(2)

DOT/FAA/AM-94/6

Office of Aviation Medicine
Washington, D.C. 20591

# Maintaining Vigilance on a Simulated ATC Monitoring Task Across Repeated Sessions

AD-A278 792

| Accesion For | |
|---|---|
| NTIS CRA&I | X |
| DTIC TAB | ☐ |
| Unannounced | ☐ |
| Justification | |
| By | |
| Distribution / | |
| Availability Codes | |
| Dist | Avail and / or Special |
| A-1 | |


D. J. Schroeder
R. M. Touchstone
Civil Aeromedical Institute
Federal Aviation Administration
Oklahoma City, Oklahoma 73125

J. A. Stern
Department of Psychology
Washington University
St. Louis, Missouri

N. Stoliarov
State Scientific Research Institute for Civil Aviation
Moscow, Russia

R. Thackray
Oklahoma City, Oklahoma




March 1994

Final Report

This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161.



U.S. Department of Transportation
Federal Aviation Administration

# NOTICE

| 1. Report No. | 2. Government Accession No. | 3. Recipient's Catalog No. |
|---|---|---|
| DOT/FAA/AM-94/6 | | |

| 4. Title and Subtitle | 5. Report Date |
|---|---|
| Maintaining Vigilance on a Simulated ATC Monitoring Task Across Repeated Sessions | March 1994 |
| | 6. Performing Organization Code |
| 7. Author(s) | 8. Performing Organization Report No. |
| David J. Schroeder, Robert M. Touchstone, John A. Stern, Nikolai Stoliarov, Richard I. Thackray | |
| 9. Performing Organization Name and Address | 10. Work Unit No. (TRAIS) |
| FAA Civil Aeromedical Institute/P.O. Box 25082<br>Oklahoma City, OK 73125<br><br>Department of Psychology/Washington University<br>St. Louis, MO 63130-4899<br><br>State Sci. Res. Institute for Civil Aviation<br>Moscow, Russia | 11. Contract or Grant No.<br><br>DTFA-02-90-R-00149 |
| 12. Sponsoring Agency name and Address | 13. Type of Report and Period Covered |
| FAA Office of Aviation Medicine<br>Federal Aviation Administration<br>800 Independence Ave., S.W.<br>Washington, D.C. 20591 | 14. Sponsoring Agency Code |

15. Supplemental Notes

This work was performed under task AM-B-HRR158.


16. Abstract

Maintaining alertness to information provided visually is an important aspect of air traffic controllers' work. Improper or incomplete scanning and monitoring behavior is often referred to as one of the causal factors associated with operational errors and deviations. This study was undertaken to assess changes in vigilance/attention across 3 separate days as subjects performed on an Air Traffic Control (ATC) simulation task. Information was gathered as part of a larger study of attention and "gaze control inefficiencies." Twenty paid subjects on 3 separate days monitored a simulated ATC task for 44 critical events over a 2-hour session. The complex monitoring task included the detection of: (a) altitude malfunctions; (b) aircraft conflict/ no conflicts where 2 aircraft were at the same altitude on an airway simultaneously; and (c) triangular targets representing VFR aircraft that appeared either centrally or peripherally on the screen during the course of each session. Changes in performance on the complex monitoring task associated with either time-on-task or repeated sessions were dependent on nature of the task. Performance on the component involving detection and decision-making (conflict/no conflict detection) evidenced a decrement associated with time-on-task on each of the 3 days. Improvement was evident from the first to the third day. Performance on the identification of the altitude malfunctions remained relatively immune to the effects of time-on-task or repeated sessions. Detection of the VFR aircraft intruders presented a mixed picture, with improvement noted in some aspects of performance but not in others, and some evidence of time-on-task effects. Outcomes were generally consistent with previous findings with this task and consistent with other literature with respect to the presence of performance decrements associated with time-on-task. The results were consistent with a view that the decrements are associated with lapses in attention or "blocks," rather than a generalized fatigue effect or a general modification in overall scanning behavior. Furthermore, the results suggest that there were aspects of monitoring performance that remain relatively immune to time-on-task effects, even during the course of the three 2-hour sessions.


| 17. Key Words | 18. Distribution Statement |
|---|---|
| Vigilance, Monitoring Performance<br>Attention, Air Traffic Control | Document is available to the public through the National Technical Information Service, Springfield, VA 22161 |

| 19. Security Classif. (of this report) | 20. Security Classif. (of this page) | 21. No. of Pages | 22. Price |
|---|---|---|---|
| Unclassified | Unclassified | 14 | |

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

# Maintaining Vigilance on a Simulated ATC Monitoring Task Across Repeated Sessions

The task of an air traffic control specialist (ATCS) in en route and tracon settings places heavy emphasis on monitoring auditory information exchanges between pilots and other controllers, visual monitoring of the planned view displays (PVDs), and information contained on the flight progress strips. Improper or inadequate scanning and monitoring behavior is frequently referred to as one of the causal factors associated with operational errors and deviations. One of the recommendations from the "Report of the Interagency Near Mid-air Collision Working Group" (1989) was for research to "Develop radar controller scanning procedures, identify effective techniques, and incorporate them into training programs." There is a clear need to identify and evaluate the factors most closely associated with changes in monitoring performance, given its importance in this work setting. Additionally, ongoing attempts to automate human-machine systems, including those involved in the control of air traffic, place a primary emphasis on the monitoring role of the human operator. This remains true despite a body of knowledge that identifies many of the difficulties that human operators experience when required to monitor displays for prolonged time periods. Thus, increased automation within air traffic control is likely to bring with it additional monitoring requirements.

