DESIGN AND VALIDATION OF A SYNTHETIC TASK ENVIRONMENT TO STUDY DYNAMIC UNMANNED AERIAL VEHICLE RE-PLANNING

Maia B. Cook, Harvey S. Smallman, Frank C. Lacson, and Daniel I. Manes
Pacific Science and Engineering Group, Inc., San Diego, CA

A key challenge facing unmanned aerial vehicle (UAV) operators is the need to re-plan routes on-the-fly when situations change. Operators must comprehend three-dimensional (3D) scenes and a multitude of potentially competing 3D mission and routing constraints in order to successfully re-plan, often under time pressure. Currently, there is significant research interest in supporting UAV operators through automation and improved visualizations. However, development and integration of these methods requires a careful understanding of the 3D spatial awareness challenges and requirements facing operators. To facilitate this understanding, here we report the design and validation of a synthetic task environment (STE) and testbed to study UAV re-planning. The STE is derived from a recent task analysis conducted with Navy UAV operators that focused on the key 3D spatial challenges entailed in re-planning. In an initial validation of the STE implemented in a re-planning testbed, several measures of re-planning performance were assessed for 36 participants working through controlled re-planning scenarios. The presence of mountainous terrain and the spatial overlap of mission constraints were parametrically varied. Performance was consistently worse in mountainous terrain, and in more highly-constrained conditions in mountainous terrain. In flat terrain, however, less constrained conditions resulted in paradoxically worse performance. Results have both basic and applied implications. Theoretically, the study provides a bridge between applied re-planning research and classic human problem solving work by allowing apparently simpler, unconstrained re-planning to be conceived of as less bounded search through re-planning problem space. For application, the results help constrain and define the requirements for future 3D visualization and automation support for UAV re-planning displays.

INTRODUCTION

Unmanned aerial vehicle (UAV) operators face a multitude of human factors challenges: their tasks are varied and complex, their workload is high, the UAV routes they plan must satisfy complex competing constraints, and their interactions with UAVs are indirect and mediated by visual displays. These challenges are exacerbated when changes to an original route are required during a mission, due to emergent events, and the complex route re-planning must be resolved on-the-fly, often under intense time pressure.

Previous task analyses have catalogued some of these emergent events requiring UAV re-planning (De Vries, Roefs, & Theunissen, 2007; Gugerty, DeBoom, Walker, & Burns, 1999), such as changes to weather and other environmental conditions, and re-tasking to track new intelligence, surveillance, and reconnaissance (ISR) targets. These new constraints must be satisfied, while simultaneously satisfying other factors, including avoiding other aircraft and terrain, remaining undetectable by hostiles, and complying with complex airspace requirements. Though similar to air traffic control (ATC) in some respects, UAV re-planning requirements are different in terms of their objective (e.g., collecting imagery), competing nature (e.g., collecting imagery while remaining undetectable), and dynamic quality (e.g., recurring ISR re-tasking).

Routes are 3D, as are the spatial configuration of events triggering re-planning and the issues and constraints associated with them. The understanding of the 3D spatial awareness UAV operators require during re-planning is a relatively unexplored issue. UAV planning and control is currently mediated by conventional, top-down, two-dimensional (2D), topographic displays, occasionally augmented with side (profile) views. However, previous research casts doubt on the suitability of such 2D displays for conveying 3D spatial awareness, because the 2D format makes it particularly challenging to reconstruct the 3D configuration and layout of the scene (see St. John, Cowen, Smallman, & Oonk, 2001). Operators' 3D spatial awareness may therefore be poorly supported by their UAV control displays.

Two sources of potential interventions have been touted to support the task and display challenges of re-planning: automated path planning algorithms to offer re-planning solutions and reduce workload (Cummings, Marquez, & Visser, 2007; Lavalle, 2006), and improved visualizations to generally support 3D spatial awareness (Theunissen, Bolderheij & Koeners, 2005). The work reported here is concerned with display interventions. The introduction of either automation or new displays into the path planning process requires a careful characterization and understanding of the 3D spatial challenges facing operators.

Recently, we conducted a task analysis focused on this issue of 3D spatial awareness challenges and requirements with Navy UAV operators (Cook & Smallman, 2008). The key challenges elicited from our task analysis were integrated with results from previous task analyses (e.g., Gugerty, 2004), and are represented in the design of a synthetic task environment (STE) for studying re-planning in a testbed. Here, we report the design of the STE embodied in a validation study that baselines re-planning performance under quasi-operational conditions.

