②

AD-A254 127

Report No. NAWCADWAR-92036-60

ADAPTIVE AUTOMATION AND HUMAN PERFORMANCE: II. EFFECTS OF SHIFTS IN THE LEVEL OF AUTOMATION ON OPERATOR PERFORMANCE

Raja Parasuraman, Toufik Bahri, Robert Molloy, and Indramani Singh
Cognitive Science Laboratory
THE CATHOLIC UNIVERSITY OF AMERICA
Washington, DC 20064

MARCH 1992

FINAL REPORT

Approved for Public Release; Distribution is Unlimited.

Prepared for
Air Vehicle and Crew Systems Technology Department (Code 6021)
NAVAL AIR WARFARE CENTER - AIRCRAFT DIVISION
Warminster, PA 18974-5000

NOTICES

REPORT NUMBERING SYSTEM — The numbering of technical project reports issued by the Naval Air Warfare Center, Aircraft Division, Warminster is arranged for specific identification purposes. Each number consists of the Center acronym, the calendar year in which the number was assigned, the sequence number of the report within the specific calendar year, and the official 2-digit correspondence code of the Functional Department responsible for the report. For example: Report No. NAWCADWAR-92001-60 indicates the first Center report for the year 1992 and prepared by the Air Vehicle and Crew Systems Technology Department. The numerical codes are as follows:

CODE	OFFICE OR DEPARTMENT
00	Commanding Officer, NAWCADWAR
01	Technical Director, NAWCADWAR
05	Computer Department
10	AntiSubmarine Warfare Systems Department
20	Tactical Air Systems Department
30	Warfare Systems Analysis Department
50	Mission Avionics Technology Department
60	Air Vehicle & Crew Systems Technology Department
70	Systems & Software Technology Department
80	Engineering Support Group
90	Test & Evaluation Group

PRODUCT ENDORSEMENT — The discussion or instructions concerning commercial products herein do not constitute an endorsement by the Government nor do they convey or imply the license or right to use such products.

Reviewed By: ______________________ Date: __________
Contracting Officer's Technical Representative (COTR)
or
Point Of Contact (POC)

Reviewed By: Johan K Schmidt, LCDR MSC USN Date: 02JUL92
Branch Head

Reviewed By: T M Mitchell, CDR, MSC, USN Date: 2 JUL 92
Division Head

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
March 1991

3. REPORT TYPE AND DATES COVERED
Final

4. TITLE AND SUBTITLE
Adaptive Automation and Human Performance: II. Effects of Shifts in the Level of Automation on Operator Performance

5. FUNDING NUMBERS

6. AUTHOR(S)
Raja Parasuraman, Toufik Bahri, Robert Molloy, and Indramani Singh

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Cognitive Science Laboratory
THE CATHOLIC UNIVERSITY OF AMERICA
Washington, DC 20064

8. PERFORMING ORGANIZATION REPORT NUMBER
CSL-N91-2

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Vehicle and Crew Systems Technology Department
(Code 6021)
NAVAL AIR WARFARE CENTER - AIRCRAFT DIVISION
Warminster, PA 18974-5000

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
NAWCADWAR-92036-60

11. SUPPLEMENTARY NOTES
NAWCADWAR POC's: Jonathan Gluckman and Jeffrey Morrison

12a. DISTRIBUTION AVAILABILITY STATEMENT
Approved for Public Release; Distribution is Unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The application of adaptive systems to cockpit automation was discussed in a recent report by Parasuraman et al (1990). Among the issues they raised were the effects of different types of adaptive automation on performance, the adaptive logic used to implement task changes, pilot consent to suggested adaptive changes, the effects of automation failure, and so on. Empirical evidence relevant to these issues (from simulator or field studies) is still meager. The present studies form part of a program of investigations of these and other human-performance issues related to cockpit automation (Parasuraman et al., (1991). Here we briefly report the results of two experiments investigating the effects of shifts in the type and level of automation on operator performance in a multi-task environment. The first study examined the benefits and costs of adaptive automation shifts on operator performance. The second study examined the effects of variations in automation reliability on operator detection of automation failures.