Studies of the ability of the human operator to maintain attentiveness and vigilance while monitoring a visual display have generally been limited to the effects of several trials administered across a single day. Lindsley, et al. (1944), using a radar simulation task evaluated the ability of operators to detect signals for 4 hours a day across 3 weeks. Their results suggest that a loss of efficiency is not readily apparent until the third day. Webb and Wherry (1960), using an auditory monitoring task, found a decline *within sessions*, but no consistent between sessions effect for 3 subjects across 5 consecutive days. Adams, Humes, and Stenson (1962) developed a simulated semi-automatic air defense system using a digital computer display. Subjects were required to detect changes that occurred in an aircraft identification number across a 3-h time period for 9 consecutive days. Using 6 stimulus sources, a significant decline in detection times within sessions was observed, but differences across days were not significant. When the visual load was expanded to 36 stimulus sources, which produced a significantly more demanding task, Adams, Humes and Sieveking (1963) found the usual decline within sessions but an improvement in performance across the 9 consecutive days. Their outcomes suggest that learning may occur and that it is possible for individuals to develop a short term tolerance for boredom.

To date, these studies provide a relatively focused, and in many cases, a relatively non-complex stimulus situation. While Adams and his colleagues used a continuous monitoring task, most studies involve the administration of a discrete number of trials. A simulated ATC monitoring task, that was developed by Thackray, Bailey, and Touchstone (1979), provides for a more complex and challenging monitoring task. That basic task, with some modifications, has been used on numerous occasions to assess the effects of task load (Thackray and Touchstone, 1991 and Thackray and Touchstone, 1989), age (Thackray and Touchstone, 1981), a simulated emergency on performance recovery (Thackray and Touchstone, 1983), and visual task load on critical flicker frequency (Thackray and Touchstone, 1985). More recently, the task has been used to assess the utility of flashing or colored targets to increase target conspicuity (Thackray and Touchstone, 1991) and the use of colored VFR intruder targets with normal and color deficient subjects (Mertens, Thackray and Touchstone, 1992).

This study was designed to assess changes in performance on the ATC monitoring task across 3 separate sessions. Assessments were made of changes in performance on both primary and secondary tasks. This report analyzes the performance data during the 3 2-h monitoring sessions. Information concerning the gaze measures is contained in a separate report (Stern, Boyer, Schroeder, Touchstone, and Stoliarov, in press).

## Methods

### Subjects

Twenty paid volunteer university students were used as subjects (16 men and 4 women). They ranged in age from 18 to 29 years and none had prior experience with the simulated air traffic controller's task used in this study or previous ATC training. All subjects were right-handed and had 20/20 vision, or corrected to 20/20 vision. Each subject came in for testing on 3 different days spaced approximately a week apart; all subjects completed the 3 sessions in 5 to 15 days.

### Apparatus and Task Description

The basic experimental equipment consisted of a Digital Equipment Corporation (DEC) VS11 19-in (49-cm) graphics display, keyboard, and joystick, all of which were interfaced with a VAX 11/730 computer (DEC). The computer was used to both generate the input to the display and process subject responses. The VS11 was incorporated into a console designed to closely resemble an ATC radar unit. Two diagonal, nonintersecting flight paths were presented on the display, along which aircraft targets could move in either direction. A given aircraft's location was displayed as a small rectangle on the flight path, and an adjacent alphanumeric data block identified the aircraft and gave its altitude and ground speed. Ground speed remained constant for all aircraft, to prevent overlapping alphanumerics. Aircraft position and the altitude information contained in the data block were updated every 6 s. There were 8 aircraft on each flight path at all times; as 1 left the screen, another appeared.

In addition to the primary targets (rectangles and adjacent alphanumerics), the display also contained stationary "clutter" in the form of 4 mm diameter green dots. Green triangular targets were used to represent non-tracked aircraft, or unidentified aircraft (designated as VFR intruders), and were similar in size to the dots. The triangles could appear in any 1 of 8

