

## **Flight Deck Automation and Task Management**

**Ken Funk**  
**Candy Suroteguh**

Dept. of Industrial and Manufacturing Engineering  
118 Covell Hall  
Oregon State University  
Corvallis, OR, USA 97331-2407  
funkk@enr.orst.edu

**Jennifer Wilson**  
**Beth Lyall**

Research Integrations, Inc.  
PO Box 25405  
Tempe, AZ, USA, 85285-5405  
jennifer.wilson@ResearchIntegrations.com

### **INTRODUCTION**

The last two decades have brought significant change to the commercial transport aircraft flight deck. Complex functions once performed by human pilots are now routinely performed by automated systems. Autoflight systems control aircraft altitude, heading, and speed as selected by the pilots. Flight management systems (FMSs) allow the human flightcrew to pre-program flight paths, complete with altitude and speed restrictions, then couple through the autoflight systems to fly those paths accurately and economically.

Flight deck automation has generally been well received by pilots and the aviation industry. There is little doubt that the addition of flight deck automation has made significant contributions to the safety and efficiency of operations, and accident records tend to corroborate this point: the hull loss accident rates for advanced technology (more automated) aircraft are generally lower than those for comparable traditional technology (less automated) aircraft [2].

Yet several accidents involving advanced technology aircraft have called some of these perceived benefits into question. For example, on April 14, 1990, an Indian Airlines Airbus A320 aircraft crashed just short of the runway at the Bangalore, India airport destroying the aircraft and killing 90 people on board. The investigators determined that the probable cause of the accident was the failure of the pilots to realize the gravity of the situation and immediately apply thrust. The pilots spent the final seconds of the flight trying to understand why the autoflight system was in idle/open descent mode rather than taking appropriate action to avoid impact with the ground [8].

The reports from the investigation of this and other recent advanced technology aircraft accidents suggest that besides some of the obvious problems with the design of human-machine interfaces, these aircraft may pose substantial challenges to the flightcrew as they perform a crucial process we call task management. Task management is the process by which operators of complex systems prioritize and perform the multiple, concurrent tasks that compete for their attention. In task management the operator chooses (consciously or automatically) what task(s) to attend to. The concept is closely related to attention allocation.

The purpose of this paper is to show that recent flight deck automation human factors research suggests that attention allocation or task management is a critical safety issue in advanced technology aircraft, to relate that finding to task management research, and to suggest a course for future research to address that issue.

### **FLIGHT DECK AUTOMATION ISSUES**

With the advent of advanced technology aircraft and the transfer of safety-critical functions away from human awareness and control, pilots, scientists, and aviation safety experts have expressed reservations about flight deck automation. Investigators like Wiener [14], Wise and his colleagues [16], Billings [1], and Sarter and Woods [12], as well as a human factors team chartered by the US Federal Aviation Administration [5], have identified a variety of automation issues related to such things as the design of automation interfaces, the complexity of automated systems, pilots' lack of understanding of automation, the possibility that automation actually increases rather than reduces pilot workload, and the tendency for automated systems to distract pilots from safety-critical flight control tasks.

Yet until recently there did not exist a comprehensive list of such issues. This prevented a full understanding of flight deck automation issues and a coordinated effort to address those issues using limited research, development, manufacturing, operational, and regulatory resources. To address that lack, we conducted a more comprehensive review of flight deck issues. This study is described briefly below and more fully in a paper [6] and a website (<http://flightdeck.ie.orst.edu/>).

## Objectives

The objectives of this automation issues study were to:

1. develop a comprehensive list of flight deck automation human factors issues;
2. compile a large body of existing data and other evidence related to those issues;
3. disseminate the list of issues and supporting data to the aviation research, development, manufacturing, operational, and regulatory communities.

## Methodology

Our methodology involved two phases. In Phase 1 we compiled a list of possible problems with, or concerns about, flight deck automation, as expressed by pilots, scientists, engineers, and flight safety experts. We reviewed 960 source documents, including papers and articles from the scientific literature as well as the trade and popular press, accident reports, incident reports, questionnaires filled out by pilots and others, and documentation from our own analyses, recording citations of problems and concerns. In Phase 1 we did not attempt to substantiate the claims made about automation problems. Rather, we merely identified and recorded citations of people's perceptions of problems and their concerns about automation as a prelude to our Phase 2 work.

