FINAL REPORT

Benefits of an information organization tool in strategic air traffic control

Scott D. Gronlund, John M. Canning, Peter M. Moertl, Joakim Johansson

University of Oklahoma,

Michael R. P. Dougherty, University of Maryland,

and

Scott H. Mills

FAA, CAMI
Abstract

A research group from the University of Oklahoma, together with the cooperation and assistance of the FAA Academy (especially Ron Andrei), completed a series of experiments examining the strategic planning abilities of en route air traffic controllers. First, Dougherty, Gronlund, Durso, Canning, and Mills (1999) showed that specialized computer tools would be important to enhance the solving of sequencing problems because sequencing problems proved to be the most strategically challenging. Next, Canning, Johansson, Gronlund, Dougherty, and Mills (1999) described a new interface that we developed and showed that it had several advantages over the traditional flight progress strips in terms of developing plans for sequencing aircraft. The interface replaced the paper flight progress strips with virtual tokens, which were dynamically-linked with their radar target. Finally, in the current paper, we describe our third study—a pilot study conducted with five controllers. In this study, we examined the effect of our new interface’s automated sequencing tool and found that it sped up the development of the plan, resulted in the creation of better plans, and reduced the workload of the tactician who implemented the plan. We conclude with recommendations for additional empirical and interface development work.
Benefits of an information organization tool in strategic air traffic control

Strategic (premeditated) decision making or planning is fundamental to the completion of many dynamic tasks. For example, Miller, Copeland, Heaton, and McCloskey (1998) found good strategic decision making to be important for successful military operations (see also Pew & Mavor, 1998), and Xiao, Milgram, and Doyle (1997) found it to be an important aspect of anesthesiologists' preparation for surgery. Air traffic control, however, has been characterized as primarily involving tactical (i.e., reactive) decision making (Durso & Gronlund, 1999; Hutton, Olszewski, Thordsen, & Kaempf, 1997). However, work at the University of Oklahoma has shown that strategic planning is also an important contributor to successful air traffic control.

Strategic decision making is important given future concepts being proposed for the air traffic control system. For example, Lawson & Thompson (personal communication, Dec., 15, 1997) have argued for the creation of a strategic controller position—sometimes called a multi-sector D-side controller. The strategic controller position is similar to one that researchers at several institutions have proposed (e.g., NASA's Ames Research Center, Vivona et al., 1996). One person would be responsible for a multiple-sector airspace, making strategic planning decisions about air traffic in that airspace, and delegating responsibility for tactical decisions to sector-level controllers. It was for this type of position that our interface tools were developed.

Before we could begin to develop new interface tools, we first needed to better understand how controllers use their current tools to develop strategic plans. The first problem to be solved involved disentangling strategic from tactical planning. Roughly speaking, strategic planning occurs further in advance than tactical planning and involves a larger number of aircraft. Tactical planning occurs in the current moment and involves the separation of (usually) pairs of aircraft. Normally, tactics and strategies are confounded because both types of decision making lie within the same head (even when a team has responsibility for a sector). We solved this problem by putting the strategist and tactician into "different heads" to get a "clean" sample of strategic behavior.
We (Dougherty, Gronlund, Durso, Canning, & Mills, 1999) examined twelve full-performance level (FPL) en route air traffic controllers serving as instructors at the FAA Academy. The experiment began with instructions that delineated the roles of the tactician (the subject matter expert or SME) and the strategist (the participant). The tactician's job was to maintain separation between aircraft; he made whatever altitude, speed, and heading changes were necessary to maintain separation; the strategist's job was to tell the tactician how to control the traffic and to make the tactician's job as easy as possible. Participants completed two different types of scenarios. In one scenario, most aircraft flew direct or straight-line routes through the sector, which meant that pairs of aircraft could intersect at any point in the airspace. The primary problem to be solved involved maintaining separation among aircraft. For the other scenario, the aircraft flew on standard routes (so-called “highways in the sky”); the primary problem to be solved involved sequencing aircraft into Dallas/Fort Worth (D/FW).

We found that the scenario that involved crossing traffic was tactically complex, but not very strategically complex. In contrast, the sequencing scenario was more strategically challenging. Because strategists were instructed to use the FPS’s only if it helped their strategic planning, we assumed that the usage of FPS’s was an indicator of strategic processing. Our participants relied more on FPS’s in the sequencing scenarios. A greater proportion of FPS’s were marked in the sequencing scenario (.45 vs..24), and more FPS’s were moved at the beginning of the sequencing problems (5.71 vs. 1.5). Ten of the 12 strategists built their sequences using only the FPS’s, reshuffling their order in the bay to correspond to who was first in line, second, etc. Two controllers used the FPS’s to remember which aircraft were going where (offset right for Houston and left for OKC, or used right bay for Houston, left bay for OKC).