14. SUBJECT TERMS
automation, adaptive automation, system reliability, aircraft automation, pilot workload, pilot situational awareness

15. NUMBER OF PAGES
6

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	**18. SECURITY CLASSIFICATION OF THIS PAGE**	**19. SECURITY CLASSIFICATION OF ABSTRACT**	**20. LIMITATION OF ABSTRACT**
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	SAR

NAWCADWAR-92036-60

EFFECTS OF SHIFTS IN THE LEVEL OF AUTOMATION ON OPERATOR PERFORMANCE[1]

Raja Parasuraman, Toufik Bahri, Robert Molloy, and Indramani Singh
Cognitive Science Laboratory, The Catholic University of America

INTRODUCTION

The introduction of intelligent automation in advanced cockpits has raised many human-factors issues. Although benefits have accrued from automation, potential costs have also been noted (Norman et al., 1988; Wiener, 1988; see Parasuraman et al., 1990, for a review). For example, automation may adversely affect a pilot's system awareness and ability to revert to manual control (Chambers & Nagel, 1985). Over-reliance on automated systems may also lead to pilot "complacency" (Wiener, 1981).

Adaptive automation [2] has been proposed as an automation concept that may preserve the benefits of automation while minimizing its costs (Rouse, 1988). In adaptive systems automated aids are invoked dynamically in response to changing mission demands. Such systems are thought to be superior to conventional automation because they provide for regulation of operator workload, maintenance of skill levels, and system involvement (Hancock & Chignell, 1988; Parasuraman, 1987). At present, however, empirical evidence for the efficacy of adaptive automation is lacking (Parasuraman et al., 1990).

The application of adaptive systems to cockpit automation was discussed in a recent report by Parasuraman et al. (1990). Among the issues they raised were the effects of different types of adaptive automation on performance, the adaptive logic used to implement task changes, pilot consent to suggested adaptive changes, the effects of automation failure, and so on. Empirical evidence relevant to these issues (from simulator or field studies) is still meager. The present studies form part of a program of investigations of these and other human-performance issues related to cockpit automation (Parasuraman et al., 1991). Here we briefly report the results of two experiments investigating the effects of shifts in the type and level of automation on operator performance in a multi-task environment. The first study examined the benefits and costs of adaptive-automation shifts on operator performance. The second study examined the effects of variations in automation reliability on operator detection of automation failures.

COST/BENEFIT ANALYSIS OF SHORT-CYCLE ADAPTIVE AUTOMATION

Adaptive automation involves *transitions* between levels of automation. The frequency of these transitions may vary between two extremes. In *long-cycle* adaptive automation, a particular flight function may be automated for long periods of time, then be carried out manually for a short period, and then revert for another long period to automated control. At the other extreme, in *short-cycle* adaptive automation, flight functions are cycled between manual and automated control more frequently, particularly if the adaptive logic is sensitive to small changes in task demands or pilot workload. The present study examined whether such short-cycle automation shifts benefit operator performance on flight-related tasks and whether there are any costs associated with the return to manual control following adaptive automation.

Methods

---

[1] Supported by contracts from NADC, Warminster, PA (John Deaton, technical monitor) and NASA Langley Research Center, Hampton, VA (Alan Pope, technical monitor). Address correspondence to Raja Parasuraman, Cognitive Science Laboratory, Catholic University of America, Washington DC 20064.

[2] Also referred to as "adaptive aiding" and "adaptive function allocation."

Twenty four young adults were tested on a multiple-task set consisting of compensatory tracking, system monitoring, and fuel management tasks presented simultaneously in separate windows of a 14-in color monitor (Comstock & Arnegard, 1990). The forcing function for the tracking task was derived from a sum of sine waves whose highest frequency was .064 Hz. Subjects had to keep a circular target symbol in the center of the window using a two-axis joystick under rate control. The performance measure was the integrated root mean square (RMS) error averaged over a 10-min block. The system monitoring task required subjects to detect "out-of-range" deviations of any one of four meters with continuously moving pointers (representing temperature and pressure). Accuracy and reaction time (RT) were the performance measures. Fuel management was a complex task requiring subjects to maintain the level of fuel in the two main tanks of the aircraft at a specified critical level by transfering fuel from several supply tanks. Fuel transfer was accomplished by activating one or more of several pumps. Pump flow rates varied and occasionally a pump was deactivated for 60 sec. Several performance measures were used for this task. Here we report the mean RMS error in fuel level deviations of the main tanks. (See Parasuraman et al., 1991, for additional details on each task and performance measure.)