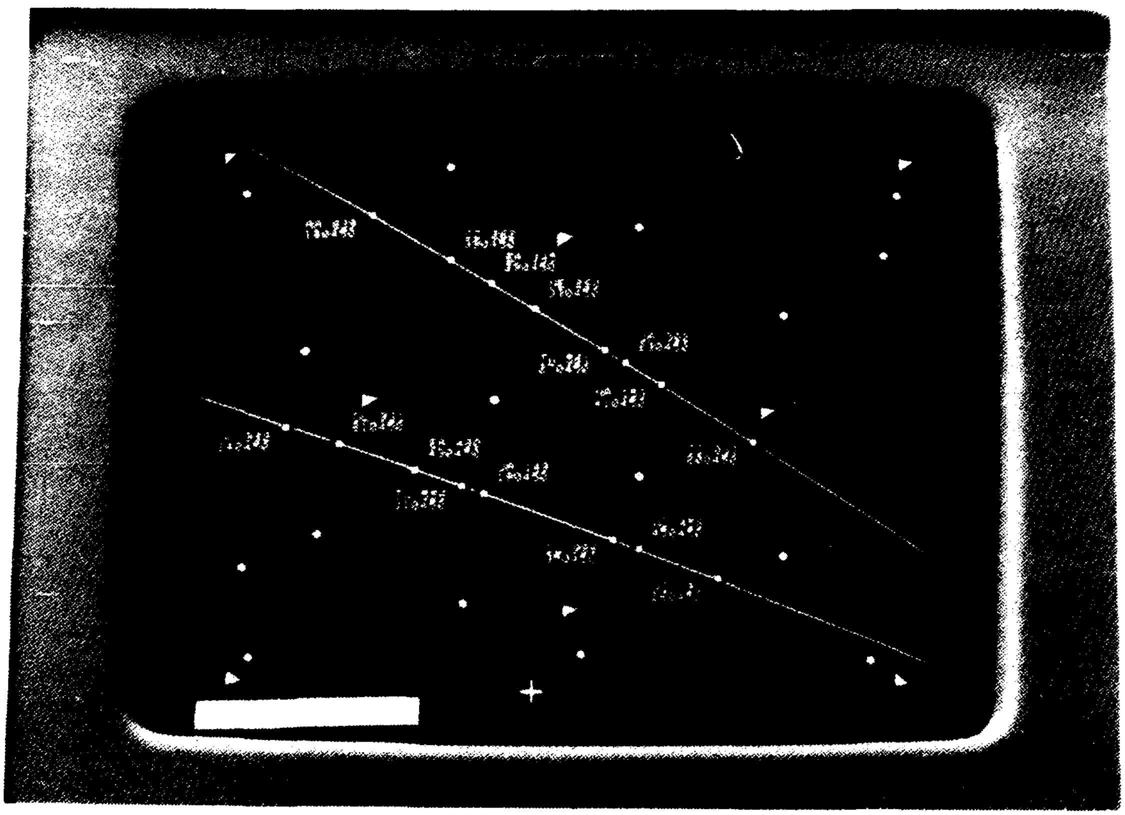


**Figure 1.** *Subject's display showing the clutter, location of each of the VFR intruders (triangles) and a typical air traffic condition with the alphanumeric targets.*

different locations on the screen; 4 locations were near the extreme corners of the display and 4 were more centrally located. All of the possible triangle locations, as well as the clutter and a typical target pattern as displayed in the subject's console, are shown in Figure 1.

The primary task consisted of monitoring for 2 types of critical events. The first was readily detectable and consisted of the appearance of 3 X's in place of the altitude information in the alphanumeric data block of a given target. It was explained to the subjects that the appearance of the 3 X's indicated that the altitude transponder on the aircraft had malfunctioned, resulting in a loss of altitude information. Upon detecting the 3 X's, the subject pressed the console button marked "XXX." The joystick and cursor were used to correct the malfunction.

The second type of critical event was more difficult to detect, in that it was not readily apparent. This type of critical event involved the occurrence of 2 aircraft at the same altitude and on the same flight path. As soon as this event was detected, the subject responded by pressing the console button marked "ALT CHECK." This, in turn, caused the 2 aircraft and alphanumerics to turn red. The subject then had to determine if the situation was 1 of 2 possible conditions. If the 2 aircraft at the same altitude were moving toward each other, this was a conflict situation. In this case, the subject responded by pressing the button marked "CONF." Using the joystick the subject moved the cursor over 1 of the conflicting aircraft and pressed a button beside the joystick. This activated a computer assist program which, in the lower left corner of the display, displayed the identification of the aircraft and suggested a new altitude. The subject, upon determining that the computer-assigned altitude did not place the aircraft in conflict with any other aircraft on that flight path, pressed the "ENTER" key on the keyboard. The new altitude replaced the conflict altitude and the color of the aircraft and alphanumerics returned to green.

A *no conflict* condition was one where the 2 aircraft with the same altitude were either flying away from each other or were headed in the same direction. When the *no conflict* button was pressed the color of the 2 aircraft and alphanumerics returned to green, although the conflict altitude remained for 60 s after the button press before changing, so each aircraft had a different altitude. The subject was informed that during this period, before the altitude changed, they had to remember that they had already responded to the *no conflict* situation. The subject was informed that if he/she forgot and responded again to the same situation it was counted as a memory error.

If the subject failed to respond to the conflict or no conflict situation within 26 s, a "conflict alert" occurred indicating a possible conflict. The "conflict alert" consisted of the 2 aircraft at the same altitude flashing at a 1 Hz/s rate and the presentation of a 60 Hz/s, 65 dBA tone pulsing at a 2-s rate.

The secondary task consisted of detecting the triangles that represented VFR intruders. When one of these triangles was detected, the subject responded by pressing the console button labeled "VFR INTRUDER" and the triangle would disappear. The triangles had a 90-s time-out period, and if not detected and responded to within that time, would disappear.

Nine primary task (4 conflict, 4 no conflict, and 1 XXX) and 2 secondary task critical events were presented during each 30-m period of the 2-h run. Only 1 critical stimulus was presented at a time. The time of each primary and secondary critical event was fixed, but the location of each of the secondary events was randomly assigned without replacement for each subject. The utilization of these stimulus conditions, based on data from Thackray and Touchstone (1991) resulted in a relatively "high" task load with use of the most difficult stimulus condition for detection of the secondary targets (non-flashing green target).

Two Sony CCD cameras were used to monitor the subjects and the visual display. The outputs of the 2 cameras were combined by using a special effects generator, and displayed on a video monitor. A small indicator light, located above the display, but out of the subject's line of sight, illuminated each time a critical event was presented.

Beckman Biopotential Skin Electrodes were attached above and below the right eye to record vertical eye movements and blinks with electrodes attached at the outer canthi of each eye, for recording horizontal eye movements.

Head movements were recorded using an adjustable inner liner from a construction worker's helmet. A strip of balsa wood with 4 LEDs mounted at the tip was attached to the liner at the back of the subject's head, so that it was at a 90-degree angle to the shoulders when the subject faced the display. Photo receptors were attached on each shoulder and the output was amplified and combined to produce a measure of head movement in the horizontal plane.

Eye movements, head movements, onset of critical stimuli, and subject's responses were recorded on a Beckman Dynograph and the analog signal recorded on a Kyowa RTP-600B magnetic tape recorder. Information concerning the analysis of the eye and head movement data is contained in a report by Stern, et al. (in press). The performance data were recorded on the DEC 11/730 computer.

