(2)

Report No. NAWCADWAR-92037-60

AD-A254 381

ADAPTIVE AUTOMATION AND HUMAN PERFORMANCE: III. EFFECTS OF PRACTICE ON THE BENEFITS AND COSTS OF AUTOMATION SHIFTS

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

OCTOBER 1991

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: K. Schmidt Date: 02JUL92

Branch Head LCDR MSC USN

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
October 1991

3. REPORT TYPE AND DATES COVERED
Final Report

4. TITLE AND SUBTITLE
Adaptive Automation and Human Performance: III. Effects of Practice on the Benefits and Costs of Automation Shifts

5. FUNDING NUMBERS

6. AUTHOR(S)
Raja Parasuraman, Brian Hilburn, 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

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-92037-60

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)*
Adaptive automation, or adaptive function allocation, is thought to maximize the benefits associated with cockpit automation while maintaining pilot involvement, enhancing situation awareness, and regulating workload. These claims have not been tested empirically. The present study examined the effects of short-cycle adaptive automation and practice on performance of flight-related functions in a multi-task environment. Twenty four non pilot subjects were tested on a PC-based flight-simulation task that included three primary flight functions -- tracking, monitoring, and fuel management. Each function could be automated or performed manually.
The results provide preliminary evidence that dynamic automation shifts over short cycles, of the type likely in adaptive systems, benefit performance of flight-related tasks, with no evidence of costs to performance following the return to manual control. Benefits are realized despite the added workload of supervisory control of automated functions. However, training procedures other than simple practice may be necessary to maximize and maintain the performance benefits associated with adaptive automation.

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

15. NUMBER OF PAGES
39

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

TABLE OF CONTENTS

SUMMARY

Adaptive automation, or adaptive function allocation, is thought to maximize the benefits associated with cockpit automation while maintaining pilot involvement, enhancing situation awareness, and regulating workload. These claims have not been tested empirically. The present study examined the effects of *short-cycle* adaptive automation and practice on performance of flight-related functions in a multi-task environment. Twenty four nonpilot subjects were tested on a PC-based flight-simulation task that included three primary flight functions--tracking, monitoring, and fuel management. Each function could be automated or performed manually. An adaptive procedure was simulated by shifting from manual to automatic control and back every 10 min in a 30-min session. Supervisory control of automated functions was simulated by requiring subjects to report the number of minor "deviations" (not malfunctions) in the automation routine at the end of some but not all blocks ("catch trials"). Practice effects were assessed by testing subjects over four 30-min sessions. Each subject performed four cycles of the form: Manual - Automated - Return to Manual ({M} - {A} - {RM}). Benefits and costs associated with changes in the level of automation of each of the three flight functions were evaluated. Automation benefits were assessed by comparing multiple-task performance in the {M} and {A} blocks. Automation costs were assessed by comparing performance in the {M} and {RM} blocks.

There were five major sets of results. (1) Operators reported about 40% of all "deviations" in automated functions in post-session queries, indicating that they maintained supervisory control of automation as required. (2) Dynamic adaptive automation resulted in performance benefits for all three flight functions evaluated. The performance benefit was realized generally across tasks and conditions. (3) However, adaptive automation benefits were largest and only statistically reliable during the early phase of performance. Benefits declined and were generally not significant in later blocks. (4) Practice effects varied for manual and automation blocks. While performance in manual blocks increased steadily over time, performance in automation blocks did not show consistent improvement. (5) Finally, no evidence of automation costs was obtained for any of the three flight functions.

The results provide preliminary evidence that dynamic automation shifts over short cycles, of the type likely in adaptive systems, benefit performance of flight-related tasks, with no evidence of costs to performance following the return to manual control. Benefits are realized despite the added workload of supervisory control of automated functions. However, training procedures other than simple practice may be necessary to maximize and maintain the performance benefits associated with adaptive automation.

INTRODUCTION

Recent technological advances have made viable the implementation of intelligent automation in advanced tactical aircraft. The use of this technology has given rise to a number of new human factors issues and concerns. While many benefits related to efficiency and safety have resulted from cockpit automation, some costs have also been noted (Wiener, 1988). Errors in highly automated aircraft have been linked to the adverse effects of automation on the pilot's system awareness, workload, and ability to revert to manual control (Chambers & Nagel, 1985; Parasuraman, 1987; Wiener, 1988). These problems have been attributed to limitations in automation design, which has been largely technology-driven, or based on technical capabilities rather than on human operator capabilities and needs (NASA, 1989).

Partly in response to these concerns, *adaptive automation*, or automation that is implemented dynamically in response to changing task demands on the pilot, has been proposed (Rouse, 1988). Adaptive automation, also known as "adaptive function allocation" or "adaptive aiding," is thought to be a superior concept to nonadaptive ("traditional" or *static*) automation in that it provides for improved pilot situational awareness, regulation of operator workload and vigilance, maintenance of skill levels, and task involvement (Hancock et al., 1985; Hancock & Chignell, 1988; Noah & Halpin, 1986; Parasuraman, 1987; Rouse, 1988). According to this view, many of the benefits of automation can be maximized and the costs minimized if automation is implemented in an adaptive manner rather than in an all-or-none fashion. Adaptive processes are thought to allow synergistic communication between the pilot and aircraft subsystems. For example, the pilot can actively control a process during moderate workload and allocate this function to an automated subsystem during peak workload if necessary. Statically automated processes, on the other hand, can impact negatively on pilot workload and lead to a loss of system awareness (Norman et al., 1988), particularly if they provide inadequate feedback to the pilot (Norman, 1991). Adaptive automation is thus believed to allow the advantages of automation to be realized while maintaining pilot involvement in the system.

Thus far these claims remain largely untested. The efficacy of adaptive function allocation has yet to be demonstrated reliably for a broad range of flight functions, whether in the laboratory, simulator, or cockpit. In an extensive review of this field, Parasuraman, Bahri, Deaton, Morrison & Barnes (1990) found that few empirical studies of the effects of adaptive automation on performance have been carried out. If adaptive automation is to be a viable cockpit design option, more needs to be learned about its effects on pilot performance under different flight conditions. Most previous research has examined effects on operator performance of *static* automation, i.e. where the set of tasks that are automated and manual remains fixed and invariant over time (cf. Parasuraman, Bahri, Molloy, & Singh, 1991b).

FOUR CHARACTERISTICS OF ADAPTIVE SYSTEMS

Parasuraman et al. (1990) discussed a number of issues relevant to the design of adaptive systems. Four issues are particularly pertinent to the present investigation. First, what is adapted to in adaptive automation? Second, how often does adaptation occur? Third, when an adaptive change occurs, how does it change the nature of the task performed by the operator? Four, what are the training needs for operators of adaptive systems? (This by no means exhausts the list of issues relevant to the design of adaptive systems: in their review Parasuraman et al. (1990, Table 1) enumerate 10 issues and several sub-issues within each issue that should be addressed in planning the design of adaptive systems.)

Bases for Adaptive System Control

On what basis should an adaptive system implement changes in the level of automation? Flight functions can be allocated dynamically to the pilot or to automated subsystems in a number of ways. Rouse (1988) categorized all adaptive schemes as based on either *measurement* or *modelling* of the operator. However, another category of adaptation, based on critical environmental events, does not require either measuring or modelling the pilot's performance (Barnes & Grossman, 1985). In general, Parasuraman et al. (1990) identified four major classes of procedures on which adaptation may be based: (1) critical events; (2) pilot performance measurement; (3) pilot psychophysiological assessment; and (4) pilot performance modelling.

Each of these methods has its merits and drawbacks. In the critical-events method, changes in the level of automation are tied to the occurrence of specific tactical events, for example an increase in the number of targets appearing on a radar display. The advantage of critical-event logic is that it can be tied closely to actual task or mission events. This method of function allocation is adaptive because if the critical events do not occur, the automation is not invoked. Its disadvantage is its potential insensitivity to the current needs of the pilot. The method assumes a priori that the appearance of some critical event necessitates automation of some functions because the pilot cannot efficiently carry put these functions and deal effectively with the critical event. But it may not be the case that pilot workload always increases beyond some acceptable upper limit when the critical event occurs.

Pilot measurement attempts to overcome this limitation. Various pilot mental states (workload, vigilance, strategies, intentions, etc) may be assessed dynamically and these measures then fed to some adaptive logic. Alternatively pilot states and performance may be modelled theoretically, with the adaptive algorithm being driven by the model parameters. Measurement has the advantage of being an "on-line" technique that can potentially respond to unpredictable changes in pilot cognitive states. Performance measurement occurs "after the fact," that is, it follows the point in time when an adaptive change may be necessary because of a sudden change in workload. Hence this technique must use the previous history of performance in order for to provide a rapid signal to the adaptive logic. In highly automated systems where the operator makes few overt responses performance measurement may be so impoverished as to not be practical. However, psychophysiological measurement, which can be obtained continuously or with a greater sampling frequency than performance measures, may allow for a higher degree of predictive capacity.

In any case, pilot measurement methods are only as good as the sensitivity, diagnosticity, and speed of the measurement technology. Performance modelling provides an alternative to measurement as a basis for adaptive automation. Modelling techniques have the advantage that they can be implemented off-line and easily incorporated into rule-based expert systems. However, this method is only as good as the theory behind the model, and many models may be required to deal with all aspects of pilot performance in a complex task environment. Hybrid systems that combine measurement and modelling, or critical-event logic and performance measurement, or other possible combinations of these methods, may maximize their relative benefits and minimize their disadvantages.

Cycles of Adaptive Automation

A second important characteristic of an adaptive system, irrespective of which type of adaptive logic it is based on, is by frequency with which the system shifts between levels of automation, i.e., between manual and automatic conditions and vice versa (Figure 1). In *long-cycle* adaptive automation, a particular flight function might be automated for long periods of time, then carried out manually for a short period of time, and then revert to automatic control for another long period. 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 either task demands or pilot workload. The influence of adaptive automation on pilot performance may vary with these two extremes of frequency of shifting between levels of automation. For example, automation "deficits" due to skill loss or degradation of pilot mental models of the task may be more likely under long-cycle adaptive automation than under short-cycle adaptive automation.

Short-cycle adaptive automation

{M} {A} {M} {A} {M} {A} {M} {A} {M} {A} {M} {A} --- etc

Long-cycle adaptive automation

{M} {M} {M} {M} {A} {A} {A} {A} {M} {M} {M} {M} --- etc

where {M} refers to a set of tasks each performed under manual control and {A} refers to the same set of tasks, one or more of which are performed under automation control. Alternatively {M} and {A} may refer to different levels of automation in the continuum between full automation and full manual performance.

Figure 1. Cycles of adaptive automation

Supervisory Control of Automated Functions

A third characteristic of adaptive automation that is important from the perspective of the present investigation concerns the nature of an automated function. What exactly does it mean to automate a function normally performed manually by a pilot? Clearly, automation does not mean that the function disappears (although the pilot may well have reduced awareness of the function). Automation also does not necessarily mean that the pilot "has less to do." As Wiener (1988) has shown, there is good evidence that automation changes the nature of the workload imposed on the pilot but does not invariably reduce it. Workload may even be increased if the automated function needs to be monitored frequently by the pilot in order to ensure that the automation is working efficiently. This has important implications for experimental studies of the effects of automation on performance. In other words, automation effects cannot be studied in the same way as standard multi-task experiments. Typically in these experiments dual-task performance is compared to a condition in which the subject is asked to ignore one of the two tasks and to only perform the other one. Superficially this is the same as if the task that must be ignored is performed by automation. But in real systems the automated device cannot be "ignored" and there is some workload associated with its supervision by the pilot. Thus, in order for experimental studies of automation and human performance to be relevant to real-world automation issues, they must be able to simulate the added workload associated with supervisory control of automated systems.

