Experiment

An experiment using the MAUVE simulation was conducted to determine if the StarVis and/or BarVis DSVs helped operators mitigate schedule problems, improve their performance, and reduce their subjective workload.

A. Apparatus, Participants, and Procedure

The experiment, including training and testing, was performed on a four screen workstation. Experimental subjects interacted with the MAUVE simulation through a standard mouse and numerical keypad. A total of 15 participants, 11 males and 4 females, took part in the experiment. Subjects consisted of students, both undergraduate and graduates, and young professionals in technical fields. All subjects were compensated for their participation either through cash or gift certificates. The age of subjects ranged from 18 to 31 years, with a mean of 22.2 years.

The subject’s task was to supervise four UAVs in the MAUVE time-critical targeting mission. Specifically, subjects’ primary objective was to guide each UAV’s actions so that all UAVs correctly executed dynamic mission requirements. Each subject’s secondary objective was to answer questions about the mission through the mission communications chat interface. The experiment format consisted of training, experimental scenarios, and post-experiment feedback. All subjects received between 90 and 120 minutes of training through three to four practice scenarios until they demonstrated competency in using MAUVE and understood mission objectives.

Post-training, subjects were tested on two consecutive 30 minute mission scenarios, one each of low and high mission re-planning levels. Each scenario represented a pre-planned mission; the low re-planning scenario contained 7 re-planning events, while the high re-planning scenario contained 13. Subjects experienced the same re-planning scenarios, but the order in which they saw them was randomized and counter-balanced. Visualization assignment was randomized for no visualization and StarVis subjects. BarVis subjects, however, were tested in a block in an experiment conducted after the no visualization-StarVis study. Each subject was provided with one type of schedule management DSV throughout the entire experiment. After each mission scenario, subjects completed a NASA Task Load Index (TLX) workload measurement survey. After completion of all experimental scenarios, subjects provided feedback about the interfaces on a post-experiment questionnaire.

B. Experimental Design

The primary independent variable of interest in this experiment was the schedule management DSV type. DSV was a between-subjects variable which had three possible assignments: StarVis, BarVis, and no DSV (as a control). The secondary independent variable, level of re-planning, represented different operational tempos of re-planning events, both in the number and their placement within a given mission scenario. Level of re-planning was a within-subjects variable, with all subjects experiencing the two re-planning levels through two different testing scenarios.

Performance and subjective workload were the two primary dependent variables studied in this experiment. Operator performance was based upon the number of targets correctly destroyed weighted by their priority and
difficulty level, as well as if certain targets were assessed for battle damage. Points were deducted from the performance score for incorrectly firing at targets specified not to be destroyed, for damage taken by UAVs while crossing threat areas, for UAVs returning to base beyond the mission time limit, and for requesting TOT delays. This latter penalty was included in performance score because a previous experiment demonstrated that abuse of TOT delay requests could have consequences for an individual and an overall mission. Subjective workload was measured using the NASA TLX subjective workload rating survey. The survey computed a workload score from operator-weighted ratings on a 1-20 scale across the dimensions of mental demand, physical demand, temporal demand, effort, performance, and frustration. Because the experiment did not involve physical demand, subjects were instructed to purposefully rank physical demand as a low workload contributor and to ignore survey portions asking about that dimension.

V. Results

A 2x3 repeated measures MANOVA was used for statistical analysis of the experimental data, which used re-planning level (low vs. high) as the within-subjects variable, and visualization type (no visualization, StarVis, and BarVis) as the between subjects variable. For all reported results, \( \alpha = .05 \). Five subjects were nested within each of the visualization types. While the performance score met normality and homogeneity assumptions, subjective workload scores were log-transformed to satisfy these assumptions. There was no significant interaction between the factors for either of the dependent variables.

Figure 6 is an estimated marginal means plot for performance score. The lines represent the visualization types, while the x-axis represents the re-planning level. For performance score, level of re-planning was significant \( (F(1,12)=5.038, p=.044) \) as was visualization type \( (F(1,12)=7.764, p=.007) \).

As Fig. 6 shows, there was a difference in the estimated means between the StarVis and the other two experimental conditions (no visualization and BarVis). Tukey post-hoc pairwise comparisons found that performance was statistically significant between the no visualization and StarVis conditions \( (p=.008) \) and the StarVis and BarVis conditions \( (p=.026) \), but not significant between the no visualization and BarVis conditions. StarVis subjects, on average, achieved the highest performance across both re-planning levels when compared to subjects using no visualization or BarVis. Tukey tests demonstrated that those subjects with BarVis performed no differently than those with no visualization, averaged across both re-planning levels. However, there is a trend in Fig. 6 that demonstrates consistent performance for BarVis across the different replanning operational tempos. A point comparison found that BarVis subjects statistically performed the same across re-planning levels \( (p=.74) \). However, BarVis

---

Figure 6. Estimated marginal means plot for performance score

Figure 7. Estimated marginal means plot for subjective workload
performance was significantly lower than the StarVis performance. StarVis and no visualization subjects had lower performance scores in their respective high re-planning scenarios than in their low re-planning scenarios, which was expected as a result of the increasing workload. So while StarVis subjects had the best overall performance across the two re-planning levels, there is a trend towards consistent, but lower, performance across the different levels of re-planning for BarVis subjects.

