

COMPARING QUANTITATIVE MODEL PREDICTIONS TO EXPERIMENTAL DATA IN MULTIPLE-UAV FLIGHT CONTROL

Stephen R. Dixon, Christopher D. Wickens, & Dervon Chang
University of Illinois Aviation Human Factors Division
Savoy, Illinois

Thirty-six licensed pilots from the University of Illinois Aviation Institute performed simulated military surveillance missions with one and/or two unmanned aerial vehicles (UAV). Pilots were responsible for navigating each UAV through a series of mission legs in one of the following conditions: 1) a baseline condition with all manual flight controls and visual displays; 2) an auditory offload condition that provided auto-alerts and other relevant information to the auditory channel; and 3) an automation condition that provided auto-pilot control of the UAV. Pilots were responsible for mission completion, target search, and systems monitoring. Results indicate that the two offloads are beneficial in reducing task interference and overall workload. Three theories, with corresponding workload models, were discussed in order to evaluate predicted pilot performance. Single channel theory was able to explain some of results in the baseline condition, while single resource theory and multiple resource theory were better able to explain reduced task interference in the automation and auditory conditions.

INTRODUCTION

The Army employs a number of UAVs (e.g. Hunter and Shadow) which are used to fulfill routine reconnaissance missions without placing pilots at risk (Barnes, Ghiradelli, Stachowiak, Hill, & Dahn, 2002). Currently, military command has designated two full-time operators per UAV during in-flight missions: 1) an AVO (aviator operator) who is responsible for aviation and navigation; and 2) an MPO (mission payload operator) who searches for possible targets while monitoring critical system parameters. By reducing the pilot to UAV ratio, the Army can increase the number of operational UAVs without a corresponding increase in the number of pilots. The current research intends to address issues of task interference and workload which arise from multiplying pilot responsibilities.

Workload refers to the relationship between resource supply and task demand (Sarno & Wickens (1995). Performance remains constant provided that the supply of mental resources exceeds the amount required to perform certain tasks; however, when competing tasks demand more resources than what is currently available, performance is expected to suffer. Intuitively, merging the responsibilities of multiple UAV control into one operator will result in increased task interference and workload.

The current study examines two design strategies that allow pilots to increase the number of concurrent tasks while holding the overall workload constant. The first strategy offers relevant information to the auditory channel, while the second strategy provides an auto-pilot to perform the tracking task. We also discuss three theories of workload that predict task performance in workload overload situations: single channel theory (SCT), single resource theory (SRT), and multiple resource theory (MRT). Using the models generated by these theories, we can compare actual pilot performance against that which is predicted by the models.

Proponents of SCT claim that the total time necessary to complete dual tasks is equal to the sum of the individual task times (Liao & Moray, 1993; Welford, 1967), and that parallel processing is impossible. Furthermore, SCT makes strong predictions about the relationship between the ISI (interstimulus interval) and the reaction time to the second-arriving task (RT_2). Specifically, if an operator is performing one task and a second task arrives shortly thereafter, the operator must complete the first task before initiating the second task. As the ISI between the tasks shortens, the RT_2 must increase linearly. This creates a -1 correlation between ISI and RT_2 (Welford, 1967).

In contrast, SRT argues that operators have access to mental resources which they can strategically allocate to multiple tasks, and that task interference is dependent on task difficulty (Kahneman, 1973). MRT expands on this concept by suggesting that tasks which use separate resource structures (e.g. different modalities), facilitate more efficient parallel processing than those which access the same resource structures (Navon & Gopher, 1979; Wickens, 2002). Both resource models would predict that, due to parallel processing, shorter ISIs have little to no effect on RT_2 .

Wickens & Dixon (2002) tested the two proposed strategies, or "offloads", in a single-UAV environment by presenting pilots with three different flight conditions: 1) a baseline condition requiring manual navigation with visual flight information; 2) an auditory offload providing auditory alerts for possible system failures (SF), with flight instructions offered via synthetic voice; and 3) an automation offload providing an accurate auto-piloting system. Participants were responsible for four concurrent, and continuous, tasks, with brief periods of intense target inspection designed to make maximal demands on available resources (Gugerty & Brooks, 2001). The authors reported benefits from both offloads. Specifically, the auditory offload improved SF monitoring and memory recall. The automation offload improved detection of

targets of opportunity (TOO). The authors also noted “cognitive tunneling” effects during the highly challenging task of target inspection that was not seen during low workload task combinations.