Training

Training is perhaps one of the most important issues relevant to adaptive systems, and one that influences each of the other issues discussed previously. Successful training of pilots to use adaptive systems will determine how effective such systems are. Automation can place conflicting demands upon pilots which they may not be well-equipped to meet (e.g. passive monitoring versus active control) unless they have been specifically trained to cope with these demands. It has been suggested that inadequate training may lead to several automation-induced problems in the cockpit. For example, the negative effect of automation on monitoring performance may be related, in part, to a lack of "automation-based" skills (Parasuraman et al., 1990). This reflects inappropriate training because automation necessitates a shift from psycho-motor skills to more cognitive and problem-solving skills, which may not be emphasized in the training program (Idaszak & Hulin, 1989). Unfortunately, training needs for operators of automated systems have not received much attention, and in fact there is virtually no literature comparing different training methods for users of adaptive systems.

Adaptive automation suggests the relevancy of *adaptive training* methods already in use. In contrast to adaptive automation research, there is a long history of research on adaptive training methods. It has long been held that adaptive training is a superior form of training, and many investigators have advocated its feasibility. However, as Lintern and Gopher (1978) showed in a comprehensive literature review, the empirical evidence for its effectiveness is quite weak. Thus, adaptive training methods alone may not be the most appropriate technique to help train users of automated systems, although they may still be helpful when combined with other methods. Development of effective training techniques for adaptive automation is a research priority because the shifting task requirements implicit in adaptive automation will require diagnostic and control skills substantially different from those of current, fixed automation aircraft. Unless these are developed, the anticipated benefits of adaptive systems may not be realized.

THE PRESENT STUDY

The above-mentioned four issues were each examined in the present study. One issue, the basis for adaptation, or the adaptive logic, was implicit to the investigation but was not examined explicitly by experimental manipulation. Cycles of adaptation were examined by shifting operators repetitively between different levels of automation, although only short and not long cycles were examined. Supervisory control of automated functions was required. Finally, an initial study of training issues was made by investigating the effects of practice.

The present study thus examined the effects of short-cycle automation shifts and practice on flight-related task performance. In particular, we investigated whether there are any benefits and costs associated with changes in the level of automation made as a result of some adaptive procedure, and the effects of practice on benefits and costs. We did not assume any particular form for the adaptive logic. Rather, we simulated the output of the adaptive logic by simply shifting the level of automation at specified times (10 minute intervals) during multi-task performance. Apart from their temporal regularity, such shifts could be generated by any one of the four classes of adaptive procedures described by Parasuraman et al. (1990). We recognized that the type of adaptive procedure used may influence the pattern of performance consequences of adaptive automation, but felt that as this was an initial study such effects could be investigated in future studies. Finally, automation of a function was simulated by leaving intact that part of the display showing the automated function, with the exception that no overt operator responses were required (all "responses" were made by the automation routine). However, subjects were required to maintain covert supervision of the automation and the efficiency of supervisory control was assessed by post-task queries (see Methods section).

To examine the effects of short-cycle automation shifts on performance of flight-related functions, we used a laboratory flight-simulation task modified from the MAT battery developed by Comstock and Arnegard (1990). An earlier report in this series described our initial study of the performance characteristics of our modified version of this task (Parasuraman, Bahri, & Molloy, 1991a). The task simulates flight-related functions tapping three broad domains of performance: perceptual-cognitive (instrument monitoring), cognitive-strategic (fuel management), and perceptual-motor (tracking). A short-cycle schedule of the form {M} {A} {RM} was used, where M = manual, A = automated, and RM = return to manual. The cycle was repeated four times over separate blocks to examine the effects of practice.

Our analytic procedures were as follows. We first investigated whether subjects supervised the automation as instructed so that we could be confident that the performance results could be related to automation per se rather than to removal of a task (i.e. as in a multi-task experiment). Then we sought to investigate three hypotheses: (1) that performance on concurrent flight-related tasks would be enhanced by dynamic automation of one task; (2) that performance would be degraded during a return-to-manual condition, when the operator was required to reassume manual control of all tasks; and (3) that practice would reduce both automation costs and benefits. By comparing performance under manual and automation control conditions, automation benefits could be assessed ({M} vs {A}). By comparing performance under manual and return-to-manual conditions, automation costs could be evaluated ({M} vs {RM}). The generality of these hypotheses was tested by investigating automation shifts involving each one of the three flight functions of tracking, monitoring, and fuel management. Results indicating that dynamic automation shifts over short cycles benefit performance with only minor or no costs would provide preliminary evidence for the efficacy of adaptive automation, at least for the laboratory tasks investigated in the present study.

METHODS

Subjects

Twenty four young adults volunteered to serve as subjects for this study. Most were students from the Catholic University of America. All had normal (20/20) or corrected-to-normal Snellen visual acuity. Subjects ranged in age from 18 to 29 years, and consisted of roughly equal numbers of males and females. None of the subjects had previously participated in any similar research conducted by the Cognitive Science Laboratory. The 24 subjects were randomly assigned to four groups of six subjects, with each group performing a different task under automation control, as described further later.

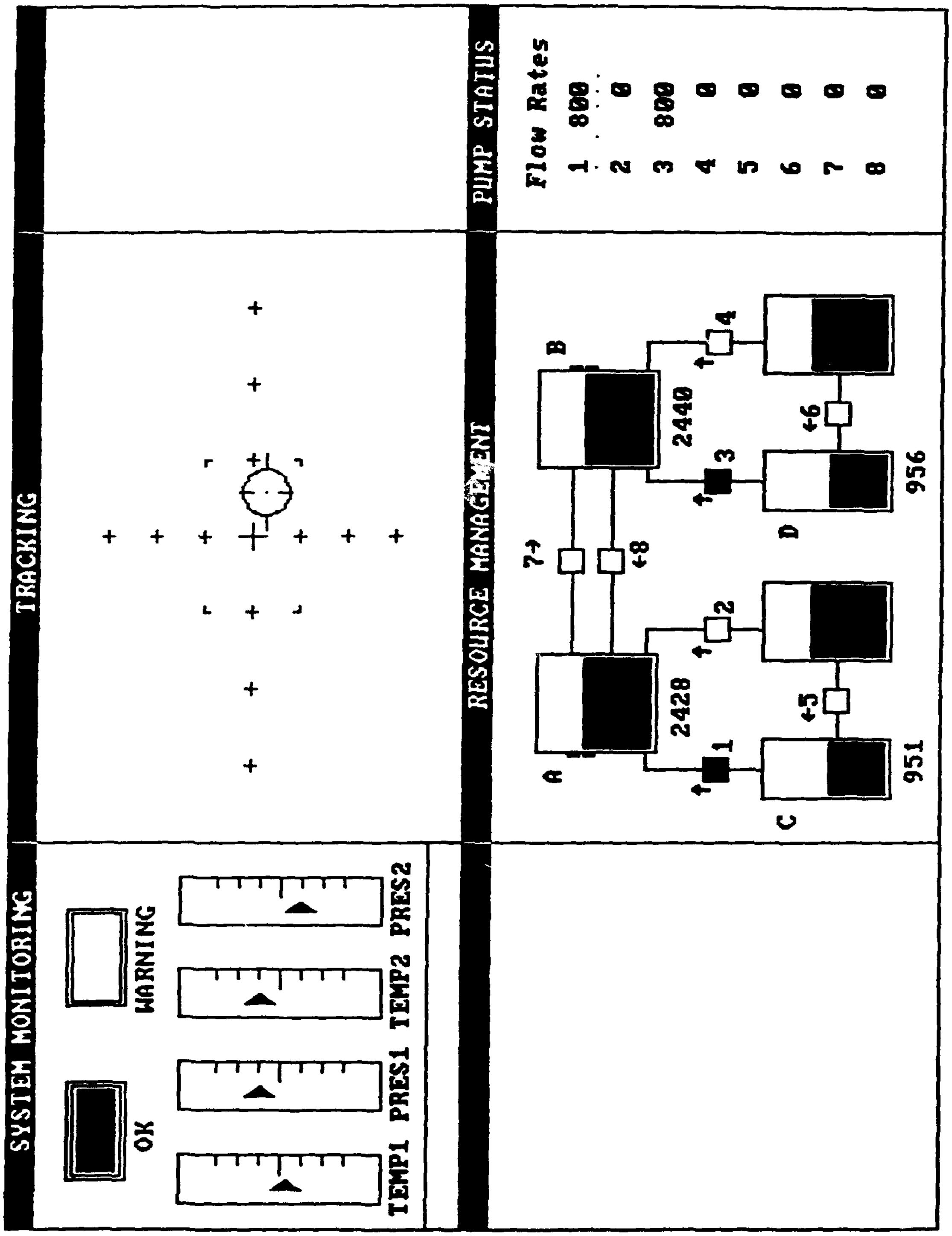


Figure 2. Flight-simulation task.

Flight-Simulation Task

A modified version of the Multi-Attribute Task Battery (MAT) developed by Comstock and Arnegard (1990) was used (Parasuraman et al., 1991a). The MAT is a multi-task flight simulation package comprising component tasks of tracking, system monitoring, fuel management, communications, and scheduling. In the present study only the tracking, monitoring, and fuel-management tasks were used. The extended version developed by Parasuraman et al. (1991a) allows each component task to be performed either manually or under automation control. Software scripts enable precise control over the timing within each session of critical events such as engine malfunctions, changes in tracking difficulty, fuel pump failures, etc. The tracking, monitoring, and fuel-management tasks were displayed in separate windows of a 13-in color display monitor (see Figure 2).

Tracking

Manual Mode. A first-order, two-dimensional compensatory tracking task with joystick control was presented in one window of the display (see Figure 2). Dashed x- and y- axes were provided for reference. A green circular target symbol representing the deviation of the aircraft from its course fluctuated within the window in the x- and y- directions according to a specified forcing function consisting of a sum of nonharmonic sine waves. The highest (cut-off) frequency of the forcing function was 0.06 Hz. Control inputs were provided by a displacement joystick using first-order or velocity control. If no control input was applied, the aircraft symbol drifted away from the center towards the edges of the window. The subject's task was to keep the aircraft within the central rectangle by applying the appropriate control inputs in the x- and y- directions.

Automated Mode. Under automation control, the joystick was disabled and the aircraft movements were compensated for by software. However, small fluctuations around the center of the window remained, to simulate random perturbations in the automatic control. Under normal automated conditions, therefore, the aircraft appeared to be anchored at the center of the window, but with very small movements about the center that give the appearance of a dynamic rather than completely static display.

As discussed previously, workload associated with supervision of automated functions is an important aspect of the automated cockpit. In order to simulate supervisory control of automated tracking, occasional "deviations" (not malfunctions) in the automatic control were built into the script for automation. Under these conditions, the aircraft symbol began a slow drift away from the center until it reached the inner rectangle and then drifted back. Four such deviations, one to each corner of the tracking window, were programmed to occur at random intervals during a 10-min block.

System Monitoring

Manual Mode. The upper left window in Figure 2 shows the system monitoring task, which consists of four vertical gauges with moving pointers and green "OK" and red warning lights. The scales for the gauges were marked as indicating the temperature (TEMP1, TEMP2) and pressure (PRES1, PRES2) of the two aircraft engines. Normally, the green OK light was on and the pointers fluctuated around the center of the gauge within a small band that extended .25 in on either side of center. Occasionally, however, a "system malfunction" occurred, and the pointer on one of the four engine gauges went "off limits." That is, independently and at intervals according to a pre-defined script, the pointer shifted its "center" position away from the middle of the vertical gauge. In each 10-minute block of the simulation, 4 such "system malfunctions" occurred at unpredictable intervals ranging from 90 to 320 seconds. These occurred in a pseudorandom sequence except that one malfunction appeared on each vertical gauge. The operator was responsible for detecting pointer shifts occurring on any of the four gauges, regardless of direction, and to respond by pressing one of the corresponding function keys T1, T2, P1, or P2, which were identified below each vertical gauge. Feedback was provided when the out-of-range status of a gauge was correctly identified by the pointer of the appropriate gauge moving immediately back to the center point and remaining there without fluctuating for a period of 1.5 seconds. Incorrect resets (i.e., false alarms) were not accompanied by a return to center. If the subject failed to detect a malfunction, the fault was automatically corrected 10 seconds from the beginning of its occurrence. Subjects were instructed to respond to deviations as quickly as possible, while keeping their error rate to a minimum.

