

THE IMPACT OF AUTOMATION USE ON THE MENTAL MODEL: FINDINGS FROM THE AIR TRAFFIC CONTROL DOMAIN

Ashley Nunes
Aviation Human Factors Division
University of Illinois at Urbana Champaign
Savoy, IL

The 'Free Flight' (FF) concept has been proposed as a means of reducing delay in the airspace system. However, concerns over lapses in a controller's situation awareness as a result of FF implementation have prompted an influx of new technologies aimed at helping controllers more effectively predict the future trajectories of aircraft thereby preserving situation awareness. Such aids are based on the principle of direct visualization, whereby future system states are presented without providing adequate information to the user as to how these states were generated. Usage of aids that employ this design principle can have adverse effects on the mental model given that the user is not forced to extensively think about the processes governing the prediction. Drawing on the results from an empirical study in the air traffic control domain, evidence of this concern is presented and it is argued that caution must be exercised when introducing these aids, given the possibility that such usage may compromise a controller's ability to problem-solve and acquire knowledge.

INTRODUCTION

Various categories of user compete for the finite available airspace (Hopkin, 1995). The duty of keeping such users apart while they share this airspace falls to air traffic controllers. Therefore, the primary goal of the air traffic control (ATC) service is to provide for the safe, orderly and expeditious flow of traffic through the National Airspace System (NAS). The provision of this service is dependent on an accurate 'mental model' of the ATC system, which is acquired as function of training and experience.

Mental Models & Situation Awareness in ATC

Kieras and Bovair (1984) defined a mental model as being the structure that embodies an understanding of how a system works, with respect to its internal components and processes. Gott, Lajoie & Lesgold (1991) suggested that the ability to predict future system states and problem-solve is highly dependent on the mental model. A study by Moray, Looftsteen, and Pajak (1986) found that the first step towards acquiring process control skills was to understand the nature of the process being controlled, highlighting the importance of mental models in learning process.

In the ATC domain, there has been limited research aimed at defining the mental model of the controller. Perhaps the most cohesive representation was that proposed by Redding, Ryder, Seamster, Purcell, and Cannon (1991), although it presented both the static and dynamic features in one unified model. Endsley (1995) defines the current state of the mental model as the situation model, a term used interchangeably here with situation awareness (SA). SA may be defined as being the perception of the elements in the environment (level 1), the comprehension of their meaning (level 2), and the

projection of their status in the near future (level 3). In the ATC domain, an operator's ability to maintain safe separation between aircraft is dependent on the achievement and maintenance of good SA. The controller must be able to effectively extract relevant information from the radar screen, comprehend what it means at the present time, and determine what it could mean in the future.

The Need for Automation

Recent efforts to move away from the constraints of the current ATC system have prompted much interest in advanced ATC concepts on both sides of the Atlantic; Free Flight (FF) in North America and various derivatives of the concept in Europe (Free Routing). Safety concerns however have prompted similar levels of enthusiasm being extended to the development and utilization of automation aimed at helping the controller perform their task at lower levels of workload without compromising SA. A 'predictive' aid is a form of automation assistance that reduces cognitive demand by providing the controller with a direct visual extrapolation of an aircraft's trajectory. Indeed, it is the effects of aid usage on SA that have been of most concern in past empirical research (see e.g. Endsley, Sollenberger, Nakata, and Stein, 2000).

THEORETICAL CONCERNS

Few predictive aids however, if any at all, actually provide the controller with information as to how a trajectory was actually generated, the designers of the system opting for a simple visual extrapolation. This can have serious implications for the underlying mental model. Wickens (1992) for example, has argued that those features of an interface that are based on direct visualization (as is the case with most predictive aids)

may reduce effort and increase performance, but may also reduce long-term retention of knowledge. This argument makes intuitive sense because, given that such interfaces provide a graphical representation of the process (trajectory prediction in this case), the human is not forced to think extensively and understand the processes governing the prediction, since as Christoffersen, Hunter & Vicente (1998) would argue, it is already apparent in the surface features of the interface.

In contrast, interfaces that do not present this information at the surface require greater cognitive effort to derive and comprehend system functionality and relationships, thereby inducing learning of deep structures (Christoffersen et al, 1998) in the underlying mental model. Therefore, it may be possible that while SA is adequate at a very superficial level when using predictive aids, the underlying mental model is being adversely affected because the controller does not take the time to reflect upon the information being provided by the aid nor does he/she comprehend the means by which that information was derived. This can have serious implications for problem-solving skill (Lee and Moray, 1994), long-term knowledge acquisition and retention. Christoffersen et al (1998) stated that Wickens (1992) was not alone in his claim, citing other researchers who have made similar suppositions (e.g. Pea, 1993; Salomon, 1993). While the claim has been supported by the results from one previous study (Vicente, 1991), empirical evidence of it has not been seen as it relates to controller performance. This paper will make a novel contribution therefore by presenting empirical evidence in the ATC domain that support Pea (1993), Salomon (1993) and Wickens' (1992) claim.