### Procedure

Subject testing was run at the rate of 2 per day (1 in the morning, the other in the afternoon). Upon arrival, the subject filled out a pay voucher, received a detailed description of the research task, and signed a consent form. He/she then completed the NEO-Personality Inventory. The EOG electrodes were attached, the head movement helmet was fitted, and photo receptors attached on top of the shoulders. After calibrating the physiological recordings, the subject was given a detailed description of each critical event and the appropriate responses to each. A short practice session followed each event description. A second practice session consisted of the presentation of all of the different types of stimuli combined as they would appear in the experiment. Following the second practice session, the subject completed a 9-point subjective rating scale assessing present feelings of attentiveness, fatigue, strain, boredom, and annoyance.

A tape recording of background noises recorded in an actual en route air traffic control radar room was played continuously during the 2-h experiment. The sound level of this noise at the subject's head location was 62dBA.

At the end of the 2-h test session, the subject completed a second version of the rating scale that contained additional items related to perceived task difficulty and the effort required to maintain task concentration.

**TABLE 1**

**Mean detection times across 4, 30-minute periods for altitude malfunction events**

| | PERIOD | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| DAY 1 | 12.2 | 19.1 | 9.6 | 6.2 |
| 2 | 7.0 | 7.6 | 9.2 | 5.2 |
| 3 | 9.5 | 8.4 | 9.2 | 8.9 |

### Data Analysis

Performance on the ATC monitoring task was determined with respect to the time required to detect each event and the number of missed events. Means were determined for both detection times and number of missed events for each 30-m period during each 2-h session. ANOVAs were conducted to determine the effects of periods and days on performance for both the number of missed events and detection times. Separate analyses were conducted for the conflict/no conflict, altitude and VFR intruder events.

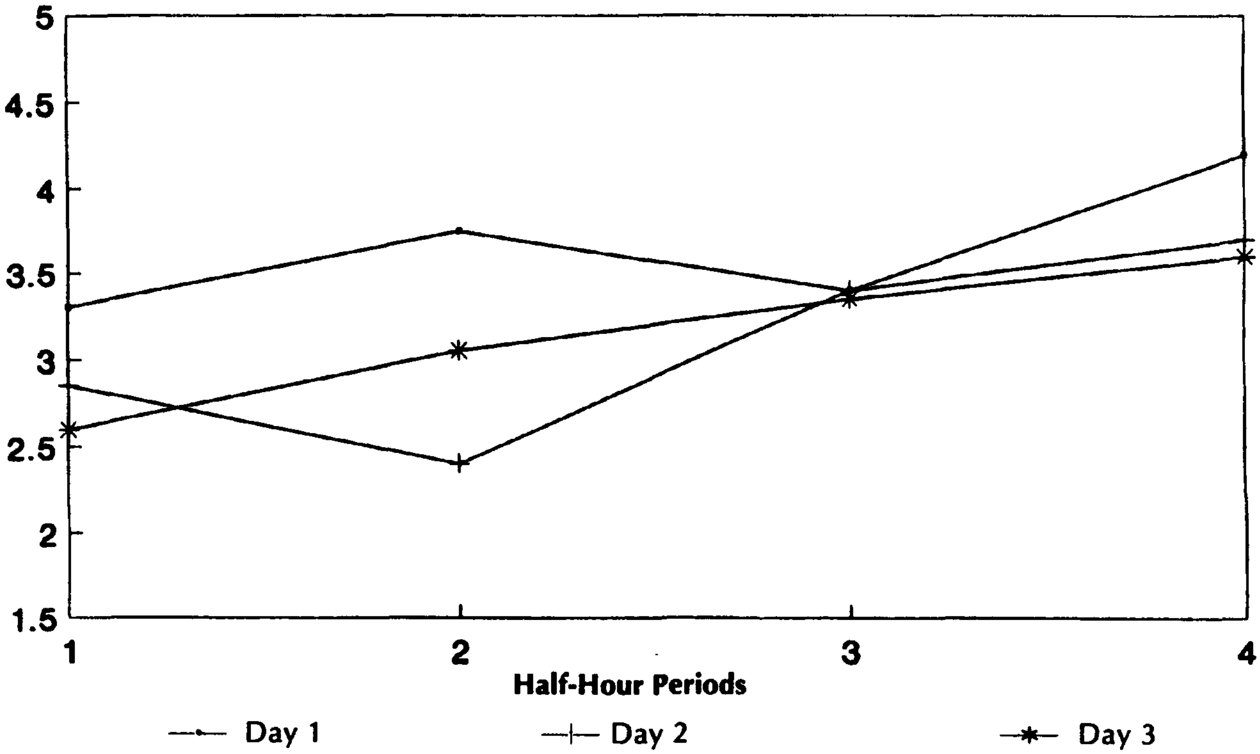
Since test times for subjects were confounded by gender and time of day (morning or afternoon), separate ANOVAs were conducted to determine if there was a differential effect of gender or time of day on performance. Since none of these analyses was statistically significant, those variables were not included in the overall analysis.