In the design of the STE, several aspects of the 3D re-planning space perception challenges are abstracted, such as the competing goals of approaching ISR targets while avoiding closed airspace and terrain, all under time pressure. To study re-planning experimentally, we derived two independent variables (IVs) from our task analysis to represent the key 3D re-planning challenges: type of terrain (flat vs. mountainous), and degree of spatial overlap of re-planning event constraints (low-constraint vs. high-constraint). The second IV is related to the relational attribute of complexity in the ATC domain (Xing & Manning, 2005). To assess re-planning performance quasi-operationally, we measured the time to simply satisfy (not optimize) competing constraints.

Re-planning was predicted to be worse in mountainous versus flat terrain due to the challenges of understanding the 3D mountainous terrain shape from conventional 2D displays. Re-planning was expected to be worse for high-constraint re-planning events due to the expected increased complexity from the spatial overlap of the competing 3D mission re-planning goals. The high-constraint disadvantage was expected to be exacerbated by mountainous terrain. Spatial ability was measured to determine its mitigating effects on re-planning performance for each IV.

EXPERIMENT

The general abstraction of the STE re-planning elements is described first, followed by a description of the specific configuration of the IVs for the experiment.

Design of UAV re-planning STE. The abstracted display elements and re-planning events derived from our task analysis were configured in a series of scenarios. The scenarios replicated the key re-planning goals and 3D challenges faced by UAV operators. In the STE, the UAV scene was shown in two separate displays, similar to conventional UAV map displays: a top-down display showing a bird's-eye view, and a side profile view showing altitude of the scene along the route, see Figure 1. The displays showed the terrain that the UAV was flying over, the UAV routes with waypoints, the UAV's speed and position along the route to convey distance from re-planning events along the route, ISR targets that needed to be within range of the route, and closed airspace that the route needed to avoid. The specific implementation of these elements in the experiment is described next. The experiment was carefully balanced to abstract these re-planning elements to a form that allowed them to be controlled and studied, yet ensured they were still generalizeable to the environment and situations faced by real UAV operators.

Method

Participants. Thirty-six college students or graduates (22 male, 14 female) with a mean age of 36.4 (range 18-65 years) were recruited from www.craigslist.org and received $30 for participating.

Stimuli. To represent the varied terrain regions where UAVs operate, the scenarios included terrain samples, created from U.S. Geological digital elevation models (DEMs), that each contained approximately equal-sized regions of flat and mountainous terrain. In the top-down view, similar to conventional UAV map displays, terrain altitude was color-coded, with darker colors progressively indicating higher altitude, and had equal-altitude contour lines indicating steepness in altitude change. The profile view showed the terrain directly beneath the route in silhouette, see Figure 1.

To reflect the ISR mission requirements of gathering imagery on targets of interest, each scenario included two targets, shown on the map display as MIL-STD red diamonds. The ISR coverage range was an invisible line-of-sight limited dome centered on the target. When the route came within range and line-of-sight of a target, the target outlined in dark red, as a proxy for another (payload) operator indicating that acceptable image quality could be achieved from that position.

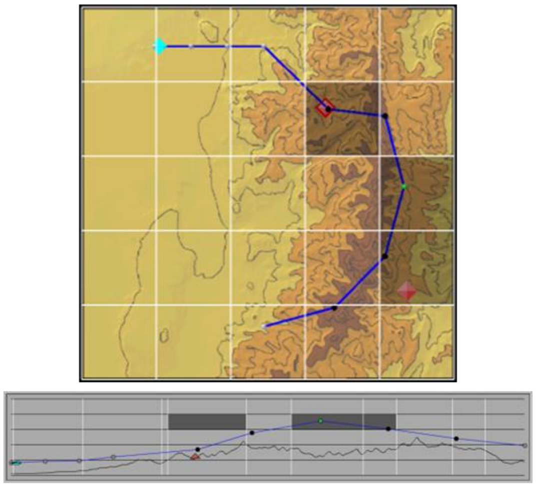

Figure 1. A high-constraint (overlap) re-planning trial in mountainous terrain illustrating the top-down (top) and profile view (bottom) testbed displays.