Because we put the strategist and tactician into different heads, and because we required the strategist to verbally communicate their plan to the tactician, the strategist’s verbalizations provided another rich source of data assessing controller performance. Analyses of these data are still underway; the goal of this part of the project (Gronlund, Dougherty, Durso, Canning, & Mills, in preparation) will be to understand the structure of controller plans as well as trace the evolution of planning (e.g., plan development, implementation, revision).
In analyzing the performance of the controllers in this first experiment, we also identified several sub-tasks being performed by the controllers related to their strategic planning. The two most important tasks were classifying the aircraft into groups and forming a clear plan of the sequence within each group. The classifications were needed to identify the set of aircraft that would be affected by a particular constraint such as “maintain 10 miles in trail between aircraft on jetroute 107.” The sequences were needed to determine in which order aircraft would pass a common ground point fix. These fixes were usually related to the route or destination of the aircraft (i.e., hand-off location for other sectors or TRACONs).

To perform their classification and sequencing sub-tasks for strategic planning, controllers needed to: 1) find current flight status information (position, altitude, speed), and 2) find current flight plan information (destination and intermediate points). Many of the controllers in the first experiment grouped their FPS’s by aircraft destination and then ordered them within each group by arrival time at a particular fix. The strips served as tokens for the aircraft, and the stacking order of the tokens made the sequence order explicit. We designed a software interface that maintained and enhanced this functionality (Canning, Johansson, Gronlund, Dougherty, & Mills, 1999). We then conducted a second experiment (Moertl, Canning, Johansson, Gronlund, Dougherty, & Mills, in press), the purpose of which was to measure the utility of various interface elements for the sub-tasks of classifying and sequencing aircraft.

We created two screens for the strategic controller: a simulated radar view and a flight organizer. The flight organizer display was where the classifying and sequencing took place. We represented aircraft using a small rectangle that we called an aircraft token. The aircraft tokens were placed into containers called queueing blocks. The queueing blocks represented the groups of classified aircraft. All the aircraft tokens placed in a queueing block were displayed in the same color. The corresponding aircraft symbol on the radar view appeared in the same color. Therefore, the classification was encoded using color on both displays. Controllers created and positioned as many queueing blocks as they wished on the flight organizer screen. Controllers manually adjusted the position of aircraft tokens within a queueing block or between queueing blocks.
The sequence within a queueing block was based on either the expected time or distance for each aircraft to reach a particular point in space. Time or distance to reach the fix was displayed to the right of each aircraft token. The aircraft tokens can be reordered automatically so that the lowest token is the closest to that block’s fix and the other tokens were in ascending order. We call this the automated sequencing tool and it was the function we evaluated in the experiment to follow.

The strategist’s radar display looked much like the Display System Replacement (DSR) screens used by the FAA. However, the radar display was enhanced to allow for the color coding, marking, and highlighting of the aircraft in coordination with the flight organizer display. For example, if the strategist had a queueing block for the aircraft heading to D/FW, the color of those tokens on the flight organizer corresponded to the color of those targets on the radar display. For additional details of the interface see Canning et al. (1999).

In the current experiment, the primary addition to the flight organizer involved the addition of the capability for the strategist to mark on the flight token the control actions they decided should be taken to achieve the planned sequence. Participants clicked to mark that speed, vector, and/or altitude control action(s) were needed. Alternatively, the strategist could indicate that no control action was needed for a given aircraft (i.e., it would fall into place given its current speed, altitude, and heading). (The aforementioned underscored letters appear within the flight token.) Also, rather than requiring that the strategist verbally convey the plan to the tactician, the strategist printed for the tactician a copy of the sequence order with the accompanying control actions.

In the experiment that follows, the participants again served as strategists for our SME tactician. Participants developed plans for sequencing scenarios and conveyed that plan to the tactician for implementation. For one scenario, the automated sequencing tool was disabled, for the other two scenarios it was enabled.

Method

Participants
Five en route air traffic controllers participated. All were instructors at the FAA Academy and were familiar with the AeroCenter airspace used in the experiment. All were full-performance level (FPL) controllers (i.e., certified to work a sector independently). They had been FPL controllers for an average of 20.5 years. They last worked in the field an average of 3.8 years ago.