## RESULTS

### Primary Task Performance

Table 1 contains the average detection times for the altitude malfunctions across the 3 days for each of the 4 30-m periods. Subject detection times declined from an average of 12.2 during the first 30-m period to 6.2 for the final period on the first day. While some improvement was noted across periods on day 2 (7.0 s to 5.2 s), average detection times for the final day

**Figure 2.** *Average number of missed conflict/no-conflict events during each of the 4 half-hour periods.*

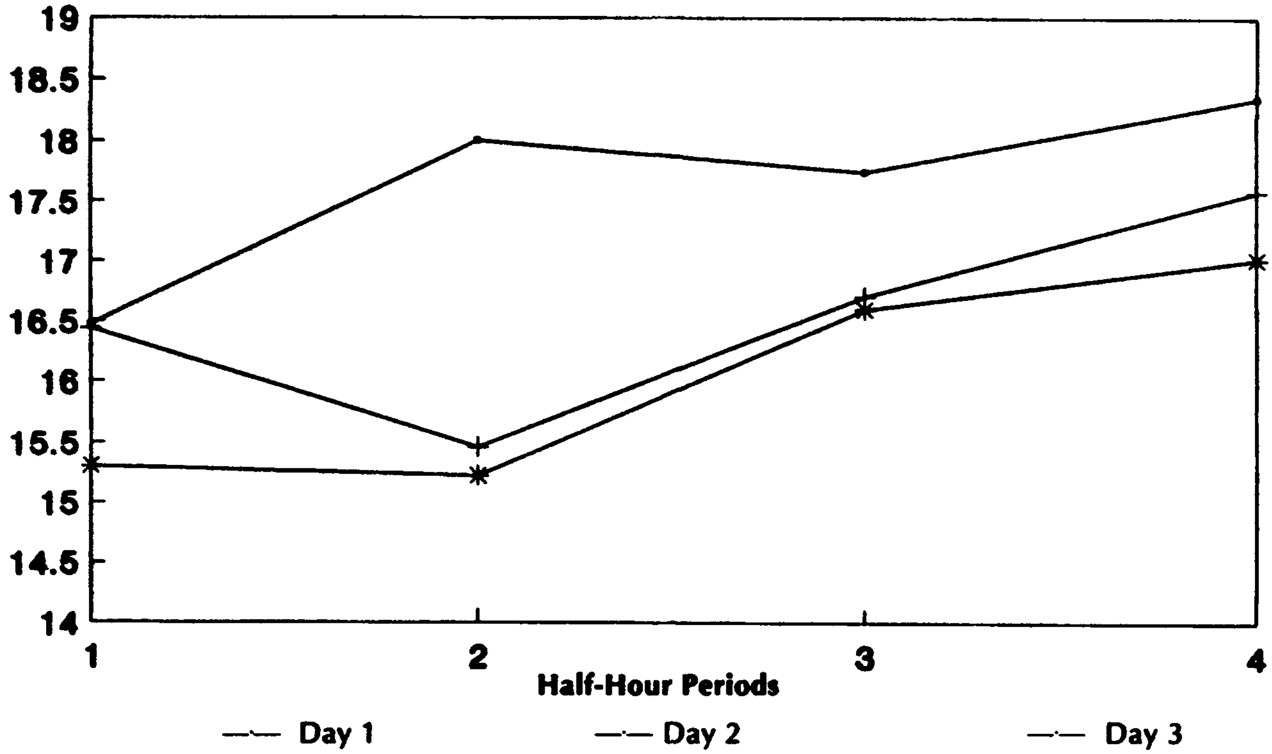



remained between 8.4 and 9.5 s. A repeated measures ANOVA revealed that the differences across days neared statistical significance ($F(2,38) = 2.75$, $p<.08$).

Average frequency of missed conflict/no conflict events during each of the 4 30-m periods, for each of the 3 days are presented in Figure 2. These values reflect the average number of conflict/no conflict events subjects failed to detect within the 26 s time period prior to the onset of the "conflict alert" alarm. Results from the ANOVA indicate a significant effect for period ($F(3, 57) = 4.47$, $p<.01$) and an effect for days that approached significance ($F(2, 38) = 2.59$, $p<.09$). When compared to the results of Thackray and Touchstone (1991), the higher averages for the missed events in this study appear related to the fact that there were 2 additional conflict/no conflict events presented during each of the 30 m sessions. For each of the 3 days, the number of missed events during the final period was higher than that of any of the previous periods. Overall, subjects missed fewer conflict/no conflict events on days 2 and 3 than during the initial day. Comparisons of the average number of "missed" events for day 1 and days 2 and 3 reveal greater differences during the first half, rather than the second half, of each 2-h session.

Mean times to detect the conflict situations for each period and day are presented in Figure 3. A value of 26 s was included in determining the average detection time for all occasions where a subject failed to detect the conflict/no conflict prior to the onset of the "conflict alert." A repeated measures ANOVA revealed significant effects for periods ($F(3,57) = 3.95$, $p<.01$) and days ($F(2,38) = 3.21$, $p<.05$). Thus, detection times generally increased from the first through the final period of each day. Detection times also improved from day 1 through day 3. However, if the detection times are compared without the inclusion of the 26 s for each of the missed events, there is little evidence of an increase across the 2-h session. Mean detection times for the first hour and second hour on the first day were 13.31 and 13.28, 13.48 and 13.10 on the second day, and 12.45 and 12.79 on the third. The ANOVA revealed no signficant changes across the 2-h session or across days in detection time when the 26-s

**Figure 3.** *Average time required (in seconds) to detect the conflict/no-conflict events.*




times were not included in the analysis. The increase in detection time evident in the initial comparisons was due solely to the inclusion of the longer detection times assigned for the "missed" events; detection times for identified events did not change across each session.