UAVs must fly in complex, congested airspace, avoiding no-fly zones and air traffic, and obeying airspace rules. To abstract these airspace availability issues in the experiment, a 5 x 5 grid was superimposed on the map to convey airspace boundaries and closures. Three closed airspace cubes operationally representing no-fly zones or regions of air traffic were shaded in gray, and available airspace was unshaded.

UAV re-planning often occurs under time-pressure. In the experiment, the passage of time was conveyed by a UAV icon travelling along the UAV route. Re-planning had to be completed before the UAV reached the events along the route necessitating re-planning.

The targets and closed airspace were always visible in the top-down view, and appeared in the profile view only if part of the route occupied the same lateral location; see targets and airspace cubes in Figure 1.

Design. Two independent variables were varied in a

2 (*terrain type*: flat vs. mountainous) x 2 (*event constraint level*: low-constraint vs. high-constraint) within-participants design, resulting in four conditions. Terrain type was classified by the region of terrain containing the re-planning events. The event constraint manipulation involved the degree of overlap, *in top-down*, between the closed airspace cubes and the two targets: airspace and targets did not overlap for low-constraint events, but did overlap for high-constraint events, see Figure 1.

There was one practice and six experimental scenarios with four trials each (one trial for each of the four conditions). The four trials were blocked by terrain, and condition and terrain order were counterbalanced.

Procedure. After informed consent and color vision screening, participants completed the Vandenberg Mental Rotation Test (MRT) of spatial ability. Next, participants were asked to role-play a UAV operator, monitoring routes for targets moving out of ISR range and for closed airspace emerging along the route, and ensuring that the route cleared terrain. Each trial consisted of two phases: a starting route phase, where the route was within range of the targets, and avoided closed airspace and terrain, and a re-planning phase, where the targets moved out of the route's ISR range and closed airspaces appeared along the planned route.

Participants re-planned routes using the top-down and profile view displays, with the goals of getting the route within range of both ISR targets, and out of closed airspace, while clearing terrain, all as quickly as possible. Participants performed lateral route changes by moving waypoints in the top-down view, and vertical changes in the profile view. As lateral movements were made in the top-down view, the profile view updated to show elements encountered along the newly-adjusted top-down route. Participants were free to make lateral and vertical movements in any order. The lateral and vertical waypoint adjustment range was limited to represent the physical kinematic limitations of a UAV.

Participants were instructed to submit a re-planned route when they had corrected all target, terrain, and airspace errors, before the UAV icon reached a set point on the route, which occurred at 90 seconds. The entire procedure lasted about two hours.

Dependent variables. Re-planning performance was measured by *correct RT*, response time for trials with correct re-planned routes submitted within time, and *# movements,* the number of waypoint movements per trial, as a measure of re-planning effort. Re-planning accuracy was measured by *overall errors*, the proportion of trials containing an error, and *specific errors*, overall errors split into the proportion of trials containing *target misses*, *terrain incursions*, and *airspace violations*.

Results

The results are first organized by effects for terrain type, event constraint level, and their interaction, to facilitate understanding of the results pattern across the multiple dependent variables. Repeated-measures ANOVAs and paired t-tests for significant interactions were conducted. Next, the relationship between spatial ability and re-planning performance is reported. Finally, the overall pattern of errors is reported to identify specific relationships between errors and to motivate future directions for the testbed and experiments.

Terrain type. Re-planning was consistently slower and more effortful in mountainous versus flat terrain (correct RT: $F(1,35) = 192.2$, $p < .001$, $\eta^2 = .85$; # movements: $F(1,35) = 189.2$, $p < .001$, $\eta^2 = .84$), see Figure 2, top. This was not due to a speed-accuracy tradeoff since there were more overall and target, terrain, and airspace errors in mountainous terrain (overall errors: $F(1,35) = 61.4$, $p < .001$, $\eta^2 = .64$), see Figure 2, bottom. Thus, the challenge of re-planning in mountainous terrain was evidenced by increased time and effort in addition to increased errors.

Event constraint level. Contrary to prediction, re-planning performance was the same across event constraint levels, and in one case (airspace errors), was worse for low- versus high-constraint events, $F(1,35) = 8.5$, $p < .01$, $\eta^2 = .20$. The predicted pattern for event constraint level was found only in mountainous terrain, discussed in the next section.