In Phase 2 we located and recorded evidence related to the possible problems and concerns identified in Phase 1 from a wide variety of sources. Since an *issue* is "[a] point of discussion, debate, or dispute ..." [9], in the following we refer to these possible problems and concerns as *flight deck automation issues* or, just *issues*, except where referring to the results of Phase 1, where we refer to them as problems and concerns.

The more than 100 sources we reviewed for evidence included accident reports, documents describing incident report studies, and documents describing scientific experiments, surveys and other studies. We also conducted a survey of automation experts, individuals

with broad knowledge related to human factors and flight deck automation. We reviewed these sources for data and other objective information related to an issue. For each instance of this evidence we assessed the extent to which it supported one side of the issue or the other, and assigned a numeric strength rating between -5 and +5. We assigned a positive strength rating to evidence supporting that side of the issue claiming that a problem truly exists (supportive evidence) and a negative strength rating to evidence supporting the other side (contradictory evidence).

For example, consider the statement of the workload issue alluded to above, *issue079: Automation may increase overall pilot workload, or increase pilot workload at high workload times and reduce pilot workload at low workload times, possibly resulting in excess workload and/or boredom*. If we found evidence in a source indicating that pilot workload is increased by automation (at least under some circumstances), we recorded an excerpt from the source document and assigned this supportive evidence a positive rating, perhaps as great as +5. If we found evidence in a source indicating that no such problem exists (at least under some circumstances), we recorded an excerpt and assigned this contradictory evidence a negative rating, perhaps as great as -5.

We developed detailed strength assignment guidelines for evidence from each type of information source. For example, in pilot surveys of automation issues, if at least 90 per cent of the respondents were in agreement with a statement consistent with an issue statement, we assigned a strength of +5. If at least 90 per cent were reported as agreeing with a statement contradictory to an issue statement, we assigned a strength of -5. We developed similar guidelines for each type of evidence source.

For each instance of evidence found, we recorded in a database the related issue, an excerpt from the source document describing the evidence, source document reference information, the type of aircraft and equipment to which the evidence applied (if specified), and a strength rating. During the process of collecting and recording evidence, we revised, updated, consolidated, and organized the issues, yielding, at the time of writing, 92 flight deck automation issues. For each one, we compiled supportive and contradictory evidence related to it in preparation for dissemination.

We also performed a meta-analysis of this data to summarize the evidence in order to identify those issues that are problems in need of solutions, those issues that

do not appear to represent problems, and those issues which require more research. We ranked the issues based on several criteria, including the number of citations (from Phase 1), the extent to which the experts in our survey agreed that the issue is truly a problem, their rating of the significance of that problem, and the sum of strength assignments made in our reviews.

## Results

In Phase 1, we found 2,428 specific citations of 114 possible flight deck automation problems and concerns. In Phase 2 we identified and recorded more than 700 pieces of evidence for 92 distinct automation issues. The issue ranking highest in supportive evidence alone was *issue105: Pilots may not understand the structure and function of automation or the interaction of automation devices well enough to safely perform their duties*. (Issue numbers greater than 92 are possible because more than 92 issues were identified, but some were eliminated or combined with others.) Other issues ranking high in supportive evidence were related to automation behavior that surprises pilots and automation-induced complacency.

We found primarily contradictory evidence for several issues. In particular, *issue079* (automation may adversely affect pilot workload, see above) ranked second in contradictory evidence despite its rank of eighth in number of Phase 1 citations.

A particularly noteworthy issue was *issue102: The attentional demands of pilot-automation interaction may significantly interfere with performance of safety-critical tasks*. (e.g., "head-down time", distractions, etc.). By virtue of its focus on how automation may influence the allocation of attention among tasks, this issue is clearly related to task management. In the meta-analysis it ranked second with respect to the extent that our automation experts agreed that it is truly a problem and it ranked first in number of Phase 1 citations. It also ranked in the top 20 of all rankings based on other criteria.

## Discussion

In general, we consider those issues with the greatest overall supportive evidence (such as *issue105*, above) and especially those issues ranking highest in multiple criteria (e.g., *issue102*, above) to indicate real problems which require solutions, and resources should be dedicated to finding those solutions. Along those same lines, we consider that those issues with the greatest overall contradictory evidence are not significant problems, and resources would be better used in solving real problems or further exploring unresolved issues,

those ranking high in neither supportive nor contradictory evidence.