## Secondary Task Performance

To determine the average detection time for the VFR intruders, each missed event was assigned a value of 90 s. This procedure was consistent with that used by Thackray and Touchstone (1991). Due to the small number of missed events and because the "inner" and "outer" location of the events was not consistent within each 30-m period, statistical comparisons for the VFR events are based on data for the entire session. Data concerning the total number of missed VFR events and the mean detection times, for each of the 3 days, appear in Table 2. The number of missed VFR events declined from 12 on the first day, to 7 and 8 on days 2 and 3, respectively. That change is due primarily to the reduced

**TABLE 2**

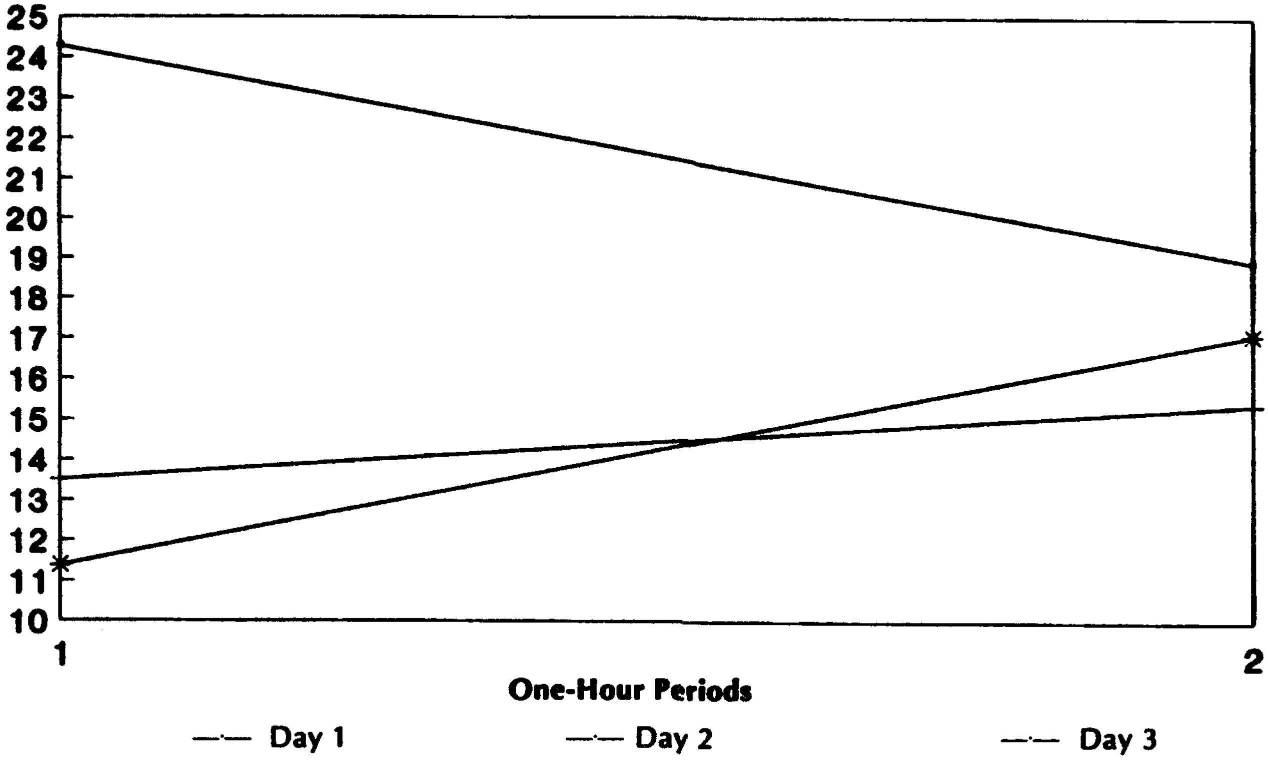
**Number of missed "inner" and "outer" VFR intruder aircraft**

| DAY | Inner | Outer |
|---|---|---|
| 1 | 4 | 8 |
| 2 | 0 | 7 |
| 3 | 1 | 7 |

**Mean reaction time for detection of VFR intruder aircraft**

| DAY | Inner | Outer |
|---|---|---|
| 1 | 18.96 | 24.60 |
| 2 | 8.76 | 20.94 |
| 3 | 8.58 | 29.16 |

**Figure 4.** *Average time (in seconds) required to detect the VFR intruders.*




number of missed "inner" VFR targets. The ANOVA for changes in frequency of the VFR events revealed a significant effect for days (F(2,38) = 5.05, p<.01) and location (F(1.19) = 10.34, p<.01).

The ANOVA for the reaction times associated with detection of the VFR targets revealed a significant effect for days (F(2,36) = 7.51, p<.01) and location (F(1.18) = 14.82, p<.01). Detection times were clearly shorter for the "inner" compared to the "outer" VFR targets. While there was little change in the average detection times for the "outer" targets (24.60 s on day 1 to 20.16 s on day 3), the difference between the average on the first day (18.96 s) for the "inner" targets was markedly above the averages for days 2 and 3 (8.76 s and 8.58 s, respectively). Thus, while subjects demonstrated an improved capability across days to detect the "inner" VFR targets, there was little evidence of any improvement in the speed with which subjects detected targets located near the edges of the display.

A final analysis concerning the VFR data involved the determination of possible time-on-task effects, using first and second hour comparisons of combined reaction times for the "inner" and "outer" targets. Results of those comparisons for each of the 3 days are presented in Figure 4. On day 1 performance improved, from an average of 24.3 during the first hour, to 18.9 on the second. Overall performance was better on days 2 and 3. For day 2, the mean reaction time increased from 13.5 to 15.4 across the 2 h, compared to an increase from 11.4 to 17.1 on day 3. Thus, while subjects demonstrated some improvement in performance across the 2 h on the initial day, on subsequent days there was some evidence of a decline in performance associated with time-on-task. Results of the ANOVA for these data revealed a significant effect for days (F(2,36) = 4.15, p<.05).