Current Study

Nunes (2002) conducted a study to assess the impact of a predictive aid on controller performance in a simulated FF environment. The predictive aid represented a form of data-link whereby pilot intent for a vertical deviation from the present trajectory could be communicated to the controller. This intent was depicted to the controller via an extrapolation of the aircraft's current lateral trajectory on a secondary screen (Figure 1).

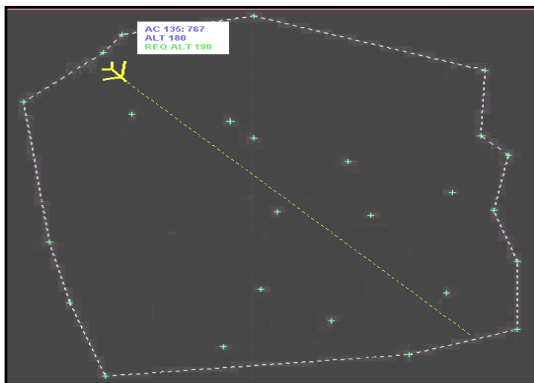


Figure 1

Twenty ATC trainees, recruited from Embry Riddle Aeronautical University, participated in the study, their ages

ranging from 20 to 25 (mean=21.4). These participants were assigned to one of two display groups, with ten subjects in each group. An incomplete within-subjects design was used for the study and two independent variables were manipulated. They were predictive aid presence (unaided and aided) and airspace load (low and high). Display presence served as the between subjects variable and airspace load as the within subjects variable. Dependent variables in the study included a controller's overall problem-solving ability, SA and workload. Problem-solving ability was measured using controller reaction time and accuracy data that was provided in response to pilot requests. SA was measured using the Situation Awareness Global Assessment Technique (SAGAT), and queries on the probe pertained to aircraft callsigns, location, altitude, heading and convergence likelihood. These queries were specifically selected given their pivotal importance in helping controllers determine the precise spatial location of aircraft in space (see e.g. Leplat & Bisseret, 1966). Finally, workload was assessed using the NASA-TLX questionnaire. Collectively, these measures were utilized to provide an accurate assessment of overall controller performance.

Task & Procedure

Prior to testing, trainees were made familiar with the protocols of the experiment and given a short briefing on the study and the nature of the task to be performed. At this time, subjects in the 'predictive aid' condition were also familiarized with the aid, including the layout of information on the secondary screen and its associated meaning. Following this, all subjects participated in a practice run, which lasted approximately 20 minutes and was designed to provide them with adequate familiarization with the task to be performed.

The study then commenced with participants viewed a one-hour ATC scenario, which was representative of a FF environment where aircraft utilized direct routes of travel through a sector. The first half of the scenario represented low airspace load and the second half, high airspace load. The controller's primary task was to evaluate a pilot's request for new altitude, determine the likelihood of a conflict and respond accordingly. Three types of altitude requests could be filed with the controller, each one varying in complexity level. At the simplest level, controllers could either grant or deny the request. Correct responses to the most complex request however, required controllers to climb or descend aircraft to intermediate altitudes and then to the target altitude, in order to avoid a conflict with over-flying traffic. In the 'no aid' condition, subjects had to extrapolate trajectories using only using the radar screen whereas in the 'aid' condition, a predictive aid was also provided which presented the controller with an extrapolation of an aircraft's future lateral trajectory. This extrapolation was only provided when pilot requests were made and no further information was provided. Therefore even those controllers that used the aid had to arrive at their own decision in response to pilot requests.

A total of eighteen requests were filed by the pseudo pilots in the study, nine under low load and nine under high load and

the accuracy of controller responses as well as the time taken to provide the response were carefully recorded. Approximately 20 minutes into the study (under low airspace load), the simulation was paused and the SAGAT probe was administered followed by the NASA-TLX questionnaire. Following completion of these questionnaires, the simulation resumed and approximately 20 minutes into the second half hour of the study (under high airspace load), the procedure was repeated. At the end of the experiment, controller comments were elicited regarding strategies employed to deal with these requests and compensation was provided.