In particular, we find the contrast between *issue102* (attention), which ranked high in multiple criteria, and *issue079* (workload), which ranked high in contradictory evidence, interesting. In Phase 2, we found that there is strong evidence and agreement that automation can and often does draw pilot attention away from safety-critical flight control tasks. On the other hand, despite the common linking of workload to attention, we also found that there is pretty strong evidence that automation does not significantly increase workload, *per se*. Clearly, the solution to the attention problem will require further investigation.

## AUTOMATION AND TASK MANAGEMENT

Coincidentally, before our meta-analysis had revealed *issue102* as so significant, we had already begun a study investigating it [15]. Results from earlier studies [3] suggested that task management (related to attention allocation, as described above) may be particularly problematic in advanced technology, highly automated aircraft. We therefore conducted a study of aircraft incident reports to determine if the level of automation was a significant factor in errors related to the proper management of flight deck tasks.

### Objective

The primary objective of the study was to begin evaluating the relationship between task management of commercial airline pilots and the level of automation on the flight deck by determining how automation affects the frequency of task prioritization errors as reported in Aviation Safety Reporting System (ASRS) incident reports.

### Methodology

We compared two samples of ASRS incident reports to determine if level of automation on the commercial aircraft flight deck affected the frequency of task prioritization errors. The first sample was composed of 210 incident reports submitted by pilots flying advanced technology aircraft and the second sample was composed of 210 incident reports submitted by pilots flying traditional technology aircraft. In total, we analyzed 420 incident reports.

To prevent confounding the level of automation with differences in experience level because the advanced aircraft are comparatively new, we divided the two samples into three sub-samples each made up of 70

reports submitted during specified time periods: 1988-1989, 1990-1991, and 1992-1993.

In an effort to collect homogenous samples, we constrained the sample populations so that the level of automation and the submission period were the only two differences between the samples. For example, all the reports from both the advanced technology and the traditional technology samples were constrained to reports submitted by a member of the flightcrew flying a two-person commercial air carrier aircraft in which the aircraft was classified as a medium-large transport, large transport or widebody transport aircraft. All reports were for incidents that occurred during the descent or approach phases of flight.

We used an incident analysis form developed specifically for this project. This form allowed us to classify the ASRS incident reports as either containing a task prioritization error or not, based on the description given in the narrative of the report. Using the form, we identified the tasks that were being performed during the incident period reported. Prioritization was evaluated by identifying whether these tasks were related to the task categories of aviate (i.e., flight control, the highest priority tasks), navigate, communicate, manage systems, or non-flight related tasks (lowest priority tasks). If a task of lower priority was active (i.e., being actively performed) while a task of higher priority was not active, the report was classified as containing a task prioritization error.

For example, consider ASRS incident report #92507, describing an incident in which the flightcrew allowed the aircraft to deviate from their assigned altitude. The narrative of this report, in which the reporter (the captain, in this case) describes the incident, contains the following. The terse text of the original narrative (in all capitals, as recorded in the ASRS database) has been elaborated by the authors for purposes of clarification.

THE [first officer] WAS FLYING THE [aircraft]. WE HAD BEEN ISSUED SEVERAL [headings] AND TURNS BY [Air Traffic Control (ATC)] FOR FLOW [control] INTO CHICAGO O'HARE. I WAS ON THE [public address system] EXPLAINING THE ... DELAY TO THE [passengers] WHEN I NOTICED THE ... [autoflight system's target altitude] HAD CHANGED FROM 35000 [feet] TO 33000'. I ASKED THE [first officer] IF WE HAD BEEN [cleared by ATC] TO [33000 feet]. HE SAID NO. THE [aircraft

altitude] WAS 34600' WHEN I NOTICED THE [problem]. ...

In this incident the flightcrew members were actively performing communication tasks (with passengers and between themselves) while not actively monitoring the task of maintaining their assigned altitude. Since the latter should have a higher priority than the former ones to insure safety of flight (based on the task prioritization strategy discussed above), we classified this report as containing a task prioritization error.

We analyzed each incident report using the incident report analysis form described above. To minimize bias during the analysis, the two samples (including the three sub-samples within each) were randomly mixed and the sample to which each incident report belonged was not specified until all analyses were complete. After all reports had been analyzed, we sorted the reports and summarized the data.

**Results**

Of the 420 incidents reports analyzed, we classified 43 (10.2%) as containing task prioritization errors. Of these, 28 were from the advanced technology sample and 15 were from the traditional technology sample (Table 1).