Automated Mode. Under automation control, the keyboard keys T1, T2, P1, and P2 were disabled and the scripted engine malfunctions were identified and responded to by software. To enable the operator to know that the automation had properly detected and corrected the malfunction, the automation responded with a "reaction time" of 4 sec. To evaluate the operator's efficiency in supervisory control of the automation, occasional "deviations" in the efficiency of control were scripted, as for the automated tracking task. When such a deviation occurred, the automation correctly identified and corrected the malfunction, but with a delayed "reaction time" of 10 sec. Four such deviations were presented at random time intervals during a 10-min block.

Fuel Management

Manual Mode. This task is a simulation of the actions needed to manage the fuel system of the aircraft. Figure 2 displays the fuel (or resource) management window. The six rectangular regions represent tanks which hold fuel, the green levels within each tank indicating the amount of fuel in each tank. The pumps connecting the tanks allow the transfer of fuel from one tank to another in the direction indicated by the corresponding arrow and fuel line. The numbers underneath four of the tanks (Tanks A, B, C, and D) represent the amount of fuel in gallons for each of the tanks. This number was updated every 2 seconds as the amount of the fuel in the tanks increased or decreased. The maximum capacity was 4000 gallons each for Tanks A and B and 2000 gallons each for Tanks C and D. The remaining two supply tanks had an unlimited capacity.

Subjects were required to maintain the level of fuel in both Tanks A and B at 2500 gallons each. This critical level was indicated by a tick mark in the shaded bar on the side of these two tanks. The numbers under each of these tanks provided another means of feedback for the subject. The shaded region surrounding the tick mark represented acceptable performance. Tanks A and B were depleted of fuel at the rate of 800 gallons per minute. Therefore, in order to maintain the task objective, subjects had to transfer fuel from the lower supply tanks by activating the pumps. Each pump could only transfer fuel in one direction, as indicated by the corresponding arrow. These pumps were turned on when the corresponding number key was pressed by the subject. Pressing the key a second time turned off that particular pump and so on. The pump status was indicated by the color of the square area on each pump. When that area was black, or lacking in color, the pump was turned off. A green light in this area indicated that the pump was actively transferring fuel. The flow rates for each pump were presented in the "Pump Status" window. The first column of numbers represents the pump number, 1 through 8. When a pump was activated, its flow rate was shown next to the pump number in this window. When a pump was off, its flow rate was shown as 0. Pump 1 and 3 transferred fuel at the rate of 800 gallons per minute, Pumps 2, 4, 5, and 6 at 600 gallons per minute, and Pumps 7 and 8 at 400 gallons per minute.

Pump faults occurred unpredictably during each 10-minute block of the simulation. This was indicated by the appearance of a red light in the square on the pump and resulted in the pump becoming inactive. Fuel could not be transferred through that pump until the fault was corrected.

The operator had no control over the fault correction; the duration of the fault was written into the pre-defined simulation script. When the fault was corrected, the status of that pump was automatically returned to the "off" mode, regardless of its status before the fault condition. Likewise, when a tank became full to capacity, all incoming pumps were automatically turned "off". For example, if all of the pumps were activated and Tank A reached its capacity of 4000 gallons, Pumps 1, 2, and 8 would automatically turn off. Furthermore, if a tank were to become totally depleted of fuel, all outgoing pumps would be deactivated. At the onset of each flight simulation, Tanks A and B contained approximately 2500 gallons of fuel each and Tanks C and D approximately 1000 gallons of fuel each. All pumps were turned off at the beginning of the task, leaving all strategic action to the operator's discretion.

Automated Mode. Under automation control, the keys for activating pumps 1 through 8 were disabled. All pump activations were executed from a script that mimicked expert performance [1], in combination with the following restrictions: (1) all fuel level changes were responded to; (2) appropriate pump activations were executed; (2) no "extra" pumps were activated (e.g., activating pump 4 when that has no direct effect on fuel level, an action that sometimes occurs during manual performance). To evaluate the level of operator supervision of the automation, occasional "deviations" in pump failure activity were programmed, as for the tracking and monitoring tasks. When such a deviation occurred, pump failures lasted for 90 sec instead of 60 sec. Four such deviations occurred during a 10-min block.

Procedure

After completing a background questionnaire and visual acuity screening test, each subject received a brief explanation of the experiment and was instructed how each of the three tasks was to be controlled. Following this initial instruction, each subject was administered a 5-min practice session, during which all three tasks were performed manually. The practice session was followed by two experimental sessions, held on separate days. Two 30-min blocks were administered on each day (with a short intervening rest break). Thus, practice effects were evaluated over four 30-min blocks. Each 30-min block consisted of a short-cycle adaptive automation schedule comprised of three consecutive 10-min sub-blocks, as follows: Manual control of all three tasks (manual condition, {M}); Automated control of one of the three tasks (automated condition, {A}); Manual control of all three tasks (return-to-manual condition, {RM}).

---

[1] Defined as the performance of two laboratory personnel who had over 30 hours of experience of manual performance on the fuel management task.

This {M} - {A} - {RM} sequence comprised one complete 30-min cycle. At the end of each block, subjects were provided screen-displayed feedback on their performance on the three tasks.

Under manual conditions, instructions for performing each task were as described previously. Subjects were also asked to devote equal attention to all three tasks. In the automated conditions, in which two tasks were manual and one automated, subjects were asked to devote primary attention to the two manual tasks. However, they were also told that they must supervise the automated task in order to make sure the automation worked efficiently. This was done by inserting occasional "deviations" in the automation routine, as described previously. Subjects were not required to make any overt response to such deviations but were simply instructed to notice and keep a mental count of them. Subjects were queried at the end of the block whether they had noticed any deviations, and if so, how many. To ensure that subjects supervised the automated task during automation blocks, but to discourage prolonged, active processing of the automation in order to detect deviations, we employed a "catch trial" procedure whereby subjects were only questioned after some (but not all) blocks to report the number of deviations they had noticed.

Design and Analysis

Supervisory control was evaluated by the percentage of automation "deviations" reported by subjects on "catch trials" of automated blocks. Tracking performance was assessed as follows. The x and y control inputs were sampled at 10 Hz to yield the x and y deviations. The root mean square (RMS) error was then computed for the samples obtained over a 1-second period. In computing the combined horizontal and vertical deviations from the target, vertical deviations were converted (in proportion to the monitor x and y resolution) to horizontal pixel units before combination with the horizontal deviations:

$$\text{RMS error} = \sqrt{\left[\sum_{i}^{N} \{\Delta x_i^2 + (K\Delta y_i)^2\}/N\right]}$$

where Δx and Δy are the x and y deviations, K is the monitor resolution ratio (horizontal/vertical), and N is sample size. RMS error scores for successive 1-sec epochs were then averaged over a 10-min period to yield a mean RMS error score for a block. Several measures were collected for the monitoring task. Percentage of correct resets of gauges (i.e., hit rate) and the absolute number of incorrect resets (false alarms) were collected for each 10 minute block. RTs were also recorded for each correct reset. Finally, a global measure of performance of the fuel management was obtained by computing the mean RMS error in the fuel levels of Tanks A and B (deviation from the required level of 2500 gallons). Fuel levels were sampled and RMS error computed over a 30-second period. RMS error scores for successive periods were then averaged over 10 minutes to yield a mean RMS error score for a block.

BLOCK	1			2			3	4
	{M}	{A}	{M}	{M}	{A}	{M}	¥'¥'¥'¥'	
	T	[T]	T	¥'¥'¥'¥'				
Tracking	MN	MN	MN					
Automated	F	F	F					
	T	T	T	¥'¥'¥'¥'				
Monitoring	MN	[MN]	MN					
Automated	F	F	F					
	T	T	T	¥'¥'¥'¥'				
Fuel Management	MN	MN	MN					
Automated	F	[F]	F					
	T	x	T	¥'¥'¥'¥				
Control	MN	x	MN					
	F	x	F					

Legend

{M} = ALL 3 TASKS MANUAL
{A} = 1 TASK AUTOMATED, 2 TASKS MANUAL
T = TRACKING. MN = MONITORING. F = FUEL MANAGEMENT.
x = INTERPOLATED ACTIVITY
[ ] = AUTOMATED TASK (UNDER SUPERVISORY CONTROL)

Figure 3. Experimental design.

The performance indices for the tracking, monitoring, and fuel management tasks were submitted to 4 (Automation Group) x 2 (Automation Condition) x 4 (Blocks) analyses of variance (ANOVAs). ANOVAs were carried out separately to assess automation benefits ({M} vs {A}) and automation costs ({M} vs {RM}). The automated task was manipulated as a between-subjects factor of Automation Group. Each subject was assigned to one of four groups that corresponded to the one task that would be automated during automation blocks: tracking, monitoring, fuel management, or none (see Figure 3). That is, subjects in three experimental groups had either the monitoring, tracking or fuel management task automated while a fourth group served as a no-automation control. Each group performed the short-cycle adaptive automation sequence {M} {A} {M} four times (see Figure 3). During the automation block of the experimental groups, control subjects performed an unrelated choice RT task. Because the control group did not perform any task under automation control, they served primarily as a control for assessing automation costs under the return-to-manual {RM} condition. The second factor, Automation Condition, was varied within subjects; factor levels for the automation-benefit ANOVAs were {M} and {A}, and for the automation-cost ANOVAs {M} and {RM}. Finally, the third factor was Blocks, for the four consecutive 30-min blocks of performance.

RESULTS

Supervisory Control of Automated Functions

The efficiency of operator supervision of automation was assessed by computing the percentage of automation "deviations" reported by subjects at the end of a block (i.e. on "catch trials"). The mean number of deviations reported for the tracking, monitoring and fuel management tasks, were 50%, 40% and 25%, respectively (Figure 4). The overall level of supervision was thus satisfactory, at least for the tracking and monitoring tasks. The variation in report accuracy across tasks probably simply reflects differences in the discriminability of the deviations in each task window.

Figure 5 shows that there was a clear practice effect in the number of automation deviations reported on catch trials over the four blocks. Averaged across the three tasks, the percentage of reported deviations increased from roughly 30% after block 1 to nearly 60% after block 4. This probably represents a cuing effect---subjects were instructed before the experiment that they might be queried about automation deviations after some of the blocks. In fact, they were queried after each of two blocks that were randomly selected from the four blocks that subjects performed.

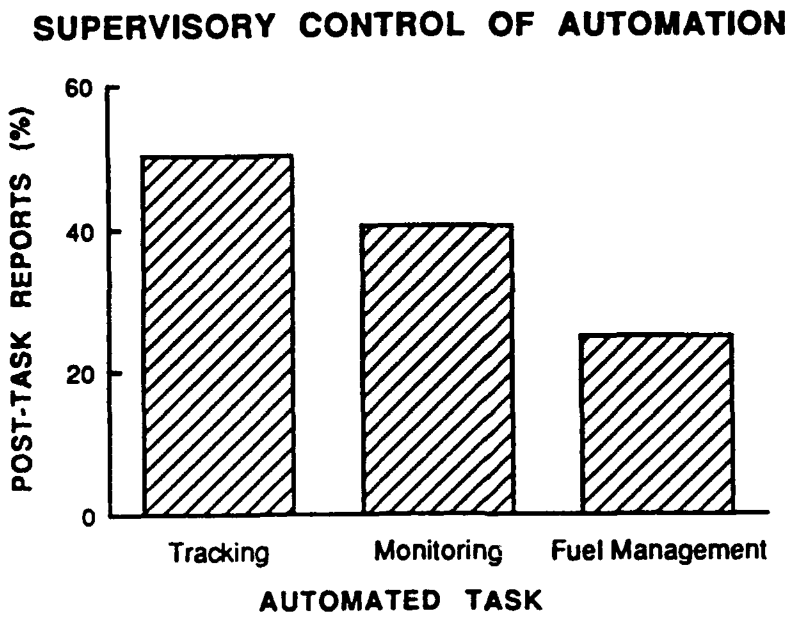


Figure 4. Percentage of reported automation deviations for different automated functions.