## TABLE 3

**Mean rating scale values for the two measurement periods across each day**

| VARIABLE | PERIOD | DAY 1 | DAY 2 | DAY 3 |
|---|---|---|---|---|
| Attentiveness | Pre | 6.89 | 5.95 | 5.16 |
| | Post | 4.47 | 5.26 | 5.05 |
| Tiredness | Pre | 5.11 | 5.89 | 5.79 |
| | Post | 6.68 | 6.53 | 5.79 |
| Strain | Pre | 4.63 | 3.74 | 3.84 |
| | Post | 4.47 | 4.74 | 4.47 |
| Boredom | Pre | 3.74 | 5.05 | 4.68 |
| | Post | 6.53 | 6.00 | 5.95 |
| Annoyance | Pre | 1.21 | 1.74 | 1.74 |
| | Post | 2.37 | 3.05 | 3.47 |
| Effort | Initial | 4.05 | 4.00 | 4.16 |
| | End | 6.42 | 6.00 | 5.32 |
| Difficulty | Initial | 3.11 | 2.32 | 2.16 |
| | End | 4.05 | 3.47 | 2.95 |

### Subjective

Outcomes from ratings provided by subjects appear in Table 3. Subjects generally rated their attentiveness as being fairly high (above the mid-point on the 10-point scale) at the start of each session; 6.89, 5.95 and 5.16 on days 1, 2, and 3 respectively. The ANOVA for attentiveness revealed a significant main effect for session ($F(1,18) = 7.51, p<.01$) and a significant session by day interaction ($F(2.36) = 8.70, p<.01$). While attentiveness at the completion of the session was considerably below the starting level on the first day (6.89 versus 4.47), on each of the next successive days the difference between the pre- and post- session levels decreased (5.95 to 5.26 on day 2 and 5.16 to 5.05 on the final day).

Subjects' ratings indicated they were more tired at the end of the session. Outcomes from the ANOVA revealed a significant main effect for sessions ($F(1,18) = 5.55, p<.05$) and a significant session by day interaction ($F(2,36) = 5.63, p<.01$). The average pre- to post-session difference in self-reported fatigue declined from 1.4 on the first day to .5 on the second day, and no difference in the 2 ratings on day 3.

While the ANOVA did not reveal any significant changes associated with the strain ratings, the main effect for session approached significance ($F(1,18) = 4.06, p<.06$). Self-reported strain was generally lower at the start, as compared to the close of each session, except on the first day.

Ratings of boredom evidenced an increase from start to completion of each day and from day 1 to days 2 and 3. Results of the ANOVA revealed a significant main effect for session ($F(1,18) = 22.86, p<.01$) and the interaction of day and session ($F(2,36) = 7.45, p<.01$).

The annoyance ratings remained relatively low across sessions and days (ranging from 1.21 to 3.47). Results of the ANOVA revealed significant effects for days ($F(2,36) = 4.54, p<.05$) and sessions ($F(1,18) = 18.76, p<.01$). Post-session annoyance ratings were significantly higher than pre- levels. Overall annoyance levels were slightly higher on days 2 and 3 than day 1.

Ratings for level of effort at the start of each session, which were gathered following completion of each session, remained relatively unchanged across days (4.05, 4.00, and 4.16). The ANOVA revealed a significant effect for sessions ($F(1,18) = 17.16, p<.01$). Level of effort required at the completion of each session was consistently above that of the pre level, yet declined from a high of 6.42 on the first day to 5.32 on day 3.

Ratings on the difficulty of the task, which were also gathered following completion of each session, revealed perceived differences in the task from start to end of each session. There was a significant main effect for session ($F(1,18) = 10.62, p<.01$) and days ($F(2,36) = 4.71, p<.05$) evident in the ratings of task difficulty. Ratings of perceived difficulty at the start of each session declined from an average of 3.11 on day 1 to 2.32 on day 2 and 2.16 on the final day. Self reported difficulty at the close of each session was above that of the initial level on each day.

## Discussion

Outcomes regarding subject performance on the primary and secondary tasks were generally consistent with those noted in a previous study by Thackray and Touchstone (1991). The detection of the altitude malfunction (XXX), with the exception of the general improvement in performance across the four 30 m time periods on the first day, remained relatively consistent across each of the 4 30-m time periods. The higher initial detection times during the first 2 30-m periods may have been associated with the reduced frequency of the altitude malfunction events in this study (1 versus the 3) compared to earlier studies (Thackray and Touchstone, 1985; Thackray and Touchstone, 1989; and Thackray and Touchstone, 1991). For this element of the overall task, there was no evidence of the typical performance decrement observed in most continuous monitoring tasks. There was also little evidence of a consistent improvement in performance across days. It may be that for situations where the detectability of the target is high, where information processing requirements are low, and the task is easily subsumed under normal scanning behavior, the task remains relatively immune to the effects of fatigue. Those same characteristics also appear to leave the task relatively unchanged with respect to learning or the development of strategies to improve overall performance across repeated sessions.

Parasuraman (1986), in his review of vigilance, monitoring and search, indicated that while initial research with complex monitoring tasks suggested that the typical vigilance decrement associated with time-on-task did not occur, more recent studies involving a variety of tasks were supportive of the vigilance decrement in both simple and complex monitoring tasks. Research with the complex monitoring task used in this study has consistently demonstrated an increase in detection times for the identification of conflicting aircraft associated with time-on-task (Thackray and Touchstone, 1989 and Thackray and Touchstone, 1991). While Thackray and Touchstone (1989) suggest that this decrement is associated with aspects of attention and are not attributable to any "gross" changes in scanning activity, per se, there is evidence from analyses of changes in the gaze measures that were collected as part of this study (Stern, et al., in press) that selected aspects of those measures do evidence changes associated with time-on-task. There are subtle changes in several gaze measures observed during monitoring (e.g., the nature and frequency of eye blinks, as well as modifications in saccades) that may be linked with declining performance. Nonetheless, it is on the more difficult components of the monitoring task that the performance changes are reflected, rather than on the more readily observed and more easily processed aspects of the monitoring task (altitude malfunctions).