We used the Chi Square ( $\chi^2$ ) test to determine if the difference between 28 Task Prioritization errors found in advanced technology incident reports and the 15 task prioritization errors in traditional technology aircraft was statistically significant. The  $\chi^2$  value calculated was 4.379 at 1 degree of freedom with a p value of 0.036. Using a significance level of  $\alpha = 0.05$ , we concluded that this difference was statistically significant.

**Table 1.** Summary of the frequencies of Task Prioritization errors.

Submission Period	Task Prioritization Error Frequency		Total Errors by Submission Period
	Advanced Technology	Traditional Technology	
1988-1989	13	7	20
1990-1991	11	5	16
1992-1993	4	3	7
<b>Total Errors by Aircraft Technology</b>	<b>28</b>	<b>15</b>	

We divided the two samples into three sub-samples, each made up of 70 reports submitted during a specified time period: 1988-1989, 1990-1991, and 1992-1993, combining the data for each submission period from both the advanced technology and the traditional technology aircraft. We used the  $\chi^2$  test to determine if the differences between the submission periods were significant. The  $\chi^2$  value was 6.891 at 2 degrees of freedom with a p-value of 0.032. Using a significance level of  $\alpha = 0.05$ , we concluded that this difference was statistically significant. We also found a significant difference between reporting periods for the advanced technology aircraft reports alone, but there was no such period difference in the traditional technology data.

### Discussion

We found that task prioritization errors occurred in both advanced technology and traditional technology aircraft, and that overall there was a statistically significant difference between the number of reports classified as containing a task prioritization error in the advanced and traditional technology aircraft. This difference suggests that task management may be more difficult in advanced technology aircraft.

When evaluating the results of this study, one must of course bear in mind the limitations of ASRS incident report data. The samples used in this study were drawn from a nonrandom sample of events occurring in aviation operations and the ASRS incident reports themselves reflect reporting biases. What we can say with confidence however, is that task prioritization errors do exist in actual line operations and their existence warrants consideration.

### UNDERSTANDING AND IMPROVING TASK MANAGEMENT IN ADVANCED TECHNOLOGY AIRCRAFT

The incident study just described tends to confirm what our automation issues study asserted, that automation can and does divert pilot attention from safety-critical flight control tasks. What it does not do is tell us why this is so. What is it about automation or about how pilots think about or use automation that distracts them, sometimes with fatal consequences? Why do task management errors occur, especially in advanced technology aircraft? Though our automation issues study identified evidence that automation interface design may be a culprit, there is as yet no satisfactory empirical data on which to base design or operational recommendations to improve task management performance in advanced technology aircraft. But findings from this and other research suggest factors that may affect task management

performance, and knowing the existence and magnitude of those effects will go a long way towards offering design guidelines and recommendations for procedures.

For example, in a part-task simulator study of the effect of workload level on task management performance, Chou and his colleagues observed that the focus of attention was strongly related to the salience of stimuli. In other words, tasks with conspicuous stimuli, such as malfunction warnings, tended to distract subjects from even higher priority flight control tasks [3]. This seems germane to the present issue, for the stimuli produced by automation interfaces are often salient and compelling. Results from our incident study suggest that automation proficiency may be an important factor. The decrease in task prioritization error rates in advanced technology incident reports over time may be due to the fact that pilots are getting more proficient with automation. Others studying flightcrew behavior have suggested factors affecting task management as well: Rogers [11], Latorella [7], Damos [4], and Schutte and Trujillo [13].

Another possible source for factors influencing task management is the attention psychology literature. Unfortunately, most studies done by psychologists in this area have involved subjects performing extremely simple stimulus-response tasks. The "ecological validity" of their findings and the potential for generalizing to the flight deck are questionable. Nevertheless, their results do in fact suggest factors worth considering. Pashler [10] provides a good overview.

Table 2 summarizes the factors we have identified that may be significant in influencing task management performance. In the table, references are given for those factors which are suggested by the literature cited above. Those factors marked with an '\*' are those we have not yet found in the literature but are suggested by our own research. Many of the factors are particularly relevant to automation system design, though they may play a role on traditional technology flight decks, as well as in other complex human-machine systems.

We must emphasize that this list of possible factors reflects speculation based on our interpretation of the literature and on informal observations we made as we conducted our own research. To formally investigate the factors, we are now developing part-task simulator experiments to study the effects of some of them and we hope that we will be joined by others seeking a better understanding of task management behavior. This understanding is essential if we are to address the task management problems posed by flight deck automation.

**Table 2.** Factors that may affect task management performance. '\*'s indicate factors suggested by our research (see text).