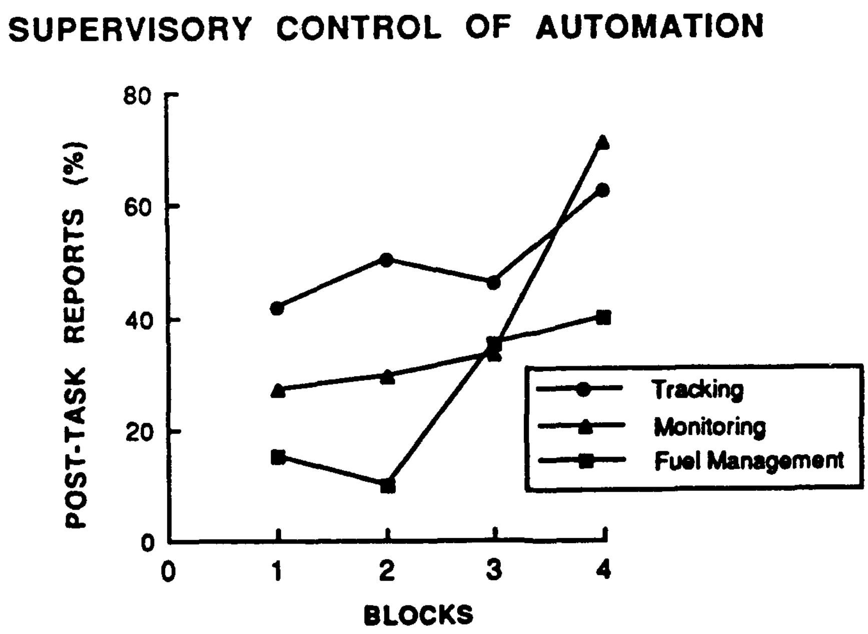


Figure 5. Percentage of reported automation deviations as a function of blocks.

Adaptive Automation Benefits

As described previously, automation benefits were assessed by comparing performance in the manual {M} and automated {A} blocks by ANOVAs. This was done separately for each dependent measure. Second, because we hypothesized that adaptive automation would benefit performance, we investigated the effects of the shift from manual to automatic control through planned multiple comparisons of performance scores for each combination of automation condition and automation group, using the Newman-Keuls procedure (alpha = .05). Third, because we predicted that practice would reduce automation benefits we examined the effects of automation at each of the four blocks and the effects of blocks for each automation condition using tests of simple effects (Winer, 1971). Finally, we examined automation benefits in general by computing composite performance scores. Theoretical and analytical tools for performance assessment in dual-task studies are available, such as the Performance Operating Characteristic (Navon & Gopher, 1979), but comparable tools for evaluating *multiple*-task performance have not yet been developed. LeMay and Comstock (1990) suggested the simple expedient of combining dissimilar performance measures for multiple tasks using standard scores, thereby arriving at a composite performance measure. We used this technique to evaluate the overall effects on multiple-task performance of automating a single task, and the changes in automation effects with practice.

Tracking Performance

Tracking efficiency significantly improved with dynamic automation of a second task, $F(1,10) = 9.15$, $p<.02$. These results can be seen in Figure 6. When the monitoring task was automated following the transition from the manual to automated condition, tracking RMS error was reduced (Figure 6a). Tracking RMS error was also reduced with automation of the fuel management task (Figure 6b). Mean tracking RMS error scores declined slightly over blocks, particularly following automation of fuel management (see Figure 6b). However, the Blocks factor was not significant, $F(3,30)=2.28$, $p>.09$, nor did it enter into a significant interaction with Automation Condition, $F(3,30)=1.11$, $p>.3$. All other sources of variance were not significant.

Multiple comparisons using Newman-Keuls revealed that significantly different pairs divided only by condition, either manual or automated. Automation conditions were associated with lower tracking error than were manual conditions. There were no significant differences by group, that is, whether monitoring or fuel management was automated. Thus, for tracking performance, adaptive automation benefits were generalized and were realized irrespective of the function that was automated. However, automation benefits varied with the level of practice. Analysis of the simple effects of automation condition yielded significant differences only for block 1, $F(1,10)=6.04$, $p<.05$, but not for blocks 2, 3, and 4, $F(1,10)<2.37$, $p>.15$. Thus, automation reduced tracking error reliably on block 1, but not subsequently. Figure 6 shows that the largest difference in tracking error between the automated and manual conditions occurred in the first block, when either monitoring or fuel management was automated.

Monitoring Performance

Reaction Time. Monitoring RT was enhanced by automation of a second task following the transition from the manual to the automated condition, although the main effect for Automation Condition was only of borderline significance, $F(1,10) = 4.86$, $p<.06$. The main effect of Block on monitoring RT was significant, $F(3,30)=10.55$, $p<.01$, indicating that the speed of detection of engine malfunctions improved over time (see Figure 7). The Automation Condition x Blocks interaction was not significant, $F(3,30)=1.93$, $p>.14$. All other sources of variance were not significant.

Planned comparisons of monitoring RT in manual and automated conditions were made for each automation group using the Newman-Keuls procedure as before. Significantly different cells were distinguished only on the basis of automation condition. Automation was associated with lower RTs, for both tracking-automated and fuel-automated groups. Thus, as for tracking performance, monitoring RT was generally facilitated by adaptive automation. The simple effect of automation condition was significant for block 1, $F(1,10)=6.95$, $p<.025$, but not for blocks 2, 3, and 4, $F(1,10)<1.50$, $p>.25$. As Figure 7 indicates, monitoring RT was facilitated by automation of either tracking or fuel management, but the automation benefit was largest at the first block and statistically reliable only for that block. Figure 7 also shows that the improvement in RT with practice was greater for the manual than for the automated condition. Simple effects analysis showed that the Blocks factor was significant for manual blocks, $F(3,30)=11.37$, $p<.0005$, but not for automation blocks, $F(3,30)=1.53$, $p>.20$.

Detection Accuracy. The detection rate of system malfunctions improved with dynamic automation of another task, as shown in Figure 8. However, the main effect of Automation Condition was not significant, $F(1,10) = 3.14$, $p>.10$. Detection rate improved over blocks, $F(3,30)=3.98$, $p<.02$. The Automation Condition x Blocks interaction was not significant, $F(3,30)=1.71$, $p>.18$. All other sources of variance were not significant.

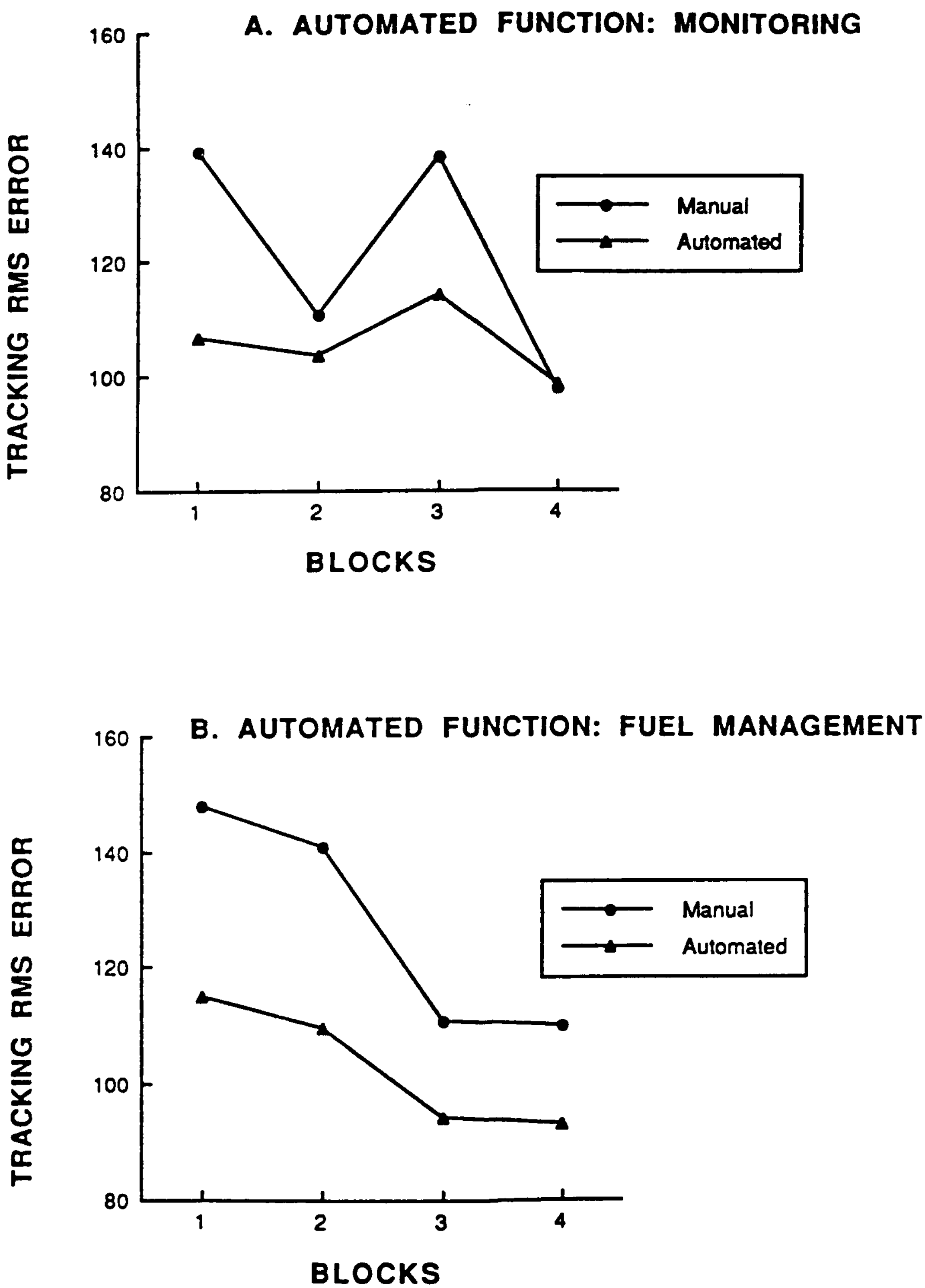


Figure 6. Effects of practice and adaptive automation of monitoring (A) and fuel management (B) on tracking performance