Each task could be performed under manual or automation control. Under automation control all control inputs were disabled and the tasks were carried out by software with slight perturbations so that the display windows were not completely static. To simulate the workload associated with "supervisory control" of automation, occasional "deviations" in the automation occurred, consisting of slow target drifts for tracking, delayed "reaction times" for system monitoring, and delayed pump deactivations for the fuel management task. The degree to which subjects "supervised" the automation could then be roughly estimated by asking them to note any such deviations and report them after completing a block. To discourage subjects from active task processing aimed at detecting such deviations, subjects were not queried after every block.

In an initial practice session each subject performed all three tasks manually. Subsequently, each subject performed the three tasks under both manual and automation control in consecutive 10-min blocks in the sequence: manual control (M)—automated control (A)—return to manual control (RM). The M—A—RM sequence was repeated four times with intervening rest periods. During the automation phase one task was automated and the other two were under manual control; this simulates *model-driven* adaptive logic (vs. *performance-driven* logic; see Parasuraman et al., 1990). Each of the three tasks was automated individually for three separate groups of six subjects. A fourth group of six subjects was a no-automation control. During the automation phase of the other groups these subjects performed an unrelated single task (choice RT).

Results and Discussion

There were four main sets of results. First, automation supervision, as assessed by the number of "deviations" reported post-block, was adequate for tracking and monitoring but less so for the fuel management task. Subjects reported noting 50%, 40%, and 25% of the deviations, respectively, for these tasks.

Second, automation benefits (M vs. A) and costs (M vs. RM) were assessed in separate 4 x 2 x 4 ANOVAs (automated task x conditions [M-A or M-RM] x blocks [1-4]). Adaptive-automation benefits were obtained. Tracking performance was facilitated by automation of the monitoring and fuel management tasks, $F(1,10) = 9.15, p < .02$, as was monitoring performance by automation of tracking and fuel management, $F(1,10) = 4.60, p < .06$, for accuracy, $F(1,10) = 6.03, p < .05$, for RT. However, while fuel management was better under automation control of tracking and monitoring, the effect was not significant due to very high inter-subject variability in the RMS error measure, $F(1,10) = 2.34, p > .05$.

forms when a person first encounters a device in a particular context and which is then reinforced when it is re-encountered in the same way. Langer identified a number of antecedent conditions that produce this type of attitude, including routine, repetition, and extremes of workload—all conditions that can occur in the automated cockpit. We therefore reasoned that automation that was unchanging (but less than perfect) would be more likely to induce a condition of complacency than automation that varied, i.e., that complacency would be a function of the consistency of reliability of the automated task in a multi-task environment. We hypothesized that complacency would be potentially high for a group of subjects encountering automation with constant, unchanging reliability, because this group was more likely to develop a premature cognitive commitment regarding the nature of the automation and its efficiency. On the other hand, subjects encountering inconsistent automation reliability would have a more open attitude concerning the efficiency of the automation and hence would be less likely to be complacent. We therefore predicted that the accuracy of operator detection of automation failure would be greater for variable-reliability than for constant-reliability automation of a system monitoring task. We anticipated that this effect would emerge with time spent under automation control, so that the performance in the two conditions would diverge gradually from an initially equal starting point.

Methods

Twenty-four adults aged 19-43 years were tested on the same multi-task flight simulation used in the first study. The tracking and fuel management tasks were always performed manually and the system monitoring task always under partial automation control. Normally the pointers of each of the four meters fluctuated around the center of the scale within one limit in each direction from center. A green "OK" light above the meters remained on. Independently and at intervals according to a script, each pointer shifted its "center" position away from the middle of the meter. These "system malfunctions" were automatically detected and reset by the automation except when the automation failed, in which case the subject was required to respond. When the automation detected a malfunction, a red warning light above the meters came on and the computer corrected the fault in 4 sec. During this time the response keys for the monitoring task were disabled. When the automation failed, the green OK light remained on and the subject was responsible for detecting and correcting the malfunction by pressing one of the four function keys for each meter. Feedback was provided on correct responses by the pointer of the relevant meter moving immediately back to the center point and remaining there without fluctuating for a period of 1.5 sec. If a malfunction was not identified (whether by the automation or by the subject), the fault was corrected 10 sec from the beginning of its occurrence. Sixteen malfunctions were programmed to occur at unpredictable intervals within each 10-min block. Malfunctions appeared in a quasi-random sequence with the restriction that an equal number of malfunctions occurred for each meter (four) and for each direction of center shift (eight).