Factors	References
advance knowledge of upcoming tasks	[10]
discriminability of task-related stimuli	[10]
differences in level of effort required to process task-related stimuli	[10]
temporal proximity of task-related stimuli	[10]
task importance: aviate > navigate > communication > manage systems	[13]
perceived urgency of task (time remaining vs. time to complete)	*
task difficulty	*
(automation) task proficiency	*
task recency	*
task momentum: tendency to continue to perform the current task	[13]
task proximity to completion	*
amount of effort already invested in tasks	*
perceived task status (satisfactory, unsatisfactory)	*

#### ACKNOWLEDGEMENTS

Part of this research was funded by the Federal Aviation Administration Office of the Chief Scientific and Technical Advisor for Human Factors. John Zalenchak, Tom McCloy, and Eleana Edens were our technical monitors and we are grateful for their support. Opinions expressed in this paper are those of the authors and do not necessarily reflect the views of their employers (past or present) or of the Federal Aviation Administration.

#### REFERENCES

- [1] Billings, C.E., *Aviation automation*. Mahwah, NJ: Lawrence Erlbaum Associates, 1997.
- [2] Boeing Commercial Airplane Group, *Statistical summary of commercial jet airplane accidents, worldwide operations, 1959-1996*. Seattle, WA: Boeing Commercial Airplane Group, Airplane Safety Engineering, 1997.
- [3] Chou, C., Madhavan, D., & Funk, K., "Studies of cockpit task management errors". *International Journal of Aviation Psychology* 6(4), 1995, pp. 307-320.
- [4] Damos, D., *The Effect of Operational Interruptions on Flight Crew Performance*. Los Angeles, CA: Damos Research Associates, Inc., 1997.
- [5] FAA Human Factors Team, *The interfaces between flightcrews and modern flight deck systems*. Washington: Federal Aviation Administration, 1996.
- [6] Funk, K., Lyall, B., Wilson, J., Vint, R., Niemczyk, M., Sureteguh, C., and Owen, G., "Flight Deck Automation Issues," *International Journal of Aviation Psychology*, in press.
- [7] Latorella, K.A., "Investigating Interruptions: An Example from the Flightdeck." in *Proceedings of the Human Factors and Ergonomics Society 40<sup>th</sup> Annual Meeting*, Philadelphia, PA, 1996, pp. 249-256.
- [8] Ministry of Civil Aviation, India. *Report on accident to Indian Airlines Airbus A-320 Aircraft VT-EPN at Bangalore on 14<sup>th</sup> February 1990*. Government of India, Ministry of Civil Aviation, 1990.
- [9] Morris, W. (Ed.), *The American Heritage Dictionary of the English Language*. Boston: Houghton Mifflin, 1969.
- [10] Pashler, H.E., *The Psychology of Attention*. Cambridge, MA: The MIT Press, 1998.
- [11] Rogers, W. H., "Flight Deck Task Management: A Cognitive Engineering Analysis." in *Proceedings of the Human Factors and Ergonomics Society 40<sup>th</sup> Annual Meeting*, Philadelphia, PA, 1996, pp. 239-243.
- [12] Sarter, N.B., Woods, D.D., "'How in the world did we ever get into that mode?' Mode error awareness in supervisory control". *Human Factors* 31(1), 1995, pp. 5-19.
- [13] Schutte, P.C., and Trujillo, A.C., "Flight Crew Task Management in non-Normal Situations." in *Proceedings of the Human Factors and Ergonomics Society 40<sup>th</sup> Annual Meeting*, Philadelphia, PA, 1996, pp. 244-248.
- [14] Wiener, E.L., *Human factors of advanced technology ("glass cockpit") transport aircraft* (NASA CR 177528). Moffet Field, CA: NASA Ames Research Center, 1989.
- [15] Wilson, J.R. and Funk, K., "The Effect of Automation on the Frequency of Task Prioritization Errors on Commercial Aircraft Flight Decks: An ASRS Incident Report Study," in *Proceedings of the Second Workshop on Human Error, Safety, and System Development*, Seattle, Washington, USA, 1-2 April 1998, pp. 6-16.
- [16] Wise, J.A., Abbott, D.W., Tilden, D., Dyck, J.L., Guide, P.C., & Ryan, L., *Automation in corporate aviation: Human factors issues* (CAAR-15406-93-1). Daytona Beach, FL: Center for Aviation/Aerospace Research, Embry-Riddle Aeronautical University, 1993.