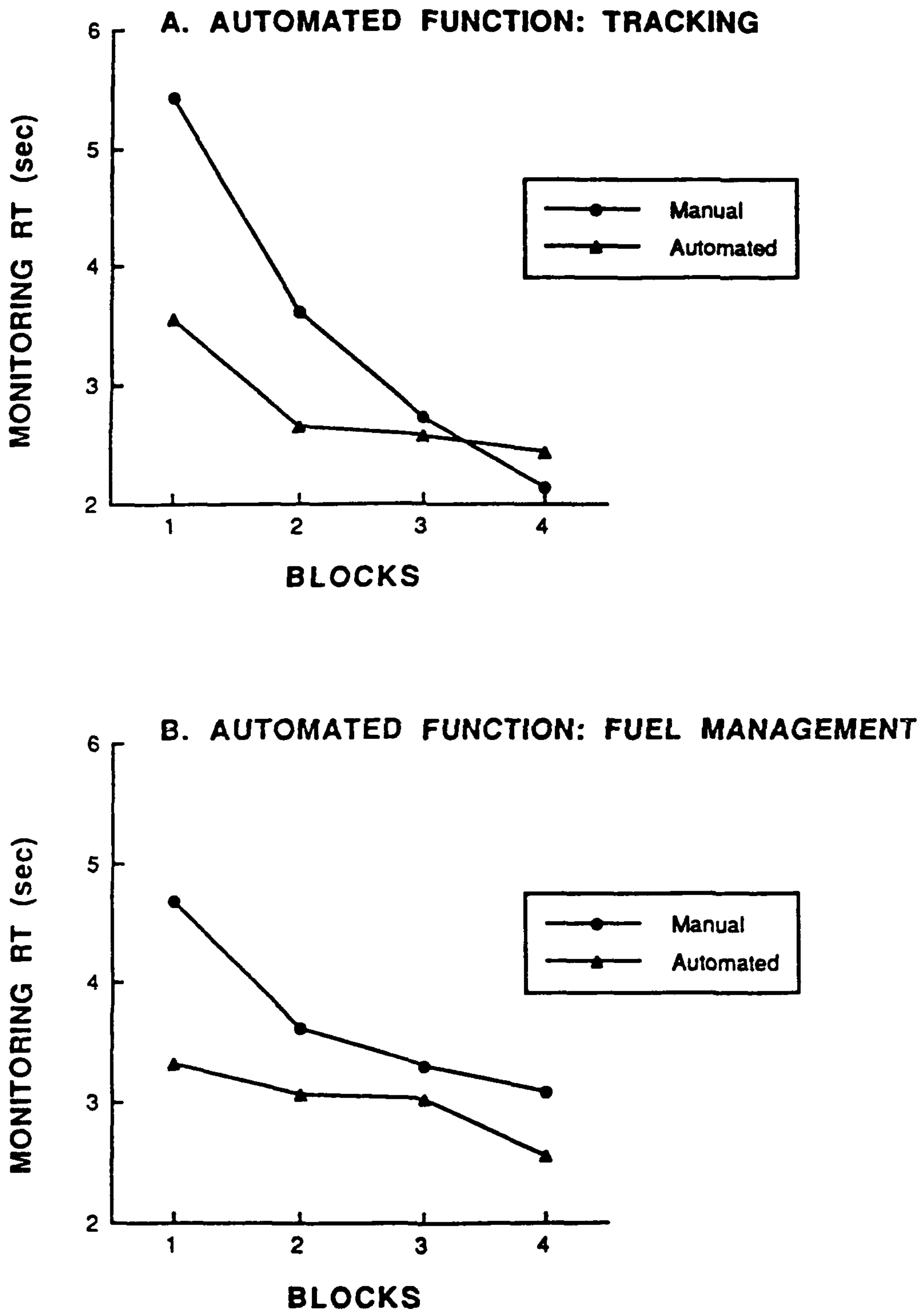


Figure 7. Effects of practice and adaptive automation of tracking (A) and fuel management (B) on monitoring RT

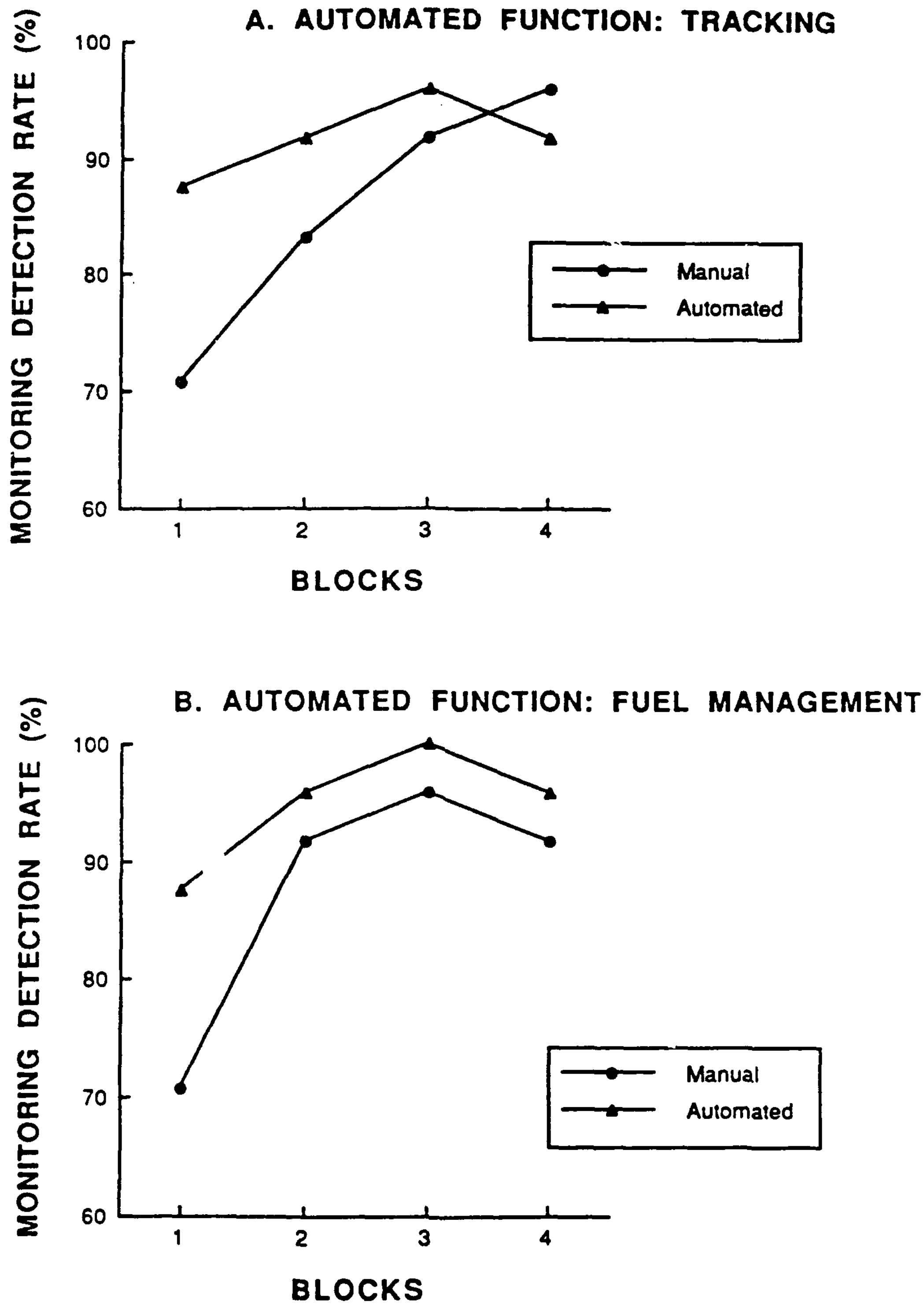
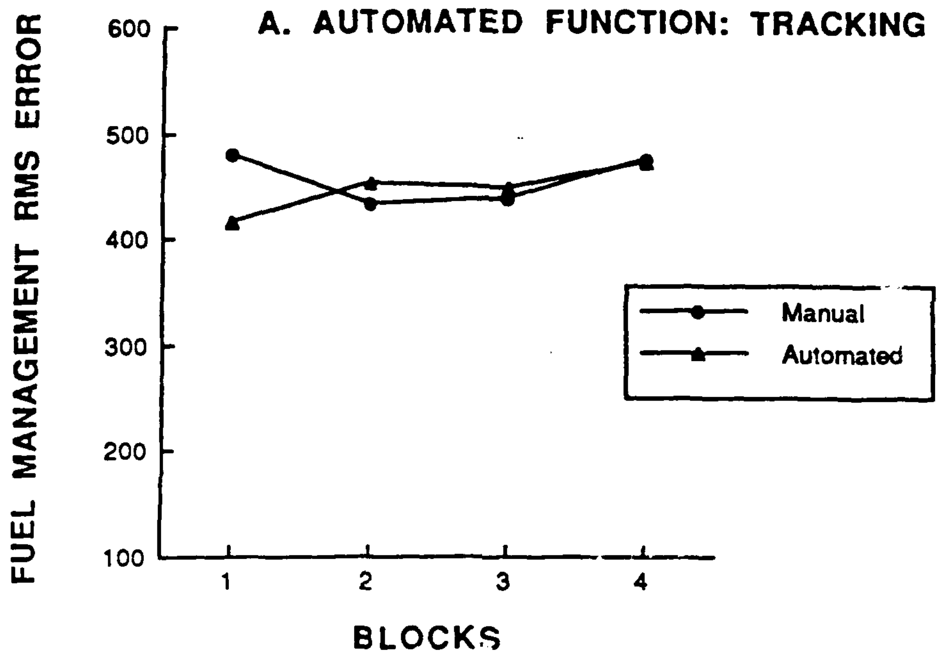


Figure 8. Effects of practice and adaptive automation of tracking (A) and fuel management (B) on monitoring detection accuracy



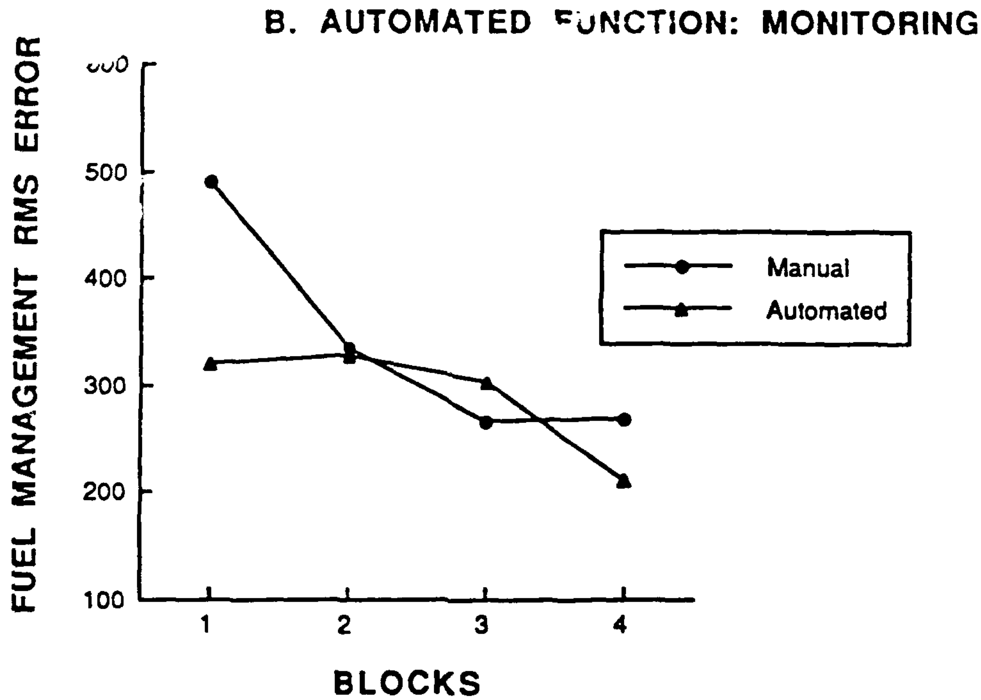


Figure 9. Effects of practice and adaptive automation of tracking(A) and monitoring (B) on fuel management performance

The simple effect of automation condition was of borderline significance for block 1, $F(1,10)=4.76$, $p<.06$, and not significant for blocks 2, 3, and 4, $F(1,10)<1$. Figure 8 shows that the automation benefit for detection rate was largest in the first block, as was the case for monitoring RT and for tracking error. Practice effects for detection accuracy were also the same as for RT: the simple effect of Blocks was significant for manual control blocks, $F(3,30)=4.22$, $p<.05$, but not for automation control blocks, $F(3,30)<1$.

Fuel Management Performance

The main effect of Automation Condition was not significant for fuel RMS error, $F(1,10)=1.72$, $p>.2$. The Blocks effects was also not significant, $F(3,30)=2.05$, $p>.1$. However, the Automation Condition x Blocks interaction was of borderline significance, $F(3,30)=2.49$, $p<.08$. The Automation Group x Blocks interaction was significant, $F(3,30)=3.35$, $p<.05$. All other effects were not significant.

The results of the multiple comparisons analysis for fuel RMS were not straightforward, insofar as the effect of automation appeared to differ by which second task was automated. For the tracking automation group, cells did not differ significantly on the basis of automation, while they did for the monitoring automation group. At either level of the automation condition, the monitoring automated group was significantly better than the tracking automated group in terms of fuel RMS error.