Results

A Multivariate Analysis of Variance (MANOVA) revealed that SAGAT scores got worse with increasing airspace load ($F(6,13) = 47.32, p < .00$) for all SA queries (callsign, altitude, heading etc). However, the main effect for display type was non-significant ($F(6,13) = 0.51, p > .79$) as was the interaction ($F(6,13)=0.53, p > .77$). An ANOVA performed on overall response time data revealed a significant interaction between airspace load and display type ($F(1,18)=34.75, p < .00$), suggesting that the temporal costs of dealing with requests under high airspace load were greatly amplified when no aid was present. The same interaction was also seen when response time was broken down by request complexity, specifically for the most complex request ($F(1,18)=16.5, p < .00$) (Figure 2). Analysis of response accuracy scores revealed a significant main effect for display conditions ($F(1,18) = 7.95, p < .05$) (Figure 3), suggesting the assessments made by controllers using the aid were much more accurate than their unaided counterparts. Despite this overall benefit to using the aid, preliminary results reported by Nunes and Matthews (2002) suggested that the ‘aided’ controllers’ response accuracy scores for the most complex request actually decreased with increasing airspace load (Figure 4). At the time the authors could not provide a robust explanation for this effect. Subsequently, to better understand the why these scores may have fallen, response accuracy data was subjected to additional 3-way ANOVA to isolate any possible significant interactions between aid usage and the time spent controlling traffic. The attained interaction (Figure 5) was marginally significant ($F(2,17)=2.24, p < .13$), suggesting that the response accuracy cost of dealing with the most complex requests over time, was amplified when the aid was present.

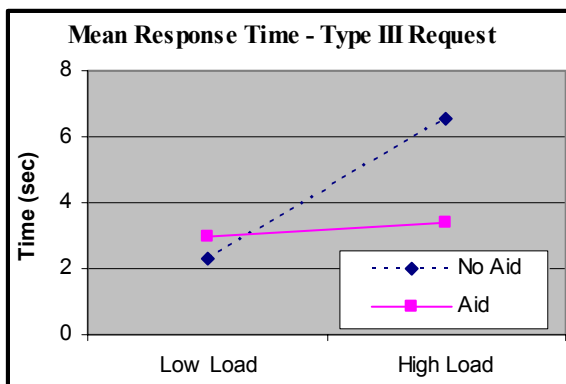


Figure 2

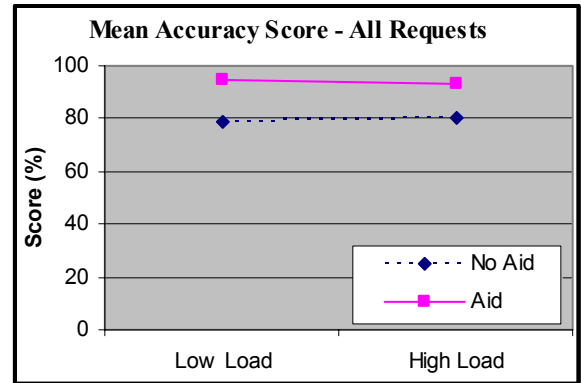
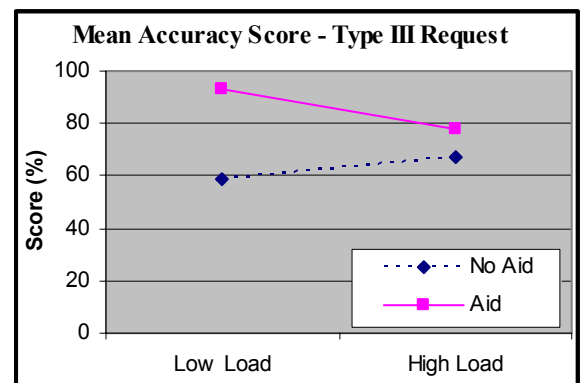