In the current experiment, we extend this paradigm to a dual-UAV flight control in order to examine these same offloads under multiple-workstation conditions. In particular, we examine the adequacy of the three theories to account for the data. In addition to doubling workload in the dual-UAV condition, the second primary difference between this, and the previous study, was the addition of motivational incentives for good performance, to determine if such incentives, by inducing a greater mobilization of effort, could mitigate the negative effects of cognitive tunneling observed in the previous study. A task timeline was generated between UAVs in order to analyze the extent to which concurrent supervision of two UAVs was accomplished by parallel processing, rapid task switching, or slow switching indicative of strict single channel theory (i.e. cognitive tunneling).

METHODS

Thirty-six licensed pilots received \$8 per hour, plus bonuses of \$10 and \$5, for 1st and 2nd place finishes, respectively, in their groups of 6 pilots. The Army typically uses pilots for UAV operation due to their specific training in flight control and visual scanning. Each simulated UAV had a separate workstation. Figure 1 presents a sample display for a single UAV.

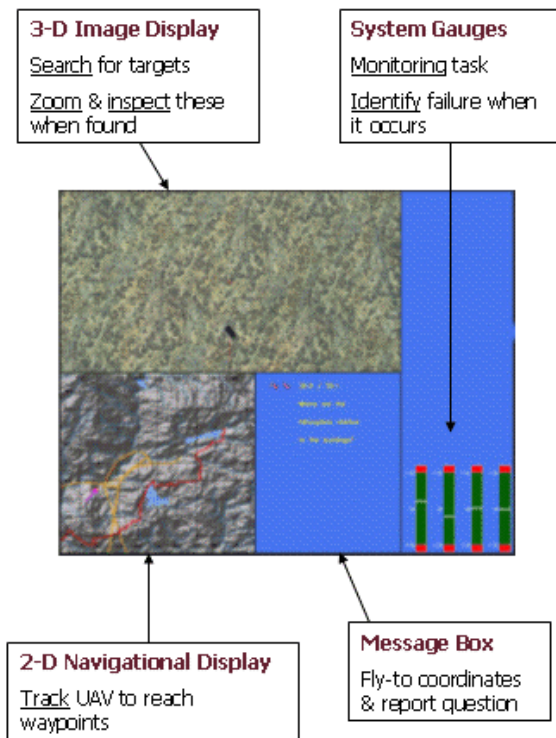


Figure 1. A UAV display with explanations for different visual areas.

Each pilot flew one or two UAVs through ten mission legs, while carrying out the following goal-oriented tasks: mission completion, target search, and systems monitoring. Beginning each leg, pilots received flight instructions via the Message Box (presented only over a set of speakers in the auditory condition). These flight instructions included fly-to coordinates to the next command target (CT) and a report question to be answered once the pilot arrived at the CT. Instructions could be refreshed by pressing a repeat button, which was used to measure pilot recall during each mission. In the baseline and auditory conditions, pilots were responsible for flying each UAV manually, but in the automation mode, they simply had to enter fly-to coordinates on the keypad and allow the auto-pilot to guide them directly to the next CT.

CT analysis involved loitering around the target area, manipulating a camera to inspect the CT, and reporting back relevant information to mission command. The task of target inspection involved mental rotation (e.g. *What weapons are located on the south side of the building?*), which has been described as a highly challenging UAV task (Gugerty & Brooks, 2001). Along each mission leg, pilots also searched for low-salience TOOs, which were reported in much the same way as CTs.

Concurrently, pilots were responsible for monitoring system gauges for possible SFs. Infrequently, a gauge would go “out of bounds”, necessitating the need for SF detection, correction, and report of flight coordinates during the SF. SFs were designed to fail either during relatively easy concurrent tasks (i.e. initiating heading or tracking) or during very difficult tasks (i.e. target inspection).

Similarly to Wickens & Dixon (2002), pilots flew in a baseline condition with all manual flight controls and visual displays, an auditory condition with SF alerts and aural flight instructions, and an automation condition which provided auto-pilot control of the UAV. Each of the flight conditions were performed in a single- and dual-UAV scenario; that is, each pilot flew both a single UAV and two UAVs in one of the conditions.