Simple effects analysis indicated that fuel RMS error was reduced by automation for block 1, $F(1,10)=5.27$, $p<.05$, but not for blocks 2, 3, and 4, $F(1,10)<1$. These results, which are similar to those for the other dependent measures, are apparent in Figure 9. The simple effect of Blocks was of borderline significance for manual blocks, $F(3,30)=2.9$, $p<.06$, but not significant for automation blocks, $F(3,30)<1$. Again, this is the same pattern as for the other measures.

Overall Benefits of Adaptive Automation

The overall benefits associated with adaptive automation were computed for each task by comparing mean performance levels on that task following automation of another task to performance without automation, averaged over blocks and automation groups. Table 1 shows the mean scores on each of the four dependent measures, under both manual and automation conditions. Performance as reflected in all four of the measures was consistently better under the automation condition when one of the other two tasks was automated (but when the automated task had to be supervised). The percentage benefit varied from abou: 8% to 19% for the different tasks and automation groups, and averaged 12.5%.

	Manual	Automation	Benefit (%)
Tracking RMS Error	124.3	104.5	15.9
Monitoring RT (sec)	3.57	2.89	18.9
Monitoring Hit Rate (%)	86.46	93.23	7.8
Fuel RMS Error	403.1	373.1	7.4

Table 1. Mean Performance Levels Under Manual and Automation Conditions

Overall benefits of adaptive automation were also assessed by computing normalized *composite performance* scores.This was done by first calculating the means and standard deviations of the four dependent measures across all the manual and automation condition blocks for all groups. The effects of automating a particular task on performance of other manual tasks were then assessed by computing the mean standard score for those manual tasks[2]. This was repeated for each of the three tasks that were automated. Figure 10 shows the results of these analyses. Automation of each of the three tasks led to consistent benefits on overall performance. There is also some indication that the maximum benefit was obtained by automating fuel management.

We also applied the composite performance analysis to examine changes in automation benefit with practice, which we previously studied using individual performance metrics. Figure 11 shows the results of this analysis. Two results are apparent in this figure. First, the automation benefit was reduced with practice. Second, practice improved performance more on manual control blocks than on automation control blocks. Both these results are consistent with those reported previously using the individual performance measures.

2 To enable meaningful averaging of standard scores across tasks, the signs of the standard scores for fuel RMS error, monitoring RT, and fuel RMS error were reversed so that a ***higher*** score reflected *better* performance, as was the case for the fourth dependent measure, detection accuracy.

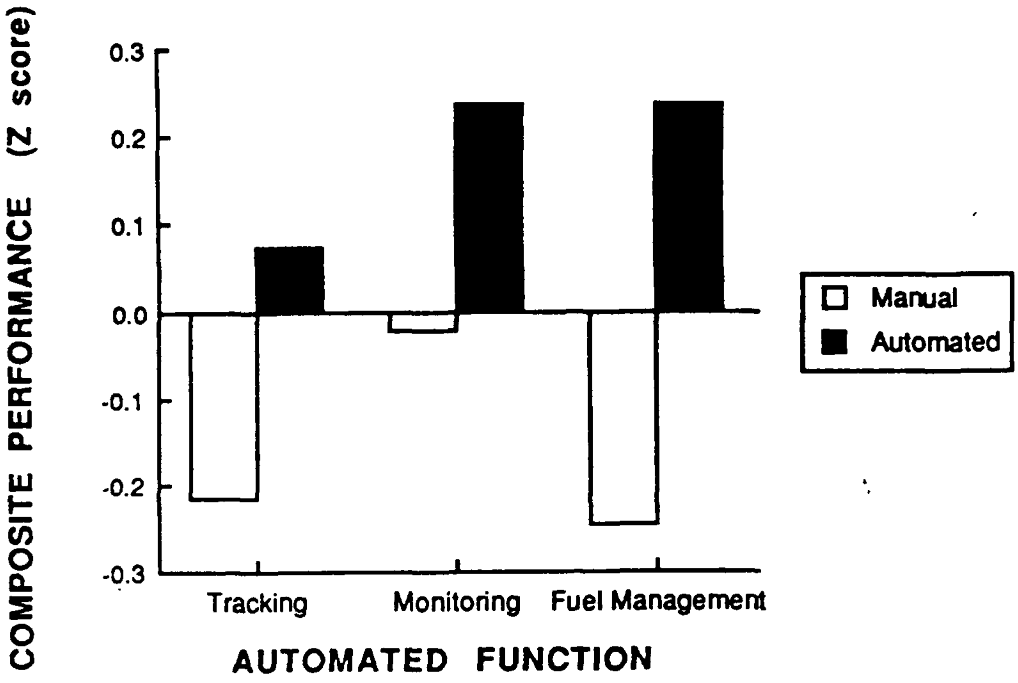


Figure 10. Effects of automation of different functions on composite performance.

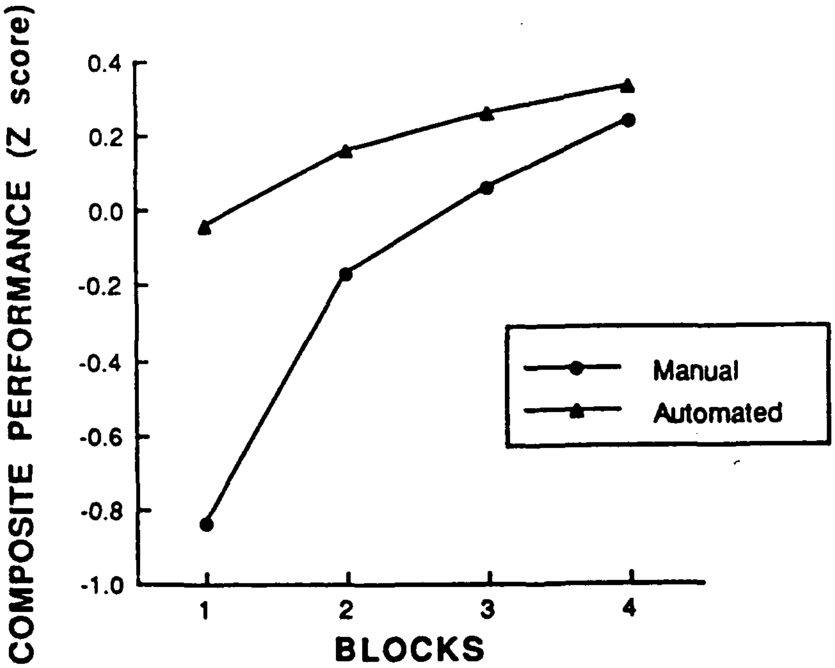


Figure 11. Changes in automation benefits with practice

Adaptive Automation Costs

Adaptive automation costs were evaluated by comparing performance in Manual {M} blocks with that in Return-to-Manual" {RM} blocks following a period of automation. The same 4 (Automation Group) x 2 (Automation Condition) x 4 (Blocks) ANOVA used in the examination of automation benefits was computed, except that the levels of the Automation Condition factor were Manual and Return-to-Manual rather than Manual and Automated. ANOVAs were computed for each of the four dependent measures. No significant effect of Automation Condition (Manual versus Return-to-Manual) was obtained for any of the four dependent measure: for tracking RMS error, $F(1,20)=3.91$, $p>.06$; for monitoring RT, $F(1,20)<1$; for monitoring accuracy, $F(1,20)=1.91$, $p>.18$; and for fuel management RMS error, $F(1,20)=1.03$, $p>.3$. Performance levels of individual tasks following automation did not differ from the normal manual performance levels for those tasks. Figures 12-15 depict the performance changes over blocks for each automation condition for tracking RMS error, monitoring RT, monitoring accuracy, and fuel-management RMS error, respectively. As these figures indicate, the major source of variance was blocks, and no consistent evidence of performance decrement following the transition from automation to manual performance was obtained.

Significant practice effects were obtained for each of the four dependent measures. The main effect of Blocks was significant in each case: for tracking RMS error, $F(3,60)=3.386$, $p<.05$; for monitoring RT, $F(3,60)=4.146$, $p<.01$; for monitoring hit rate, $F(3,60)=5.182$, $p<.01$; and for fuel RMS error, $F(3,60)=4.727$, $p<.01$. As Figures 12-15 show, performance generally improved with practice in both manual blocks and return-to-manual blocks. ANOVA of fuel RMS error revealed a significant Automation Group x Block interaction, $F(3,60)=2.19$, $p<.05$. This interaction resulted because practice effects were greater when some tasks but not others were automated. The simple main effect of Block was significant for both the monitoring $F(3,57)=3.19$, $p<.05$ and the control automation groups $F(3,57)=7.01$, $p<.005$, but not for the tracking and fuel management groups, $F(3,57)<1$.

The control group, which did not perform under automation, served as a true control only for the comparison between manual and the return-to-manual conditions. ANOVA revealed no differences between these conditions. If differences had been obtained, we had planned to compare them to performance changes associated with returning to manual performance following unrelated interpolated activity (i.e., performance of the control group). This was so that we could ascribe any automation-related performance decrement to automation per se rather than to interruption of manual performance by any other activity in general. However, no evidence of automation deficit was obtained in the experimental groups. As a result, there was no need to draw further comparisons between the experimental and control groups.

Overall costs associated with adaptive automation were assessed in the same manner as was done for automation benefits. Table 2 shows the mean values of the four dependent measures under both manual and return-to-manual conditions, averaged over all automation groups. As this table indicates, no evidence of costs associated with adaptive automation was obtained. In fact, across all measures, there was a mean *improvement* (rather than cost) of 3.75%, a value sufficiently close to 0% to suggest that no costs were incurred.

	Manual	Return-to-Manual	Cost(%)
Tracking RMS Error	135.39	127.85	- 5.6
Monitoring RT (sec)	3.79	3.47	- 8.4
Monitoring Hit Rate (%)	86.20	83.07	3.6
Fuel RMS Error	464.44	442.85	- 4.6

Table 2. Mean Performance Levels Under Manual and Return-to-Manual Conditions

DISCUSSION

This first empirical study of the effects of practice on the benefits and costs of short-cycle adaptive automation has yielded five major sets of results. First, at a methodological level, we were relatively successful in simulating the requirement for supervisory control that is typical of automated functions in real settings and which is reported to be an added source of workload (Sheridan & Farrell, 1984; Wiener, 1988). Subjects were asked to note (covertly) slight "deviations" (not malfunctions or failures) in the performance of the automation routine at the end of a series of blocks. This provided a rough evaluation of the ability of subjects to monitor an automated function while performing other manual tasks. Subjects reported about 40% of all "deviations" in automated functions in post-session queries, indicating that they maintained supervisory control of automation as required. This methodological control was important because it allowed us to determine the impact on performance of *automation* per se, rather than task removal (as would be the case in dual-task or multiple-task studies in which one task is removed and performance of the remaining task or tasks is analyzed).