Figure 3



Source: Nunes & Matthews, 2002
Figure 4

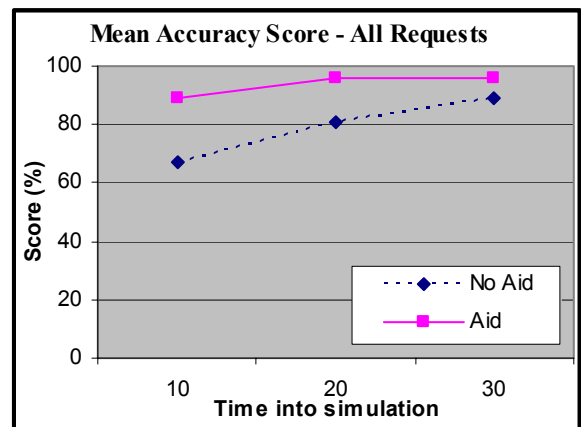


Figure 5

DISCUSSION

The results from this study are quite intriguing. SA as measured by the SAGAT did not vary between display conditions, suggesting that use of the predictive aid did not adversely affect controller SA. This may be attributed to the fact that the aid only presented trajectory information when the pilot was making requests. When no request was made, ‘aided’ controllers still had to project trajectories for themselves by perceiving, comprehending and projecting information present in the environment, thereby keeping them ‘in-the-loop.’ Therefore, there is little reason to doubt that the

SAGAT data between display conditions was inaccurate.

The overall response accuracy data for all requests combined would suggest that controllers using the aid were however, able to arrive at more accurate responses compared to their baseline counterparts. These responses were also provided in a reduced amount of time that was not affected by increases in airspace load. These effects highlight two primary benefits of using predictive aids in ATC; namely, the fact that the effectiveness of aid usage is independent of airspace load and that the usual, limited predictive capability of the controller is enhanced as a result of using the aid (see e.g. Hopkin, 1995). The results from the present study also clearly show that the response times of unaided controllers, went up significantly as airspace load increased. This makes sense given the increased number of trajectory predictions that have to be made when evaluating pilot requests. However, the marginally significant time interval / display condition effect suggests that baseline controllers were better able to first understand the problem (i.e. the nature of the request being made), predict the outcome and then problem-solve; tasks that draw on information from the underlying mental model. This would explain why their overall scores increased as time progressed. The same cannot be said for 'aided' controllers, who given the provision of the aid, were not required to develop such strategies, and may not have developed them. This would explain why their scores remained fairly constant (Figure 5). This inability to formulate effective strategies may also explain why their scores for the most complex type of request fell over time versus their counterparts whose score actually increased. This would suggest that even though SA did not differ between both groups, the problem-solving ability exhibited by the 'non-aided' group was superior to the 'aided' group when dealing with requests that had the highest cognitive processing costs associated with it.

The presence of these effects for controllers using the aid may be explained by the fact that the predictive aid only provided the controller with a direct visualization of the extrapolation process. Consequently, these controllers may have only had a shallow understanding of the prediction because the aid did not force them to extensively think about the processes governing the prediction, a point that Wickens (1992) has made before, citing concerns over long-term learning and retention. The SAGAT would not be sensitive to this reduced level of information processing because it taps a controller's perception, comprehension and projection of information in specific instances, rather than their general problem-solving and learning ability, which is highly dependent on the underlying mental model. Therefore, while SA may not have varied between groups because both aided and unaided controllers processed the same information, it is the *level* at which this processing occurred that ultimately affected a controller's response accuracy scores. Unaided controllers processed information at a greater level of depth for complex problems, which resulted in a high score being attained (for more on levels of processing, see Craik & Lockhart, 1972).

It is worth noting however that efforts aimed towards designing interfaces that are based on the principle of direct visualization, particularly in ATC have been grounded in the notion of reducing cognitive demand for the controller, but as previously suggested, following this thought process results in the human not being forced to extensively think about the processes governing the system and as a result, the human may not completely understand the system's functionalities and relationships. It is postulated that detriments to a person's mental model as a result of using such aids, may compromise their ability to problem-solve and learn, given the dependency that these two skills have on the mental model (cf. Gott, Lajoie & Lesgold, 1991; Moray, Lootsteen, & Pajak, 1986) and that the 'lack' of learning and reduced problem solving ability seen in the present study provides evidence in support of the notion that the nature of the interface used can hurt the underlying mental model even though SA is not affected.

Recommendations

Based on the results from the present study, the following set of recommendations are provided as a means of tackling the issue of mental model concerns that arise from automation usage. Efforts to address this issue should include:

- An active attempt to develop measures of the underlying mental model and utilize these measures when evaluating the effects of new technologies on human performance.
- Development of tools that reduce cognitive load but not at the expense of reduced information processing. This development should also be coupled by training aimed at making the users not only aware of how to use the tool, but the manner in which the tool arrives at a solution; i.e. the logic behind the tool.
- Finally, attempts to make the users of such tools aware of the pitfalls of automation usage, particularly when this usage occurs over a prolonged period of time. In doing so, users would be aware of these shortcomings and may therefore make a more active attempt to ensure that the level at which information is processed does not change.