RESULTS

This section will only present relevant performance measures; see Wickens, Dixon & Chang (2003) for complete results. Many of the single-UAV results replicate findings from Wickens & Dixon (2002).

Mission completion

Flight path tracking performance benefited from automation, but neither benefited nor suffered from auditory offloading in either the single- or dual-UAV flight controls. CT response times and report accuracy did not differ across conditions or number of UAVs. Analysis of pilot use of the recall button for flight coordinates and report questions revealed fewer repeats in both the auditory ($p < .01$) and automation ($p < .001$) conditions compared to the baseline condition, presumably due to improved parallel processing and automated tracking, respectively. A dual-UAV decrement compared to single-UAV flight control ($p < .01$) was also

revealed, although an interaction between condition and number of UAVs ($p < .05$), suggests that the automation condition did not suffer from dual-UAV flight control as much as the other two conditions because pilots did not have to refresh fly-to coordinates for either UAV.

TOO monitoring

TOO detection rates improved under the automation condition ($p < .001$), for both single- and dual-UAV flight controls [$F(2, 33) < 1.0$], relative to baseline. While part of this improvement can be attributed to the auto-pilot, which flew directly over each TOO and facilitated more opportunities to detect TOOs in the 3D display, at least part of the improvement is also due to reduced workload. Of the TOOs that did appear in the 3D display during the baseline condition, the percentage of detections (76% for single-UAV and 52% for dual-UAV) was still lower than in the automation condition (92% and 79%) ($p < .01$). The auditory condition did not support any improvements in TOO detections ($p > .15$). TOO response times were essentially equivalent across conditions.

System Failures

The auditory condition produced significant benefits to SF detection rates ($p < .05$) and SF detection times ($p < .001$), while the automation condition provided no benefits to SF monitoring ($p > .15$). Figure 2 presents the SF detection times as a function of the condition and number of UAVs.

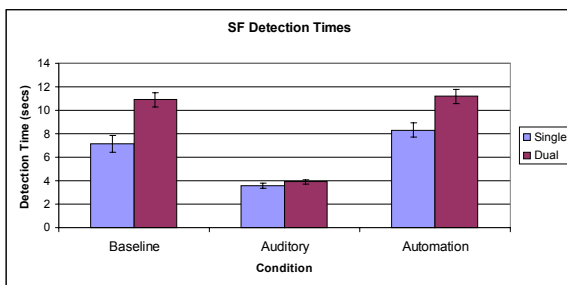


Figure 2. SF detection times across condition and workload of side task. Standard error (SE) bars are included.

As seen in Figure 2, dual-UAV decrements were found for the baseline ($p < .05$) and automation ($p < .05$) conditions, but were absent in the auditory condition; that is, performance during dual-UAV flight did not suffer compared to single-UAV flight when SFs were offloaded to the auditory channel. This is consistent with MRT model predictions, but not with SCT or SRT predictions. Further analysis of individual SFs revealed shorter response times for SFs which occurred during periods of simple monitoring than SFs which occurred during target inspection ($p < .001$), lending credence to the SRT and MRT models which state that difficult tasks use more mental resources than simpler tasks. However, an interaction effect ($p < .05$) showed a much greater penalty of concurrent target inspection for the baseline and automation conditions (3-7 sec) than for the auditory condition (1-2 sec). In short, auditory offloading appeared to preserve performance despite: a) the

extremely demanding task of target inspection, and b) doubling the number of UAVs. Only MRT would predict that improved parallel processing was made possible by offloading one of the tasks to a separate modality.

Theory predictions

Strict SCT predicts that parallel processing is impossible during dual-task conditions, particularly if events are widely separated in space such as in the dual-UAV layout. A timeline analysis between UAV workstations examined the effects of multi-tasking across UAVs. Task combinations of SF monitoring combined with target inspection were considered, with each task appearing on a separate workstation. According to SCT, the total time for both tasks should be at least the sum of the individual tasks, and this is indeed seen for many task combinations; however, in the auditory condition, some parallel processing is evident as pilots completed both tasks in less time than predicted if they were done serially.