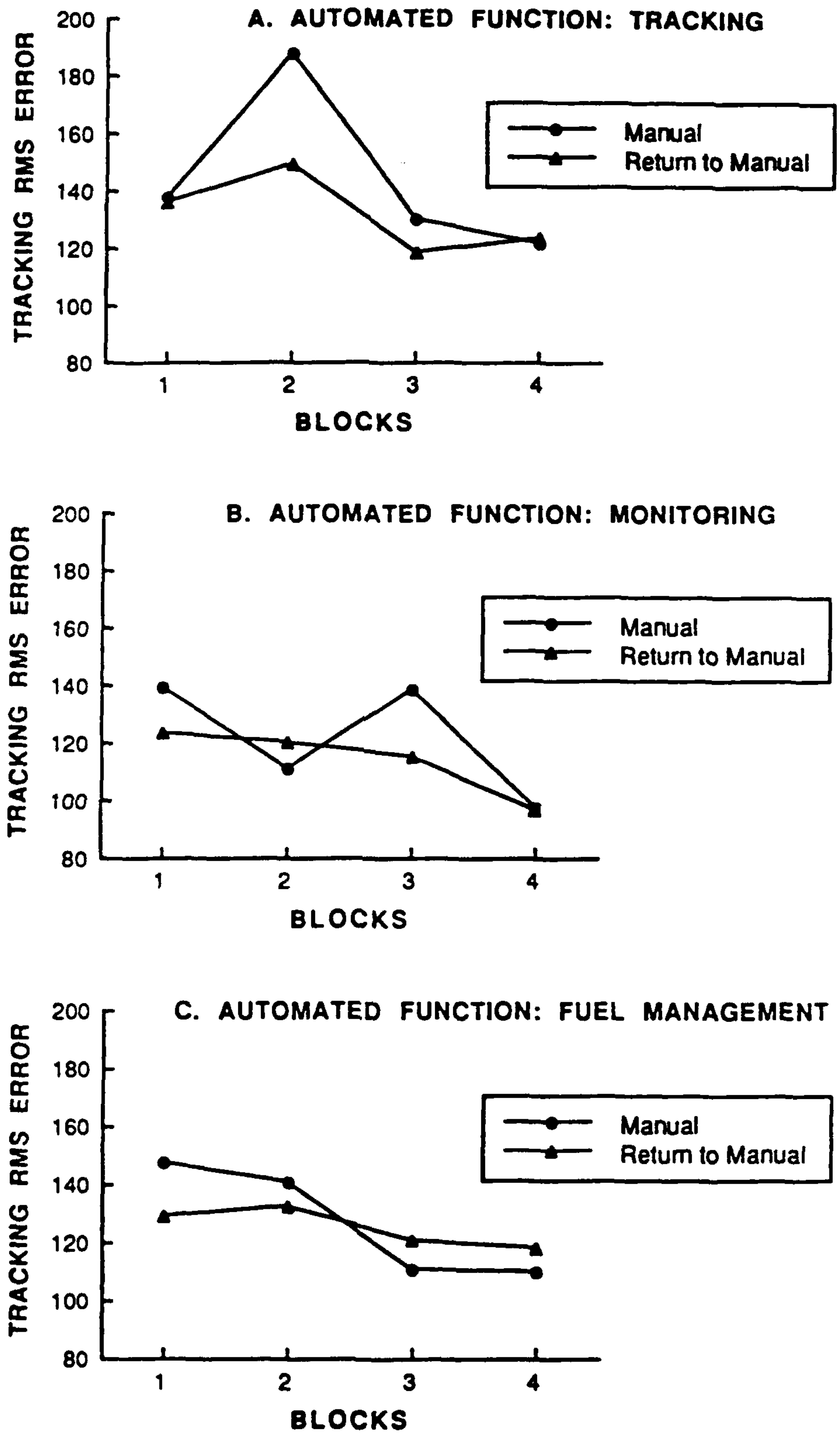


Figure 12. Effects on tracking performance of returning to manual control following automation of tracking (A), monitoring (B), and fuel management (C).

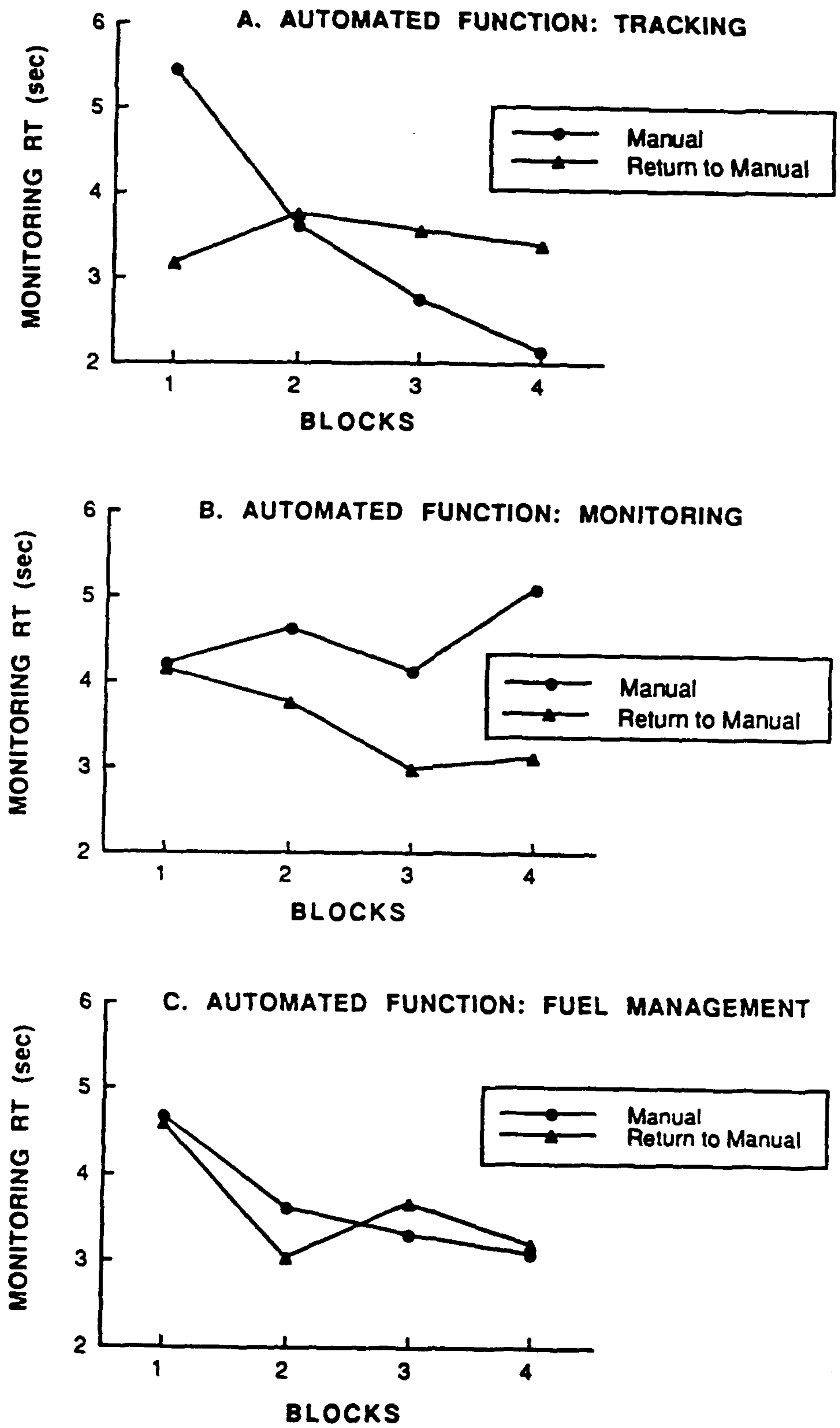


Figure 13. Effects on monitoring RT of returning to manual control following automation of tracking (A), monitoring (B), and fuel management (C).

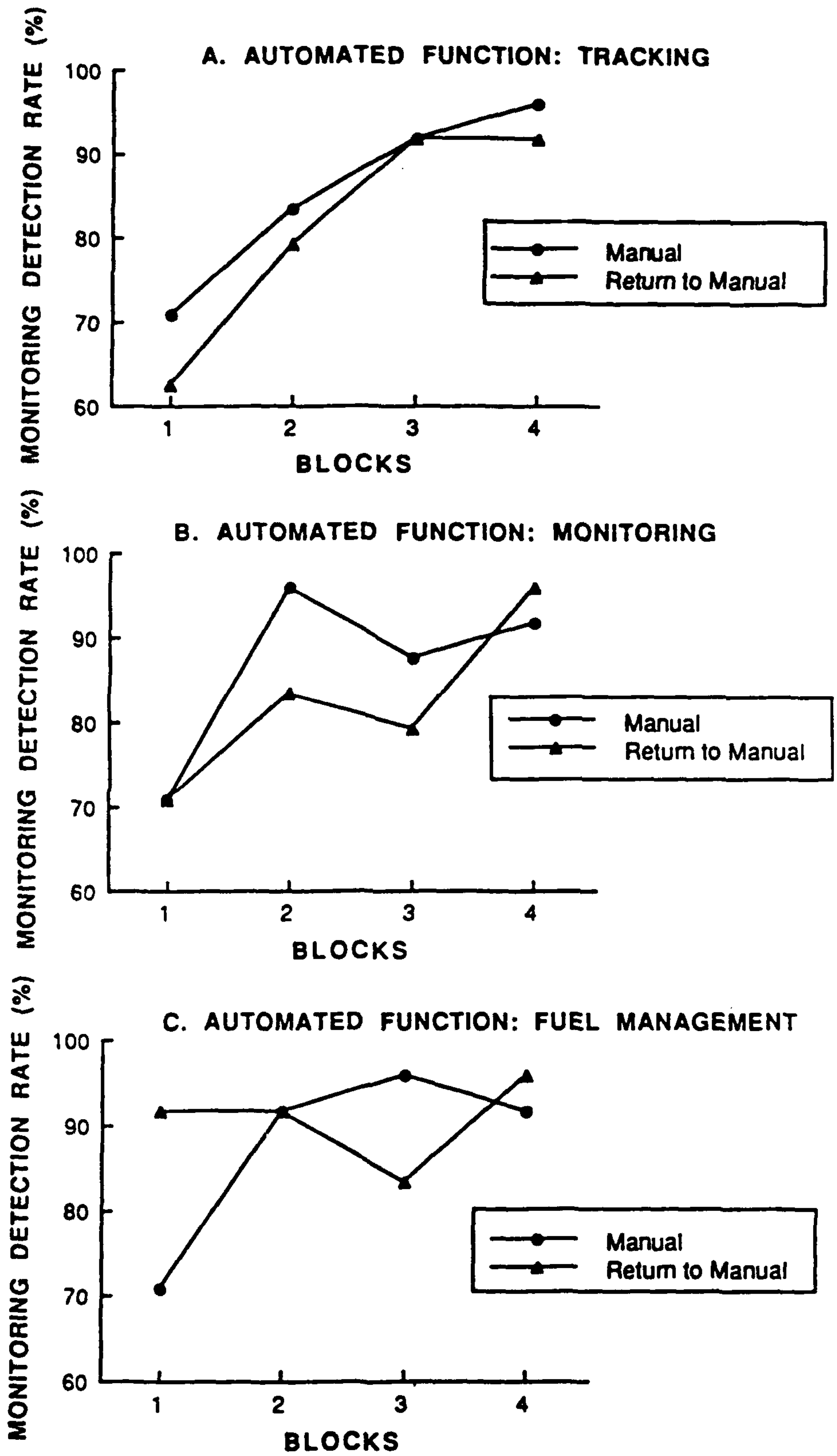


Figure 14. Effects on monitoring accuracy of returning to manual control following automation of tracking (A), monitoring (B), and fuel management (C).