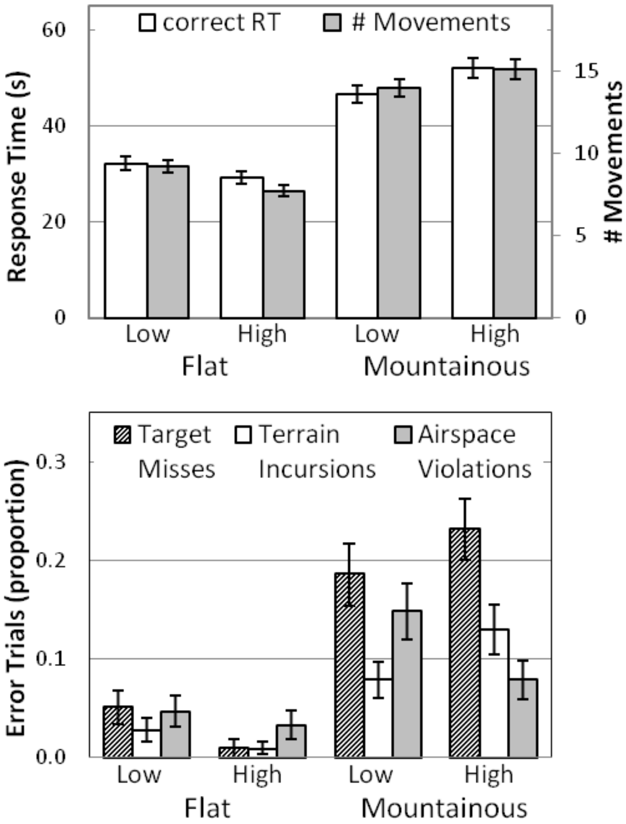

Figure 2. Re-planning time and waypoint movements (top), and error rates (bottom).

Terrain type by event constraint interactions. For re-planning time and effort, terrain type interacted with event constraint level (correct RT: $F(1,35) = 12.8$, $p = .001$, $\eta^2 = .27$; # movements: $F(1,35) = 19.0$, $p < .001$, $\eta^2 = .35$); see Figure 2, top. More constrained routing through mountainous terrain led to the predicted longer RT and greater effort (correct RT: $t(35) = 2.7$, $p < .05$; # movements: $t(35) = 2.3$, $p < .05$). The pattern reversed in flat terrain, however. More time and effort was required for low- versus high-constraint events (correct RT: $t(35) = 2.6$, $p < .05$; # movements: $t(35) = 5.2$, $p < .001$).

Overall errors, and target and terrain errors also exhibited a similar interaction pattern (overall errors: $F(1,35) = 4.3$, $p = .05$, $\eta^2 = .11$; target errors: $F(1,35) = 6.0$, $p < .05$, $\eta^2 = .15$; terrain errors: $F(1,35) = 4.9$, $p < .05$, $\eta^2 = .12$). For each, the high-constraint advantage generally existed in flat terrain, but either was canceled (overall and target errors) or exhibited a reversal trend in mountainous terrain (terrain errors). There was no terrain type by event constraint interaction for airspace errors.

In flat terrain, the high-constraint events may have bounded the re-planning problem space in such a way that performance improved. However, the beneficial boundedness effects of high-constraint in flat terrain did not extend to mountainous terrain, possibly due to the additional 3D shape understanding challenges imposed by mountainous terrain.

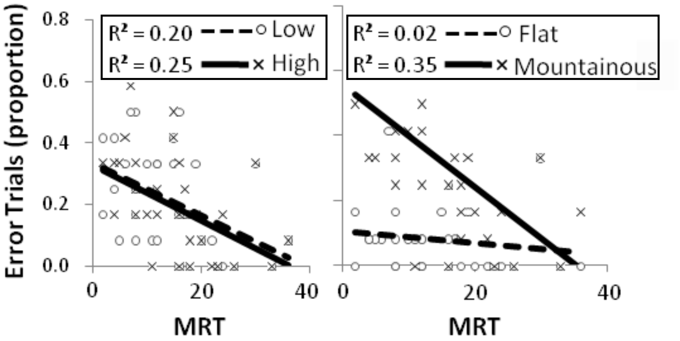

Figure 3. Overall errors and MRT split by event constraint (left) and terrain type (right).