Materials

The experiment was conducted at the RECON lab at the Civil Aeromedical Institute in Oklahoma City. The strategist sat at one end of this room, the tactician and two ghost pilots sat at the other end (the ghost pilots on either side of the tactician). The ghost pilots controlled the simulated aircraft based on the tactician's instructions. The strategist and tactician communicated over the radio, the tactician and the ghost pilots communicated over the tops of the cubicles that separated them. The tactician had a single screen on which was shown the FAA's DSR radar display. The strategist had the two display screens described above.

Four scenarios were developed with the help of our SME. The scenarios were judged by the SME to exceed the workload level typically experienced in the field. The primary problem to be solved involved the sequencing of aircraft into two different destination airports. We called these the two primary sequences. Time constraints limited us to using only three of the scenarios for a given participant.

Procedure

Each participant completed a 45-minute training session on the use of flight organizer functions. This was followed a day or two later by a 45-minute practice session with one of the trainers. The participant was asked to perform all relevant interface functions. They participated in the experiment 2-3 days later.

The strategists completed three scenarios during the course of the experiment. For two of the scenarios, participants were told that they could terminate their sequencing at a point of their choosing. For example, they could decide which aircraft was first, second, and third, but not decide about the rest of the sequence. We call this the self-terminated sequence. In the remaining scenario, participants sequence all the aircraft in the two primary sequences. We call this the exhaustive sequence. In the self-terminated
sequence condition, marking control actions for the two primary sequences was optional; in the exhaustive sequence, control actions had to be indicated for all aircraft. In one of the self-terminated sequence problems, the fix function was disabled. This meant that the automated sequencing algorithm was inoperable, which meant that a controller would have to construct their sequence “by hand”. The fix function was enabled in the other two conditions. This resulted in three conditions, self-terminated—fix enabled, self-terminated—fix disabled, and exhaustive—fix enabled. The ordering of the conditions was randomized, as was the assignment of the four scenarios to the three conditions.  

The instructions began by delineating the roles of the tactician (the SME) and the strategist (the participant). The tactician’s job was to maintain separation between aircraft; he would make whatever altitude, speed, and heading changes necessary to maintain separation. Only the tactician would communicate with the pilots. The strategist’s job was to sequence the aircraft in the two primary sequences. They also were to indicate how the sequencing should be achieved by marking whether speed, altitude, and/or vector control action(s) (or no control action) were to be used to achieve the necessary spacing and sequencing. The strategist could print a copy of the plan for the tactician anytime they wished.

The experiment began with the first scenario in paused mode. The strategist then configured the flight organizer to their liking. In particular, the strategist created at least two queuing blocks, one for each primary sequence. Aircraft going to the appropriate destinations were then placed into the designated queuing blocks and sequenced (either manually, or using the automated sequencing tool).

After the strategist had completed their plan, they pressed a button on the interface and printed two hardcopies of the planned sequence. One copy was given to the tactician, who would use this information to implement the strategist’s plan. The other copy was given to the strategist, who was asked to circle groups of aircraft that they thought of as groups. Then, for each circled group of aircraft, they rated their confidence that the aircraft in this group were in their final sequence order (1=very certain to

---

1 Time constraints forced us to drop the exhaustive—fix disabled condition. In a pilot study, plan preparation took too long in this condition. We also felt that it added little to experiment in terms of the conclusions we could draw.
5=very uncertain). Finally, they rated their confidence that the control actions marked were sufficient to establish the sequence order (again, 1 to 5).

Once the strategist had completed this process, the scenario was started. The strategists were instructed to modify their plan as necessary, and to print-out the revised plan for the tactician whenever needed. Every 2.33 minutes, a tone sounded. Both the tactician and the strategist had been instructed to rate their workload on a 7-point scale by pressing the appropriate key on the keypad in front of them (called the workload assessment keypad or WAK, Stein, 1985).

At the 10-minute mark, the scenario was stopped. The tactician completed the NASA-TLX that assessed various aspects of his workload during the preceding 10 minutes (Hart & Staveland, 1998). The strategist and the tactician then each answered five questions related to the strategist’s plan (e.g., quality of the plan, quality of revisions, strategic awareness). A 10-minute break followed. After all three scenarios were completed, the strategist completed a final questionnaire that collected biographical information, asked about the usefulness of the automated sequencing function, what equipment they used to develop and monitor their plans, etc. The participants were then debriefed and released. The experiment session took approximately 2.5 hours.

Results and Discussion

Due to the lack of participants, traditional statistical analysis of the data did not seem worthwhile. Instead, we present effect sizes in order to arrive at some tentative conclusions about what happened in the experiment. Effect size is a general measure of the degree to which the null hypothesis is false, or of the size of an effect in the general population. The effect size, \( d \), is defined mathematically as the difference between two means divided by their pooled standard deviations. That is, \( d = \frac{M_1 - M_2}{\sigma} \). By convention, effect sizes are generally judged according to the following scale: small effects = .2, medium effects = .5, large effects \( \geq .8 \) (Welkowitz, Ewen, & Cohen, 2000).