Automation reliability (the percentage of malfunctions correctly detected by the automation in each 10-min block) was varied as a between-subjects factor. For the constant-reliability condition, reliability was constant across blocks, whereas for the variable-reliability condition, reliability varied from block to block. In the constant-reliability condition, automation reliability was "high" for half the subjects (87.5%; 14 out of 16 malfunctions detected by the computer) and "low" for the other half (56.25%; 9 out of 16 malfunctions detected). Thus, for each 10-min block, the subject had to intervene for 2 malfunctions in the high-reliability condition and 7 malfunctions in the low-reliability condition. These levels remained constant across each 10-min block. In the variable-reliability condition, reliability alternated every 10 min from low (56.25%) to high (87.5%) for half the subjects and from high to low for the other half of the subjects.

Every subject completed four sessions, each consisting of three continuous 10-min blocks. In all conditions the automation routine was shut off for the last 20 min of session four (blocks 11 and 12) in order to simulate catastrophic automation failure. Subjects were trained and given 10

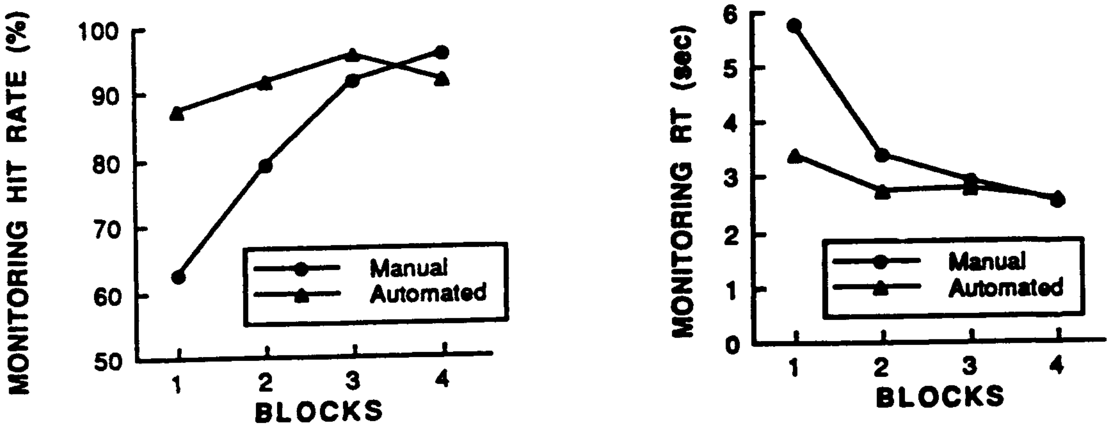

Figure 1. Effects of short-cycle automation of tracking on monitoring accuracy and RT.

Third, practice generally reduced automation benefits. As an example, see Fig. 1 for the monitoring task. Tracking automation increased monitoring accuracy but the benefit dissipated with blocks. While a performance ceiling could have limited the benefit in accuracy, monitoring RT had not reached a floor in the later blocks in which no automation benefit was obtained.

Fourth, no evidence of sustained automation cost was obtained. Mean performance levels on the M and RM blocks did differ for some conditions, but not significantly, and were not significantly different to the corresponding M and RM blocks for the no-automation group.

These results confirm that short-cycle adaptive automation can have beneficial effects on performance, as found previously with conventional, nonadaptive automation (e.g., Tsang & Johnson, 1989). No evidence was obtained for impaired efficiency associated with the return to manual control following automation. Training and practice reduced the automation benefit, but extended performance did not produce any automation costs. Although preliminary, these findings are encouraging with respect to the positive effects of adaptive automation. It remains to be seen, however, whether these benefits persist with long-cycle adaptive automation and whether there are any costs associated with long-term automation transitions.