Spatial ability and re-planning performance. Out of a possible maximum of 40, the mean MRT score was 14.8 ± 8.4 (range 2 - 36), 16.8 for males, and 11.5 for females. When combined across conditions, those with higher MRT made fewer errors, $r(34) = -.51$, $p = .001$, though a strikingly different pattern emerged for each IV. While the correlation for overall errors and MRT was similar across low- and high-constraint conditions, it was markedly different for flat and mountainous terrain, see Figure 3. Overall errors and MRT were unrelated for flat terrain, but were strongly related for mountainous terrain, $r(34) = -.59$, $p < .001$. This pattern in mountainous terrain persisted for terrain ($r(34) = -.50$, $p = .001$), target ($r(34) = -.52$, $p = .001$), and airspace errors ($r(34) = -.28$, $p = .05$). Thus, spatial ability was more helpful for the 3D perceptual challenges of mountainous versus flat terrain. It was equally helpful across the low- and high-constraint problems, which may have involved more similar spatial demands.

Frequency and distribution of errors. Re-planning errors occurred when the allotted re-planning time was exceeded (timeout errors) and, surprisingly, when routes were submitted within time (submit errors). The majority of target, terrain, and airspace errors were timeout errors (110), though an unexpectedly high number were submit errors (66). We investigated the error pattern overall, and whether the target, terrain, and airspace error pattern was similar across timeout and submit trials, see Figure 4.

First, note that more errors were related to targets than to airspace or terrain. This may have been due to the smaller goal region for targets relative to airspace and terrain, and the lack of an explicitly displayed target coverage region.

Second, a different pattern emerged for target misses across timeout and submit errors: target misses accounted for the majority of timeout errors (85%), but the minority of submit errors (18%). Submit errors were higher for goals that did not provide explicit feedback on success: terrain (44%) and airspace (41%).

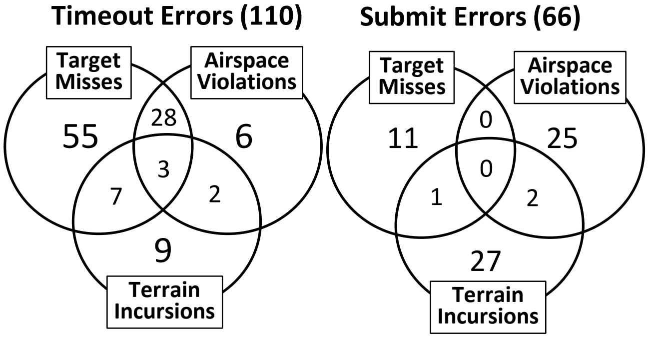

Figure 4. Distribution of specific errors across timeout and submit error trials.

Third, more than one-third of timeout trials involved multiple error types (e.g., airspace and target), while almost all submit error trials involved only a single error type (e.g., airspace). Participants failed to notice these submit errors, especially those without explicit success feedback (airspace and terrain). Further analysis is needed to determine if these oversights for submit errors are attributable to the reduced spatial overlap of constraints for low-constraint events, and whether this might contribute to the higher airspace errors for low- versus high-constraint trials.

DISCUSSION

The dynamic re-planning of UAVs is challenging, and the constellation of underlying 3D space perception challenges was evident in the experiment performance patterns. For all measures, performance was consistently worse in mountainous terrain. The challenge of extracting the 3D shape of mountainous terrain from two separate 2D views, and the more variable profile of the mountainous versus flat terrain, are among the factors likely contributing to this result. An integrated view is a natural display intervention to support the 3D understanding of mountainous terrain and other objects in the UAV scene. However, such an intervention must be introduced with care, and may support only a subset of re-planning tasks (Haskell & Wickens, 1993; St. John et al., 2001).

Relative to terrain type, the effects of the event constraint IV were less robust, and generally differed depending on the type of terrain. The predicted performance advantage for low-constraint events was observed in mountainous terrain, but was reversed in flat terrain, where performance was better for high- versus low-constraint events. The benefit for high-constraint events in flat terrain, though seemingly unintuitive when high-constraint is conceived of as more complex than low-constraint, can be related to Newell and Simon's (1972) classic original conception of problem solving as a search through a problem space. The search through UAV re-planning space can be considered analogous, with the goal attained as soon as a viable constraint-satisfying path is found. The high-constraint condition in flat terrain bounded the search for re-planning solutions in a way that movements more quickly and directly satisfied the re-planning goals. However, this boundedness became an obstacle in mountainous terrain, restricting the solution space in a way that required more movements and exploration to discern. Thus, degree of constraint does not have a constant effect, or meaning, across all environments. The event constraint results have implications for predicting the types of events that present the greatest challenges to UAV re-planning. The complex moderating effects of constraint level should be explored in detail in future studies to carefully elucidate its impact. Additionally, the event constraint findings suggest design interventions that usefully bound the problem space, to help guide users towards viable regions in re-planning space.