Another prediction of strict SCT was tested by assessing the interstimulus interval (ISI) between tasks. SCT predicts that unless there is switching between tasks, every one-second shortening of this ISI will produce a concomitant one-second lengthening of response time to the second-arriving task (RT_2) (Pashler, 1998; Welford, 1967). Similarly, an absence of increased RT_2 must correspond with an increase in RT_1 . Figure 3 presents the values of RT_2 plotted against the ISI. While SCT would predict a -1 correlation between ISI and RT_2 , the data reveal correlations of $r = +0.13$ and $r = +0.68$ for RT_2 and RT_1 , respectively, suggesting that instead pilots were putting more effort into finishing the first task quickly when they detected a second-arriving task.

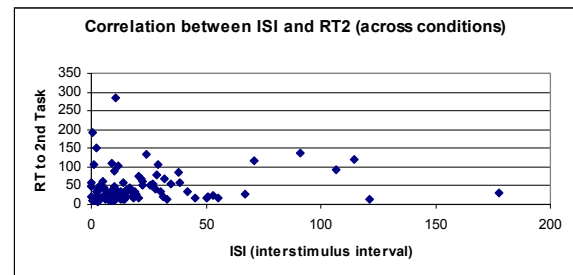


Figure 3. A scatterplot between interstimulus interval (ISI) and RT to the second-arriving task.

According to a pure version of single resource theory, the interference between tasks should be a direct and exclusive function of their task demand value. Values for each task were assigned on the basis of the least integer ordinal heuristic, by which ordinal relations between demand values are preserved, but the smallest integer (or half integer) values possible are chosen. In order to test the predictions of single resource models, the total concurrent demand value of each of the three single and three dual task combination was computed, simply by summing the demand values of concurrent tasks. Correlations ($N=6$) between predicted workload and obtained data revealed the following: +0.96 (Repeats), +0.93 (TOO detection rate), +0.49 (SF detection rate) and +0.18 (SF

detection time). Not surprisingly, the data for the SF task measures reflect a poor model fit, as SRT does not account for differences between auditory and visual displays. Only MRT would make such predictions.

GENERAL DISCUSSION

Single-UAV effects

Results from the single-UAV flight control generally replicated findings from Wickens & Dixon (2002). Interference between concurrent visual tasks created significant performance decrements in the baseline condition, while auditory offloading benefited SF monitoring and flight instruction recall, and automation offloading facilitated improved TOO monitoring. However, in contrast to Wickens & Dixon (2002), the auditory offloading appeared to eliminate the “cognitive lockup” found previously during target inspection, as SF monitoring suffered very little in the high workload conditions. We attribute this difference to added incentives provided for excellent performance.

Dual-UAV effects

Dual-UAV decrements were seen in all tasks except the primary tasks of tracking and report accuracy. Given the importance of the primary mission task and the low bandwidth necessary to monitor flight controls (i.e. no turbulence), the latter null results are not surprising. Dual-UAV decrements were revealed particularly in the two monitoring tasks (TOO and SF), but were mitigated somewhat by the two offloads in different ways.

Auditory offload

While the auditory offload had no discernible effect on tracking or TOO monitoring, it did facilitate dramatic improvements to the SF monitoring task, as well as the recall task. Given the availability of the auditory presentation, pilots were able to visually monitor other displays while using separate perceptual resources for SF alerting and the retention of flight instructions. Benefits to the SF task were not limited to the single-UAV flight control; auditory offloading was also able to buffer any negative effects of dual-UAV flight control.

Automation offload

While the automation benefit to tracking was an intentional artifact of the simulation, benefits to the TOO monitoring task and the recall task were also revealed, presumably because pilots were able to reallocate resources to these tasks. These results were even more pronounced in the dual-UAV flight control.

Theoretical Interpretations and relevance for workload models

As mentioned previously, the current study seeks to compare model predictions with pilot performance, with the

goal of understanding these predictive qualities so that future researchers can forecast operator behavior through simulation generated by the models.

SCT was able to predict some pilot behavior, particularly in the baseline condition. Specifically, the 3-second dual-UAV decrement to SF detection times indicate that pilots were unable to visually scan both workstations simultaneously. Furthermore, the timeline analysis reveals that the time necessary to complete certain task combinations (except during the auditory condition) between workstations is greater than the sum of the task times, indicating some switching costs (Miller, 2002). This model, however, failed to predict the mitigation of the offloads, particularly the auditory benefits found in the SF monitoring task.