OPERATOR MONITORING OF AUTOMATION FAILURE

High levels of automation increase the demand on the operator to monitor for possible automation failure (Chambers & Nagel, 1985; Parasuraman, 1987). Monitoring requirements may also increase if adaptive systems come into use because many automation transitions occur in such systems. For these and other reasons, the effects of different characteristics of automation on the ability of the operator to monitor automated systems need to evaluated.

Pilot "complacency" is often mentioned as representing one potential effect of automation relevant to monitoring performance (Thackray & Touchstone, 1989; Wiener, 1981). Although the term remains poorly defined (but see Singh et al., 1991), complacency is one of several behavioral coding categories used to classify aircraft incidents in NASA's Aviation Safety Reporting System (ASRS). The ASRS Coding Manual defines "complacency" as "self-satisfaction which may result in non-vigilance based on an unjustified assumption of satisfactory system state" (Billings et al., 1976).

We examined the role of "automation-induced complacency" by investigating the effects of automation reliability on operator monitoring of system failures during multi-task performance. Langer (1989) proposed the concept of *premature cognitive commitment* to refer to a mindset that

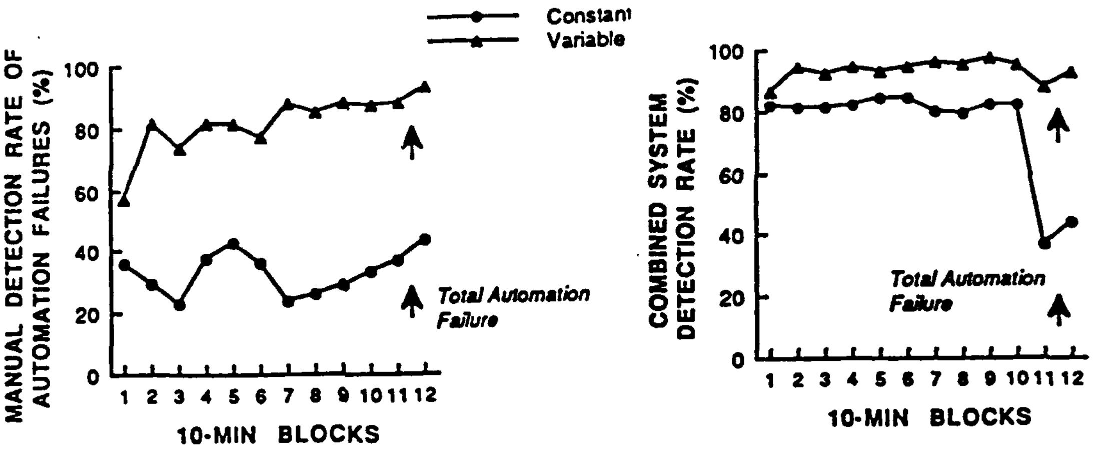

Figure 2. Effects of automation reliability on operator (left) and system (right) monitoring performance.

min of practice on all three tasks performed under manual control. They were asked to give equal attention to all three tasks during training and practice. At the end of the training session, performance feedback on all three tasks was given. The experimental phase began with subjects being informed that the monitoring task would be automated, so that they should focus their attention on the tracking and fuel management tasks. However, they were informed that the automation routine was not completely reliable and that they were required to supervise the automation and detect and respond to any malfunction that the automation failed to detect.

Results and Discussion

We predicted that initially the accuracy of manual intervention in the automated monitoring task should be roughly the same for the two experimental groups, but that later the performance of the variable-reliability group would be superior to that of the constant-reliability group. The results supported these predictions. The efficiency of manual intervention in the automated monitoring task was significantly higher in the variable-reliability condition than in the constant-reliability condition (see Fig. 2, left). The performance advantage for the variable-reliability group was present early and grew in size across the 12 10-min blocks. The effect of automation reliability (constant vs variable) was not significant in Block 1, $F(1,40) = 2.67, p > .05$, but was significant for all subsequent 11 blocks, $F(1,40) > 8.99$, $p < .01$. Detection accuracy for the variable-reliability group reached a high level quickly, exceeding 80% after block 4, but remained at a low level for the constant-reliability group. Following the total automation failure in block 11, performance of the constant-reliability group improved, but still did not recover to the level of the variable-reliability group. *Combined* system (i.e., human+computer) monitoring performance (cf. Thackray & Touchstone, 1989) was also superior with variable-reliability automation, particularly following total automation failure (Fig. 2, right). This suggests that operator complacency could limit system performance following total automation failure even if automation reliability was much higher (e.g., between 99% and 100%) than the "high" value of 87.5% used in this study.