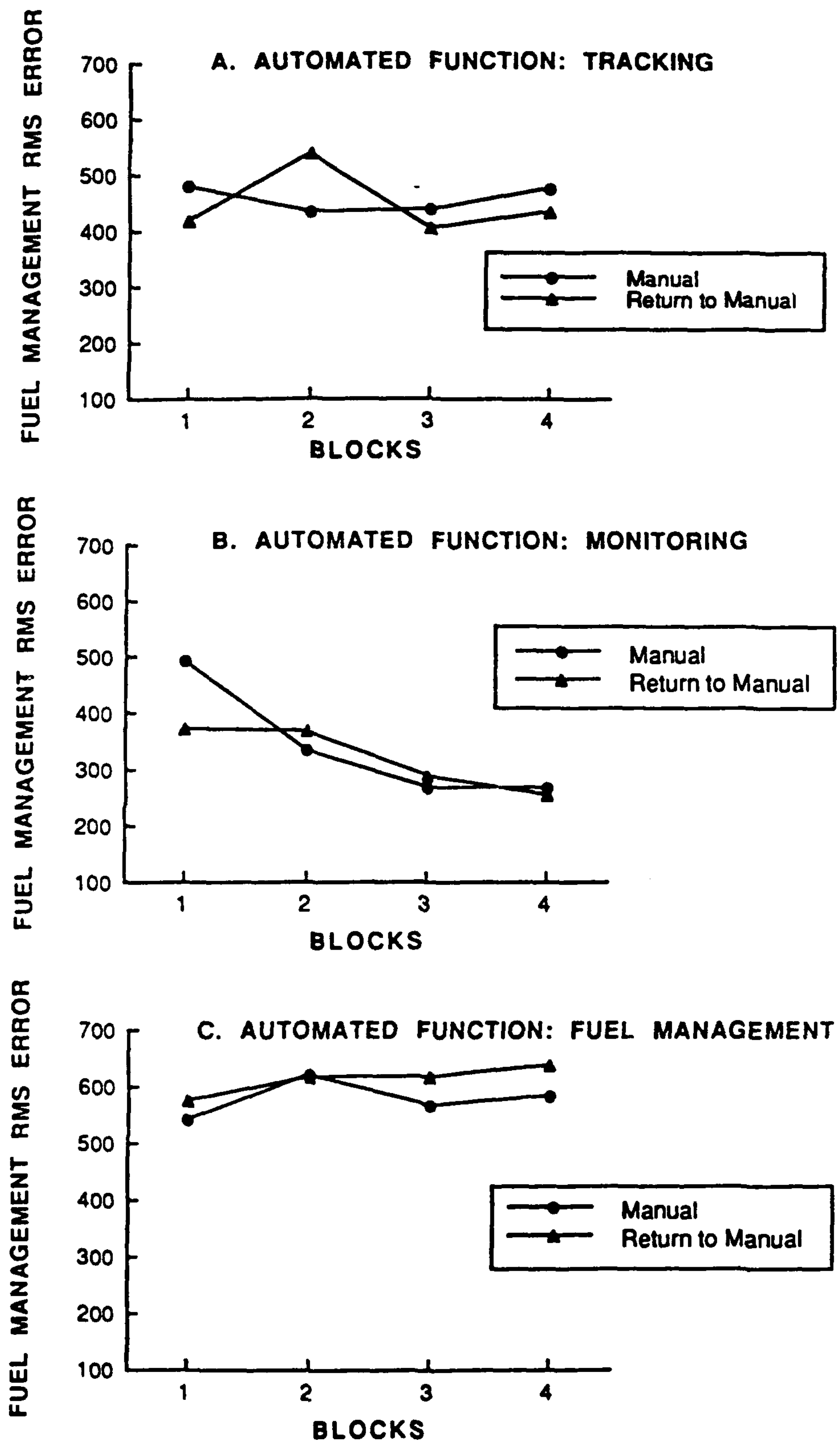


Figure 15. Effects on fuel management performance of returning to manual control following automation of tracking (A), monitoring (B), and fuel management (C).

Second, the results provide clear evidence that automation of a previously manual function (on the basis of some, unspecified, adaptive procedure) enhances performance on the remaining tasks performed manually, at least temporarily. These effects are not simply the result of task "subtraction" as in multiple-task studies (e.g., Tsang & Johnson, 1989), because we required subjects to supervise the automated function, and, as indicated by the "deviation" report percentages, subjects were able to do this satisfactorily. Thus, adaptive automation benefits were obtained even with the additional workload imposed by the requirement for supervisory control of the automated function. Thus, short-cycle adaptive automation can have beneficial effects on performance. The benefits appear to be general across tasks and conditions, at least for the set of flight-related tasks examined in the present study. Clearly, automation of a previously-manual function will impact on performance only to the extent that the function is resource sensitive (Norman & Bobrow, 1975; Wickens, 1984), or if automation frees up input or response channels that would otherwise be tied up (Navon, 1984). In a previous technical report we showed that the three tasks used in the present study--tracking, system monitoring, and fuel management--are mutually resource sensitive and interfere with one another when performed simultaneously (Parasuraman et al., 1991a). These results are consistent with the present finding that performance of a given function, e.g., tracking, is facilitated by automation, irrespective of which function is automated, e.g., monitoring or fuel management.

Third, although significant benefits of adaptive automation were obtained, the benefits dissipated with practice. For each of the three functions (and for all four of the performance measures for these three functions), automation benefits were consistently reliable only in the first 30-minute session and not subsequently. It is possible that a ceiling effect could be responsible for the performance advantage for the automation condition to decline with practice. However, this cannot be the only factor. For example, while a performance ceiling could have limited the automation benefit for monitoring accuracy (which reached ~100% in block 3), monitoring RT had not reached a floor in the later blocks in which no automation benefit was obtained. Furthermore, there was no consistent evidence that subjects had reached performance asymptotes after the four sessions (120 min) of practice that they received. If confirmed, the finding that performance benefits are transient could point to a possible limitation in the efficacy of adaptive automation, although at present we are unsure of the generality of this result and its underlying mechanism.

Whatever the underlying factors, the reduction of automation benefits with practice suggests the importance of training for obtaining optimal benefits from adaptive automation. A fourth result of the present study was that while practice led to an improvement of performance on all tasks under manual control conditions, performance under automation control conditions showed less or no improvement. For example, subjects showed no improvement with time in their ability to perform fuel management when either tracking or monitoring was automated. Again, although ceiling effects could have been a contributory factor (i.e. performance on automation blocks could not get any better), they could not be the only factor. What other factors could be responsible? One possibility is the type of training subjects received. In the present study initial training was followed by training in the form of simple practice. However, all subjects were initially trained under manual control conditions and were not specifically trained under automation control conditions. Repetitive practice alone may provide insufficient training when performing multiple tasks in which one task is automated (and must be supervised). This line of reasoning would suggest that training subjects specifically to monitor and supervise automated tasks may be required if the full benefits of adaptive automation are to be realized. This type of training may also lead to other benefits. For example, training may reduce operator "complacency" in monitoring for automation failure (Parasuraman, Molloy, & Singh, 1991c).

The fifth major result concerned possible costs of adaptive automation. No evidence was obtained for impaired efficiency associated with the return to manual control following automation. Practice reduced automation benefits, but extended performance did not produce any automation costs and practice did not influence costs. The use of a short-cycle adaptive automation schedule of the form {M} {A} {RM} allowed us to examine whether performance on the {M} and {RM} blocks, where identical tasks were performed, differed in any respect. Analysis revealed no such differences other than practice effects. Furthermore, even if performance on a {RM} block were found to be poorer than on a {M} block, it would have to be shown that such a decrement were due to the interpolation of the automation block {A} rather than to *any* interpolated activity. To this end we included a control group that performed an unrelated choice RT task in between the {M} and {RM} blocks. However, because no automation deficit or costs were found for the experimental groups in the first place, comparison with the control group was not needed.

If confirmed, the finding of no adaptive automation costs following the transition from automated to manual control would point to the superiority of adaptive automation over conventional, nonadaptive automation. A number of previous studies of nonadaptive automation (in which the set of manual and automated tasks remains fixed) have examined whether automation results in an "automation deficit"--i.e., a reduction in manual performance of a task when one aspect of the task is automated (Bortolussi & Vidulich, 1990; Fuld, Liu, and Wickens, 1987; Idasak,& Hulin, 1989; Kibbe & Wilson, 1989; Wickens & Kessel, 1981). For example, Bortolussi & Vidulich (1990) and Wickens & Kessel (1981) found that manual detection of deviations in tracking dynamics was impaired when tracking was controlled by automation rather than manually by the operator. However, other studies found no evidence of automation deficit. For example, Kibbe & Wilson (1989) found no effects of automation on the retention of information from a threat-detection display following completion of performance under automation or manual control. Although the literature is inconsistent, one generalization that can be made is that automation deficits are found for tracking and other motor tasks. In the present study, adaptive automation of tracking did not result in performance costs following the return to manual control.

In conclusion, although preliminary, these findings are encouraging with respect to the positive effects of adaptive automation. It remains to be seen, however, whether adaptive automation benefits persist with long-cycle adaptive automation and whether there are any costs associated with long-term automation transitions. The results also point to the importance of training for adaptive automation. Simple practice is an insufficient training method. Training under conditions of automation control may be necessary to optimize automation benefits.

REFERENCES

Barnes, M., & Grossman, J. (1985). *The intelligent assistant concept for electronic warfare systems*. China Lake, CA. NWC December (NWC TP 5585).

Bortolussi, M.R., & Vidulich, M.A. (1990). The benefits and costs of automation in advanced helicopters: an empirical study. *Proceedings of the Fifth International Symposium on Aviation Psychology*.

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.

Fuld, R.B., Liu, Y. and Wickens, C.D. (1987). The impact of automation on error detection: Some results from a visual discrimination task. *Proceedings of the Human Factors Society, 31st Annual Meeting*, 156-160.Santa Monica, CA: Human Factors Society.

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.

Hancock, P.A., Chignell, M.H., & Lowenthal, A. (1985). An adaptive human machine system. *Proceedings of the IEEE Conference on Systems, Man and Cybernetics*, 15, 627-629.

Idaszak, J.R., and Hulin C.L. (1989). *Active participation in highly automated systems: Turning the wrong stuff into the right stuff.* Technical Report, Aviation Research Laboratory. Institute of Aviation, University of Illinois at Champaign Urbana.

Kibbe, M.P., and Wilson, S.B. (1989). *Information transfer from intelligent EW displays.* Technical Report. Naval Weapons Center, China Lake, CA.

Lintern, G., & Gopher, D. (1978). Adaptive training: issues, results and future directions. *International Journal of Man-Machine Studies. 10,* 521-551.

LeMay, M. K., & Comstock, J. R. (1990). An initial test of a normative figure of merit for he quality of overall task performance. In *Proceedings of the Human Factors Society, 34th Annual Meeting,* pp. 81-85. Santa Monica, CA: Human Factors Society.

NASA (1989). *Aviation safety automation: A plan for the research initiative.*Washington DC: Office of Aeronautics and Space Technology, NASA.

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

Navon, D. (1984). Resources: A theoretical soup stone? *Psychological Review, 91*, 216-234..

Navon, D., & Gopher, D. (1979). On the economy of the human processing system. *Psychological Review, 86*, 214-255.

Noah, W. & Halpin, S.M. (1986). Adaptive user interfaces for planning and decision aids in C3I systems. *IEEE Transactions on Systems, Man and Cybernetics, SMC-16*, 909-918.

Norman, D. A. (1991). Cognitive science in the cockpit. *CSERIAC Gateway, 2 (2),* 1-6.

Norman, D. A., & Bobrow, (1975). On data-limited and resource limited processing. *Cognitive Psychology*, 7, 44-60.

Norman, S., Billings, C. E., Nadel, D., Palmer, E., Wiener, E. L., & Woods D. D. (1988). *Aircraft automation philoscphy: A source document.* NASA Ames Research Center. CA.

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

Parasuraman, R., Bahri, T., Molloy, R., & Singh, I. (1991b). Effects of shifts in the level of automation on operator performance. In *Proceedings of the International Conference on Aviation Psychology*, Columbus, OH.

Parasuraman, R., Molloy, R., & Singh, I. (1991c). *Performance consequences of automation-induced "complacency."* (Technical Report CSL-A91-2). Washington DC: Cognitive Science Laboratory, Catholic University of America.

Rouse, W.B. (1988). Adaptive aiding for human/computer control. *Human Factors, 30*, 431-438.

Sheridan, T., and Farrell, B. (1974). *Man-machine systems*. Cambridge: MIT Press.

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

Wickens, C.D. (1984). Processing resources in attention. In D.R. Davies & R. Parasuraman (Eds.) Varieties of Attention. Orlando: Academic Press.

Wickens, C.D. and Kessel (1981). Failure detection in dynamic systems. In J. Rasmussen and W.B. Rouse (Eds.) *Human detection and diagnosis of system failures*. New York: Plenum Press.

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

Winer, B.J. (1971). *Statistical principles in experimental design*. New York: McGraw-Hill.