Moertl et al. (in press) found that strategists developed their plans significantly more quickly using our new interface than when using the FPS’s. In the present study, we examined this same variable and found that the presence versus absence of the automated sequencing tool also affected this variable.
For the self-terminated condition, participants planned more quickly when the automated sequencing tool was enabled than when it was absent (\(\bar{x} = 541\) s, \(s = 180.9\) vs. \(\bar{x} = 766\) s, \(s = 275.8\), effect size = .98). Access to the automated sequencing tool sped up the development of the initial plan by nearly four minutes. Not surprisingly, when the fix was present, developing an exhaustive plan took longer than developing a self-terminated plan (\(\bar{x} = 722\) s, \(s = 230.2\) vs. \(\bar{x} = 541\) s, \(s = 180.9\), effect size = .88).

We next looked at the relationships among the tactician’s TLX ratings, his ratings of the strategist’s plan quality, and the strategist’s rating of their plan quality. High values on the TLX scale indicated higher workload; higher values on the plan quality scale indicated better planning. These data are shown in Table 1. The tactician rated his workload as higher when the strategist did not have access to the automated sequencing tool.\(^2\) (The highest scores in the two conditions in which the tool was enabled did not reach the mean of the disabled condition.) Consistent with the tactician working harder when the automated sequencing tool was disabled, the tactician rated plan quality as lowest in this condition, as did the strategist (barely). Unfortunately, the tactician was not blind to the participants’ condition.

Every computer interaction of the strategist was logged. We computed the total number for each action and compared these across conditions for ideas about how the strategists’ compensated for the absence of the automated sequencing tool, how the controllers used the interface, etc. More actions were logged in the disabled condition (890) than in both fix-enabled conditions (795.5 on average). Many of these additional actions involved moving tokens, something the automated sequencing tool did for the strategist at a push of a button. Strategists made greater use of the distance-measuring tool when the automated sequencing tool was disabled. One final interesting occurrence was that the strategists printed the plan for the tactician twice as often when the automated sequencing tool was disabled, signaling more revisions of their plan (3.4 vs. 6.8 print-outs per strategist on average).

There were small differences among conditions regarding the size of the aircraft groupings circled in the two primary sequences. Consequently, we present the means collapsed across condition. On

\(^2\) Workload ratings from the WAK differed very little across conditions for either the tactician or the participant. The ratings were uninformative regarding either independent variable.
average, there were 5.2 aircraft in a group ($s = 2.7$). The majority of the control actions marked were speed (mean frequency of $\bar{x} = 13.3$) or vector ($\bar{x} = 9.4$). The number of speed or vector changes ranged from 1 to 27 and 0 to 27, respectively. However, there was a strong positive correlation between the number of speed and the number of vector control actions ($r = .48, p < .07$). Apparently, rather than a preference for a controller to set-up their sequences using speed or vectors, the positive correlation indicated that some controllers issued more of both types of control actions. A median split on the number of control actions issued showed that the tactician perceived those strategists that issued fewer control actions to have a better plan. Strategists indicated that no control actions were necessary for 5.8 aircraft on average (range 1 to 15). Altitude changes were not important for solving the sequencing problems ($\bar{x} = 0.33$). Strategists expressed high confidence in both the sequence order they presented to the tactician and the control actions selected ($\bar{x} = 1.68$, 1 to 5 scale, 1 = very certain).

**Recommendations and Conclusions**

There were consistent and beneficial effects of the automated sequencing tool. Strategic plans were developed more rapidly when participants were able to use this function. The tactician’s workload also decreased, and both the tactician and the strategist agreed that plan quality increased. The automated sequencing tool appears is a beneficial feature of the interface. Without that tool, the strategist compensated by measuring distance with the distance measuring tool and had to revise their plan more often (signaled by printing the plan more often). It also required more movement of individual tokens to reorder sequences within a queuing block.

No differences in workload or planning quality were found due to the depth of planning manipulation (i.e., self-terminating vs. exhaustive). Of course, this could be due to the small sample size, but there was little indication in the data that one type of planning was better. This independent variable was not ignored by the participants. It did result in control actions indicated for 94% of the aircraft in the exhaustive condition, but only 65.5% of the aircraft in the two self-terminated conditions. But future attempts to examine whether more complete, explicit planning is beneficial or harmful should employ a different means to manipulate depth of planning, and measure additional aspects of planning.
performance. For example, depth of planning may enhance situation awareness by forcing the controller to look further into the future.