Experimental results showed that for subjective workload, visualization type was only marginally significant (F(1,12)=2.932, p=.092), and level of re-planning was not significant. The estimated marginal means plot for subjective workload is shown in Fig. 7. A point comparison between the StarVis and BarVis subjects under only the low re-planning condition found a significant difference in subjective workload (p=.025). Under only the high re-planning scenario, a point comparison between no visualization and StarVis showed a significant difference in subjective workload (p=.035). Thus in the high replanning scenario, StarVis and BarVis performed no differently.

The re-planning factor essentially represented objectively increasing workload, so the lack of statistical difference in subjective workload between the two re-planning levels across all visualization conditions was unexpected. Thus, subjects using a specific visualization on average reported the same subjective workload for each re-planning level.

VI. Discussion

The significance of re-planning level on performance score in this experiment was expected, as previous research showed the quantity and tempo of re-planning events significantly influences command and control mission complexity10. While the significance of visualization type on performance score was also expected, as it was hypothesized that a decision support visualization would improve operator performance, the equivalence in overall performance between BarVis and no visualization subjects, both inferior to StarVis, was unexpected. StarVis subjects, on average, outperformed both BarVis and no visualization across both operational tempos. Additionally, StarVis subjects experienced lower subjective workload than BarVis subjects under low re-planning, and lower subjective workload than no visualization subjects under high re-planning. These specific results are not surprising as lower subjective workload can often result in higher performance, two characteristics which StarVis exemplified. However, what is surprising is that BarVis, a perceptually and cognitively less demanding decision support tool, did not perform at least as well as StarVis, and even more surprisingly was its equivalence to the no visualization conditions in performance and subjective workload.

Curiously, although StarVis provided more information than BarVis since it represented both schedule problem types, StarVis subjects experienced lower perceived workload under the low re-planning condition over subjects using the simpler BarVis. However, this difference disappeared under high workload. StarVis was designed to promote preferential selection (one half represented late target arrivals and one half represented TOT conflicts), thus operators could easily ignore that information not deemed salient. These results indicate that providing information on mitigating both late target arrivals and TOT conflicts did not increase the subjective workload of StarVis subjects over subjects provided with less information in BarVis, under the low re-planning condition. It was originally hypothesized that BarVis, a display easier to understand than StarVis, would be more helpful to operators in managing their schedule, thus increasing their performance and decreasing their perceived workload. However, this was not the case – statistically, BarVis had the same performance and workload as no visualization subjects, averaged across re-planning. However, it should be noted that these results should be carefully interpreted, as they may have been influenced by the experimental confounds of StarVis and BarVis being tested in different experiments, although under the exact same protocol, as well as by the small sample sizes. Ideally, in the future more subjects will be used in a single experiment testing these visualization conditions to verify the results presented here.

One of the big questions arising from these results is why the more complex StarVis display facilitated superior performance and marginally lower workload over the other visualization conditions. A potential reason could relate to the visual presentation of StarVis itself; perhaps because the display was more graphically novel and complex as compared to BarVis, operators may have spent more time examining StarVis and understanding the information it provided, thus causing them to carefully consider their decisions. By spending more time trying to figure out StarVis, operators may have been more careful in using the information StarVis provided. Another potential reason as to why StarVis had the best performance may be due to the additional information on TOT conflicts, a schedule problem not featured in BarVis. Although a previous experiment with StarVis appeared to show that TOT conflict mitigation was not important to subjects who achieved high performance in the multi-UAV supervision task11, perhaps the presence of this information on StarVis gave subjects an awareness of potential high workload areas. This awareness could have aided them in achieving mission objectives, even though the information did not prompt
them to actively make changes to their schedules. Thus, the TOT conflict information on the StarVis visualization may have served as a situation awareness tool, rather than a decision support tool.

While BarVis did not meet performance and subjective workload expectations in this experiment, one interesting BarVis trend not seen with the other visualization conditions was relative consistency for performance the different re-planning levels. Although BarVis subjects, on average, did not achieve the highest performance scores, this consistency across very different operational tempos is intriguing since ideally, the best decision support system would allow operators to perform at consistent levels across a wide variety of operational conditions. More research is needed to determine what aspects of BarVis promoted this consistent trend such that perhaps they can be leveraged in the StarVis design, so as to facilitate high and consistent performance.

VII. Conclusion

This research effort highlights how different amounts of information encoded into graphical decision support can have different and surprising effects on human performance and workload in tasks involving multi-UAV schedule management. In a multiple UAV simulation, use of the graphical StarVis decision support visualization produced superior performance over a no visualization condition as well as a simpler graphical tool, BarVis. While StarVis produced the best performance and marginally lower perceived workload across different re-planning levels, BarVis usage resulted in consistent, although lower, performance across the different scenarios. However, across both re-planning levels, BarVis’ performance and workload was statistically no different from the no visualization condition, a surprising result since it was hypothesized that this simpler tool should perform as well as StarVis.

Future work stemming from this experiment needs to examine why StarVis, a complex display visualizing multiple variables, facilitated better performance over BarVis, considered more straightforward and easier to use. Because these results flow from two different experiments performed months apart, although under the exact same protocol and conditions, another experiment is needed, but with larger sample sizes under each of the visualization conditions, so as to verify the results presented here. Future research stemming from this work will examine how the different display designs support the operator information processing loop to determine whether or not the difference in performance is due primarily to the perceptual differences in the decision support display, or because of the difference in information content and reasoning support. From this research, it may be possible to design a new visualization that supports consistent, yet high performance across different operational tempos and conditions in the multi-UAV supervision task.

Acknowledgments

This research was partially sponsored by the MITRE Corporation. We would like to acknowledge Dave Pitman and Albert Leung for programming MAUVE and its interfaces.

References