CONCLUSION

The traditional need to design interfaces based on direct visualization have been grounded in the notion of reducing cognitive demand for the user. However, many researchers (Pea, 1993; Salomon, 1993; Wickens, 1992) have argued against this approach citing concerns over the detrimental effects that such aids can have on the underlying mental model, which in turn can compromise an operator's problem-solving and learning ability. Results from the present study certainly support this view. Therefore, the goal for system designers must be to strike a balance between an interface that both, reduces cognitive effort and at the same time forces the human to extensively think at a functional level about the processes being displayed by the aid. Given the current trend to utilize automation in the ATC domain, it is imperative that

this recommendation be considered, for the cost of not doing so in the event of automation failure, will be measured not monetarily, but in the number of human lives lost.

ACKNOWLEDGEMENTS

The author gratefully acknowledges Olivier St-Cyr, University of Toronto, Michael L. Matthews, University of Guelph, Canada and Christopher D. Wickens, University of Illinois for their insight and discussion of the ideas discussed herein. Funding for this research was provided by the J. Armand Bombardier Internationalist Fellowship program administered by the Canadian Bureau for International Education (CBIE) and the Natural Science and Engineering Research Council of Canada (NSERC).

REFERENCES

- Christoffersen, K., Hunter, C. N., & Vicente, K. J. (1998). A longitudinal study of the impact of ecological interface design on deep knowledge. *International Journal of Human-Computer Studies*, 48(6), pp. 729-762.
- Craik, F.I.M & Lockhart, R.S. (1972). Levels of Processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behaviour*, 11, pp. 671-684.
- Endsley, M. R. (1995). Towards a theory of situation awareness. *Human Factors*, 37(1), pp. 32-64.
- Endsley, M. R., Sollenberger, R. L., Nakata, A., & Stein, E. (2000). *Situation awareness in air traffic control: Enhanced display for advanced operations* (Report DOT/FAA/CT-TN00/01). Atlantic City International Airport, NJ: U.S. Department of Transportation
- Gott, S. P., Lajoie, S. P., & Lesgold, A. (1991). Problem solving in technical domains: How mental models and metacognition affect performance. In R. F. Dillon (ed.), *Instruction: Theoretical and applied perspectives*. Praeger Publishers, pp. 107-117.
- Hopkin, V. D. (1995). *Human Factors in Air Traffic Control*. London and Bristol, PA, Taylor & Francis.
- Kieras, D. E., & Bovair, S. (1984). The role of a mental model in learning to operate a device. *Cognitive Science*, 8(3), pp. 255-273.
- Leplat, J. & Bisseret, A. (1966). Analysis of the processes involved in the treatment of information by the air traffic controller. *The Controller*, 1(1), 13-22.
- Moray, N., Lootsteen, P., & Pajak, J. (1986). Acquisition of process control skills. *IEEE Transactions on Systems, Man, and Cybernetics*, 16(4), pp. 497-504.
- Nunes, A. (2002). An assessment of the effects of a conflict detection aid on controller performance in an advanced air traffic control environment. *Unpublished MA Thesis*. University of Guelph, Department of Psychology.
- Nunes, A and St-Cyr, O (2003). Predictive Aids & Mental Models under Free Flight: Proceed with Caution. Proceedings of the 12th Biennial International Symposium on Aviation Psychology. April 14-17, 2003, Dayton, OH: USA.
- Pea, R.D. (1993). Practices of distributed intelligence and designs for education. In G. Salomon, Ed. *Distributed Cognitions: Psychological and Educational Considerations*, pp. 47-87, Cambridge, UK: Cambridge University Press.
- Redding, R. E., Ryder, J. M., Seamster, T. L., Purcell, J. A., & Cannon, J. R. (1991). *Cognitive task analysis of en route air traffic control: Model extension and validation*. Report to the Federal Aviation Administration, McLean, VA, Human Technology Inc.
- Salomon, G. (1993). No distribution without individuals' cognition: a dynamic interactional view. In G. Salomon, Ed. *Distributed Cognitions: Psychological and Educational Considerations*, pp. 111-138. Cambridge, UK: Cambridge University Press.
- Vicente, K. J. (1991). Supporting knowledge-based behavior through ecological interface design. Unpublished Doctoral Dissertation, University of Illinois at Urbana Champaign, Urbana, IL, USA.
- Wickens, C. D. (1992). Virtual reality and education. *Processing of the 1992 IEEE International Conference on Systems, Man, and Cybernetics*, pp. 842-847.