Despite the lack of evidence for an effect of depth of planning, we think that this construct is still worth additional investigation. Planning aids like URET (MITRE, 1997) and CTAS (Vivona et al., 1996) allow controllers to make more detailed plans further into the future. The reason we included a manipulation of depth of planning in our study was because we hypothesized that more detailed plans of the sort made possible by URET might be particularly prone to cognitive rigidity (Taylor et al., 1997). Cognitive rigidity, a form of functional fixity, would occur when a controller stuck with an existing plan despite evidence (as judged by an outside observer) that the plan required modification. Cognitive rigidity can also be viewed as a form of sunk costs (Arkes & Blumer, 1985). The time and cognitive effort devoted to developing the plan would make a controller more willing to stick to an already created plan than to revise it. On the other hand, a plan you invested less time and cognitive effort constructing would be readily modified in light of new data.

Dougherty et al. (1999) argued that controllers do not create detailed plans using the current paper FPS interface. The complex and dynamic nature of their environment makes construction of detailed plans very difficult. Instead, we argued that controllers constructed hybrid plans: detailed for those aircraft in relatively close proximity to one another, sketchy for those that are farther apart from one another. For example, when sequencing aircraft to a destination airport, a controller might decide which aircraft are 1st, 2nd, and 3rd in line (the detailed part), but only determined that the 4th through 6th aircraft will come from among the next wave of incoming aircraft. They have not yet decided which one is 4th, 5th, and 6th, although they have chosen the three candidate aircraft. The cognitive effort required to decide, at this point in time, which aircraft is 4th, is great. However, tools like URET make it easier to make these decisions. On the one hand, it seems like a good idea for controllers to look further into the future. On the other hand, having more detailed plans, plans that a controller might be more invested in, could result in greater cognitive rigidity, and hence, poorer performance. Additional studies that manipulated the level of
detail in a plan, and the degree of commitment to it, should be conducted to verify that planning aids like URET actually are beneficial.

In sum, an information organization tool has been developed and tested. We showed that there is a strong strategic component to job of the en route air traffic controller, at least when faced with certain type of problems (i.e., sequencing aircraft to a destination airport). We found that a tool that provides an active linkage between tokens of flight data and their representation on a radar display, provided several benefits for a controller serving the role of a strategist. In the current workplace environment, much of this strategic planning is accomplished using paper FPS's. Our research demonstrated that planning ability can be facilitated and enhanced through the use of electronic representations of the strips on our flight organizer interface. Questionnaire responses also indicated a high degree of satisfaction by the strategist participants with the new interface.
References


Acknowledgments

We gratefully acknowledge the support of FAA grant 97-G-037. We are indebted to Ron Andrei and Paul DeBenedittis for their help on this project. Correspondence concerning this article should be addressed to Scott D. Gronlund, Department of Psychology, University of Oklahoma, Norman, OK 73019, sgronlund@ou.edu.
Table 1. Tactician's subjective workload, and tactician and strategists' rating of plan quality for the three experimental conditions. Standard deviations given in parentheses.

<table>
<thead>
<tr>
<th></th>
<th>Fix-Exhaustive</th>
<th>Fix-Self-terminated</th>
<th>NoFix-Self-terminated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tactician's TLX&lt;sup&gt;a&lt;/sup&gt;</td>
<td>51.0</td>
<td>44.9</td>
<td>63.6</td>
</tr>
<tr>
<td></td>
<td>(7.08)</td>
<td>(14.27)</td>
<td>(17.39)</td>
</tr>
<tr>
<td>Tactician’s plan quality&lt;sup&gt;b&lt;/sup&gt;</td>
<td>4.85</td>
<td>5.15</td>
<td>4.45</td>
</tr>
<tr>
<td></td>
<td>(1.49)</td>
<td>(1.95)</td>
<td>(1.72)</td>
</tr>
<tr>
<td>Strategist’s plan quality&lt;sup&gt;b&lt;/sup&gt;</td>
<td>5.65</td>
<td>5.96</td>
<td>5.52</td>
</tr>
<tr>
<td></td>
<td>(0.99)</td>
<td>(1.43)</td>
<td>(1.37)</td>
</tr>
</tbody>
</table>

<sup>a</sup> Collapsed across five subscales, performance excluded. Scale ranges from 0-100 with 100 indicating maximum workload.

<sup>b</sup> Collapsed across initial plan quality, quality of revisions, strategic awareness, and strategic planning (scale 1 to 9, 9 indicating superior planning).