The different accounts of the effects of the two IVs are reflected in, and validated by, the spatial ability results. The strong relationship between spatial ability and re-planning had the greatest impact for the variable with the greatest difference in 3D space understanding challenges: terrain type. There was no differentiation between the contributions of spatial ability to performance on low- and high-constraint events. This suggests that performance differences across event constraint levels may be more attributable to other abilities, such as those related to cognitive problem solving, consistent with our bounded problem solving account for that variable.

Finally, relative timeout error rates were partially mitigated for re-planning goals whose spatial boundaries were shown explicitly (airspace and terrain) versus only implied (targets). Relative submit error rates were partially mitigated for re-planning goals whose success was explicitly shown (targets) versus confirmed by visual inspection (airspace and targets). These two results taken together suggest the importance of making goal shape boundaries explicit and using feedback to convey re-planning success, but these two approaches did not eliminate all errors. Although a critical system component, warnings should not be considered a panacea for error prevention: users will ultimately have to understand whether the re-planning goals are met in cases where automation fails or needs validation. The error rates are suggestive of an explanation for the unique pattern of airspace error results, and our future work will examine the differences between the specific error types in more detail.

This initial study begins to identify possible leverage points for design interventions to support re-planning. Our future work will develop design interventions discussed here, assess these interventions in our testbed, and inform the design of future displays to support re-planning. Additionally, interventions through visualizations of path planning algorithm outputs to support re-planning are under development, and will also be assessed in our testbed environment.

ACKNOWLEDGEMENTS

This work was sponsored by the Office of Naval Research, program manager Dr. Gerald Malecki. The authors thank Kevin Aeling and Kathryn Imler for technical assistance, and Heather Shaver for data collection.

REFERENCES

Cook, M.B., & Smallman, H.S. (2008). When plans change: task analysis with navy UAV operators, display requirements, and UAV re-routing taxonomy. San Diego, CA: Pacific Science & Engineering Group, Inc., Technical Report.

Cummings, M.L., Marquez, J. J., & Visser, M. (2007). Shadow TUAV Single Operator Consolidation: Display Assessment. HAL2007-01 Technical Report, MIT Cambridge, MA.

De Vries, M., Roefs, F., & Theunissen, E. (2007). Route (re-) planning through a hostile, dynamic environment: human biases and heuristics. In *Proceedings of the Digital Avionics Systems Conference*, DASC '07. IEEE/AIAA 26th, 21-25 October 2007.

Gugerty, L., DeBoom, D., Walker, R., & Burns, J. (1999). Developing a simulated uninhabited aerial vehicle (UAV) task based on cognitive task analysis: Task analysis results and preliminary simulator performance data. In *Proceedings of the 43rd Annual Meeting of the Human Factors and Ergonomics Society* (pp. 86-90).

Gugerty, L. (2004). Using cognitive task analysis to design multiple synthetic tasks for uninhabited aerial vehicle operation. In S. Schiflett, L. Elliott, E. Salas & M. Coovert (Eds.), *Scaled Worlds: Development, Validation, and Application*, (pp. 240-262). London: Ashgate Publishers.

Haskell, I.D., & Wickens, C.D. (1993). Two- and three-dimensional displays for aviation: A theoretical and empirical comparison. *International Journal of Aviation Psychology*, **3**, 87-109.

LaValle, S.M. (2006). *Planning algorithms*. Cambridge: Cambridge University Press.

Newell, A., & Simon, H.A. (1972). *Human problem solving*. Englewood Cliffs, N.J.: Prentice-Hall.

St. John, M., Cowen, M.B., Smallman, H.S., & Oonk, H.M. (2001). The use of 2D and 3D displays for shape understanding vs. relative position tasks. *Human Factors*, **43**, 79-98.

Theunissen, E., Bolderheij, F., Koeners, G.J.M. (2005). Integration of threat information into the route (re-) planning task. In *Proceedings of the 24th Digital Avionics Systems Conference*, Washington, D.C.

Xing, J., & Manning, C.A. (2005). Complexity and automation displays of air traffic control: literature review and analysis. Washington DC: Federal Aviation Administration; Report No: DOT/FAA/AM05-4.