The performance advantage in monitoring accuracy was not a function of speed-accuracy tradeoff because RT was unaffected by automation reliability. Initial performance level was also not a factor because both groups had similar, high detection rates under manual control. Moreover, there were no significant differences between the two subgroups of the constant-reliability condition, indicating, somewhat surprisingly, that the absolute level of automation reliability, within the range studied here, did not affect the pattern of results. This shows that the increased performance of the variable-reliability group was not a function of a higher "signal rate" afforded by less reliable automation. Finally, there were no significant differences in tracking or fuel-management performance between automation reliability conditions, showing that the performance advantage of the variable-reliability group was not due to their allocating processing resources away from tracking or fuel management to system monitoring.

These results provide what we believe is the first empirical evidence of the performance consequences of "complacency," unconfounded by other factors such as speed-accuracy tradeoff, initial performance capabilitities, signal rate effects, differential allocation of processing resources to other flight tasks, or experimenter effects (all subjects received the same instructions). Performance consequences of complacency in system monitoring were demonstrated to be related to characteristics of the monitoring task automation, namely automation reliability and consistency. In our ongoing research we are examining how other features of automation, including adaptive automation, influence operator efficiency in monitoring for automation failure.

References

Billings, C. E., Lauber, J. K., Funkhouser, H., Lyman, G., & Huff, E. M. (1976). *NASA Aviation Safety Reporting System.* (Technical Report TM-X-3445). Moffett Field, CA: NASA.

Chambers, N. & Nagel, D. C. (1985). Pilots of the future: Human or computer? *Communications of the ACM, 28,* 1187-1199.

Comstock, J. R. & Arnegard, R. J. (1990). *Multi-attribute task battery.* (Preliminary Technical Report). Hampton, VA: NASA Langley Research Center.

Hancock, P. A. , and Chignell, M. H. (1988). Mental workload dynamics in adaptive interface design. *IEEE Transactions on Systems, Man and Cybernetics, 18,* 647-658.

Langer, E. (1989). *Mindfulness.* Reading, MA: Addison-Wesley.

Norman, S., Billings, C. E., Nadel, D., Palmer, E., Wiener, E. L., and Woods D. D. (1988). *Aircraft automation philosophy.* (Technical Report). Moffett Field, CA: NASA Ames Research Center.

Parasuraman, R. (1987). Human-computer monitoring. *Human Factors, 29,* 695-706.

Parasuraman, R., Bahri, T., Deaton, J., Morrison, J., & Barnes, M. (1990). *Theory and design of adaptive automation in aviation systems.* (Technical Report). Warminster, PA: Naval Air Development Center.

Parasuraman, R., Bahri, T., & Molloy, R. (1991). *Adaptive automation and human performance. I. Multi-task performance characteristics.* (Technical Report CSL-N91-1). Washington DC: Cognitive Science Laboratory, Catholic University of America.

Rouse, W. B. (1988). Adaptive aiding for human/computer control. *Human Factors, 30,* 431-443.

Singh, I. L., Molloy, R., & Parasuraman, R. (1991). *Automation-induced "complacency": Development of a complacency-potential scale.* (Technical Report CSL-A91-1). Washington DC: Cognitive Science Laboratory, Catholic University of America.

Thackray, R. I., & Touchstone, R. M. (1989). Detection efficiency on an air traffic control monitoring task with and without computer aiding. *Aviation, Space, and Environmental Medicine, 60,* 744-748.

Tsang, P. S., & Johnson, W. W. (1989). Cognitive demands in automation. *Aviation, Space, and Environmental Medicine, 60,* 130-135.

Wiener, E. L. (1981). Complacency: Is the term useful for air safety? In *Proceedings of the 26th Corporate Aviation Safety Seminar.* (pp. 116-125). Denver, CO: Flight Safety Foundation, Inc.

Wiener, E.L. (1988). Cockpit automation. In E. L. Wiener & D. C Nagel (Eds). *Human factors in aviation.* San Diego, CA. Academic Press.