A second weakness in the SCT model was revealed in the absence of a -1 correlation between the ISI and RT_2 . If pilots were choosing instead to switch to the second-arriving task before completing the first task, then RT_1 should increase in one-second increments corresponding to the decrease in the ISI. This also was not seen, rather it appeared that pilots were allocating more effort to the first-arriving task in order to accelerate its processing when they detected the second-arriving task, a behavior consistent with resource theories (Kahneman, 1973).

In contrast with Wickens & Dixon (2002), and again consistent with resource theory, added incentives for superior performance appeared to eliminate “cognitive tunneling” with the auditory offload, even in dual-UAV flight control. As our analysis shows, correlations between model predictions and experimental data for TOO detection rate and Repeats were quite high, with the single resource demand model accounting for over 85% of the variance in performance across the six conditions. Single resource model fit was less impressive in accounting for variance in SF monitoring performance. This is not surprising, given the strong role of multiple resources in influencing performance on the SF task.

It appears that MRT is responsible for predicting substantial aspects of time sharing variance between visual displays and workstations, results not accounted for in the other models. Therefore, an effective computational model needs to incorporate all three characteristics: resource demand (SRT) and resource structure (MRT), along with some single channel switching assumptions (SCT).

Practical Conclusions

The current data suggest that simultaneous control of two UAVs by one pilot is possible, particularly if certain tasks are either offloaded to the auditory channel or assisted by automation. Most importantly, the primary mission task did not appear to suffer from an additional UAV; on the other hand, most of the “secondary” tasks did suffer in the dual-UAV flight control, although these performance decrements were mitigated somewhat by the two offloads. Lastly, we note that this evaluation was carried out in a “best case” scenario; that is, both forms of offloading were completely reliable. It is quite likely that imperfections or unreliability characteristics of the hostile world in which UAVs are expected to operate, would amplify considerably the demands for both visual and

cognitive aspects of attention. At the same time, it is true that use of more qualified and experienced UAV pilots might overcome some of these performance decrements.

ACKNOWLEDGMENTS

This research was sponsored by a subcontract # ARMY MAAD 6021.000-01 from Microanalysis and Design, as part of the Army Human Engineering Laboratory Robotics CTA, contracted to General Dynamics. David Dahn was the scientific/technical monitor. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the Army CTA. The authors also wish to acknowledge the support of Ron Carbonari and Jonathan Sivier (in developing the UAV simulation), of Bobby Bernard and Mark Juntenen for assisting with data collection, and of Dr. Michael Barnes of the Army Research Lab at Ft. Huachuca, Arizona for assisting in interviewing RPV pilots within the E CO, 305th Military Intelligence Battalion to carry out the cognitive task analysis that underlies the simulation developed.

REFERENCES

Barnes, M.J., Ghiradelli, T., Stachowiak, C., Hill, S. & Dahn, D. (2002). *Future combat systems*. Unpublished paper. Ft Huachuca, AZ: US Army Research Laboratory.
Gugerty, L. & Brooks, J. (2001). Seeing where you are heading. *J. Expt. Psych: Applied*, 7(3), 251-266.

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice Hall.
Liao, J. & Moray, N. (1993). A simulation study of human performance deterioration and mental workload. *Le Travail humain*, 56(4), 321-344.
Navon, D. & Gopher, D. (1979). On the economy of the human processing systems. *Psychological Review*, 86, 254-255.
Pashler, H.E. (1998). *The psychology of Attention*. Cambridge, MA: MIT Press.
Sarno, K.J., & Wickens, C.D. (1995). The role of multiple resources in predicting time-sharing efficiency. *IJAP*, 5(1), 107-130.
Welford, A.T. (1967). Single channel operation in the brain. *Acta Psychologica*, 27, 5-21.
Wickens, C.D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomic Science*, pp. 1-19.
Wickens, C.D. & Dixon, S. (2002). *Workload demands of remotely piloted vehicle supervision and control* (ARL-02-10/MAD-02-1). Savoy, IL: University of Illinois, Aviation Research Lab.
Wickens, C.D., Dixon, S., & Chang, D. (2003). *Using interference models to predict performance in a multiple-task UAV environment: 2 UAVs*. Savoy, IL: University of Illinois, Aviation Research Lab.