An important consideration in the increased detection times associated with this task is that the increase is not uniformly observed across each of the stimulus events. As Thackray, et al. (1979) indicated in their study where detection times were not limited, the distribution of reaction times across the 30-m periods becomes skewed. The number of occasions where delayed detection times were observed increased significantly, while for most events detection times remained unchanged. Those findings are consistent with results from this study where the mean detection times for the conflict events that were detected prior to the 26-s warning did not evidence any change across the 2-h session. As Parasuraman (1986) indicates, these differences suggest that the increase in detection times is due not to any modification in scanning, per se, but to temporary lapses in attention or "blocks." The changes in gaze parameters that occurred as a result of time-on-task (Stern, et al., in press) may be an indication of temporary time periods where subjects are more susceptible to "blocks" or attention lapses while monitoring a complex display. Additional investigation is needed to determine the precise relationship between the eye movement parameters and "blocks" or attentional lapses.

In support of previous research with generally less complex monitoring tasks (Lindsley, et al., 1944; Webb and Wherry, 1960; Adams, Humes, and Stenson, 1962; and Adams, Humes, and Sieveking, 1963) subjects in this study demonstrated improved performance across days in their ability to detect altitude conflicts between aircraft. Despite their overall improved performance across the 3 days, the nature of the performance decrement associated with time-on-task remained relatively consistent for each day. That finding was also consistent with outcomes reported by Adams, et al. (1962, 1963). It should also be noted that improvment was not uniformly evident in all aspects of performance on the primary and secondary tasks. No change was evident in the detection times associated with the altitude malfunctions, and there was little evidence of improvement in the detection of the "outer" VFR intruder events. Thus, performance improvement in monitoring is, to some extent, dependent on the nature of the task and perhaps, to some degree, in the extent to which the task has an associated decision-making component.

Detection of the VFR intruders was most influenced by the location of the targets. The "edge effect" (Baker, 1958; Baker, Morris, and Steedman, 1960; Enoch, 1959), referring to the tendency of observers to fixate more attention on the center, as opposed to the edge of the display, was evident in subjects' performance in detecting the VFR intruders. The "edge effect" has also been found in an operational task involving the detection of flaws by experienced industrial inspectors (Schoonard, Gould, and Miller, 1973). Both the frequency of missed targets and the average time required to detect targets was higher for those located closer to the edges of the display than those more centrally located. Evidence for a time-on-task effect relative to the detection of the VFR intruders was inconsistent. During the initial day, an improvement in performance (combined averages for the "inner" and "outer" events) was noted from the first to the second hour. However, on days 2 and 3 there was some evidence in support of an increase in the average detection time from the first to second hour. This contrasts somewhat with the findings of Thackray and Touchstone (1991), who observed a significant increase in reaction times between the first and second hour. The effects of repeated sessions were dissimilar for the "inner" versus "outer" targets. While subjects demonstrated a significant improvement from the first to second day in detection of the "inner" targets, detection of the "outer" targets remained relatively unchanged. Whether this difference is due to an improved strategy for identifying targets more centrally located, to a change in scanning techniques or decrease in task dificulty as a function of learning, is unknown. It does suggest that performance improvement in these tasks across sessions is not uniform for all elements of task performance. It is not known whether specific strategies could be developed and taught that would improve detection of both "inner" and "outer" targets. On the other hand, on the basis of a study by Thackray and Touchstone (1991), the "edge" effect can be eliminated in monitoring performance by increasing the conspicuity value of the VFR intruders by using a flashing target.

Subjective ratings of the task were consistent with previous findings (Thackray and Touchstone, 1991). Subjects generally perceived the task as being relatively demanding with respect to maintaining attentiveness. With the exception of the "strain" measure, subjects reported an increase in tiredness, boredom, and irritation from start to close of each session. The improvement in performance (learning effect) observed in the detection of the altitude conflict events was apparently a factor in subjects perceiving the task as requiring less attentiveness and effort and involving less difficulty during the course of the sessions on days 2 and 3.

## References


Adams, J. A., Humes, J. M., and Sieveking, N. A. 1963. Monitoring of complex visual display: V. Effects of repeated sessions and heavy visual load on human vigilance. *Human Factors.* 5, 385-389.

Adams, J. A., Humes, J. M., and Stenson, H. H. 1962. Monitoring of complex visual displays: III. Effects of repeated sessions on human vigilance. *Human Factors.* 4, 149-158.

Baker, C. A., Morris, D. F., and Steedman, W. C. 1960. Target recognition on complex displays. *Human Factors.* 2, 51-61.

Baker, C. H. 1958. Attention to visual displays during a vigilance task. *British Journal of Psychology. 49, 279-288.*

Enoch, J. M. 1959. Effect of the size of a complex display upon visual search. *Journal of the Optical Society of America.* 49, 280-286.

Federal Aviation Administration. 1989. Report of the interagency near midair collision working group. Washington, DC. FAA Office of Aviation Safety.

Lindsley, D. B., et al. 1944. Radar operator "fatigue." The effects of length and repetition of operating periods on efficiency of performance. OSRD Report No. 3334.

Mertens, H. W., Thackray, R. I., and Touchstone, R. M. 1992. *Effects of color vision deficiency on detection of color-highlighted targets in a simulated air traffic control display.* Report No. DOT/FAA/AM-92-6. Washington, DC: Federal Aviation Administration, Office of Aviation Medicine.

Parasuraman, R. 1986. Vigilance, monitoring, and search. In *Handbook of Perception and Human Performance* (Edited by K. R. Boff, L. Kaufman, and J. P. Thomas) New York: Wiley.

Schoonard, J. W., Gould, J. D., and Miller, L. A. 1973. Studies of visual inspection. *Ergonomics,* 16, 365-379.

Stern, J. A., Boyer, D., Schroeder, D. J., Touchstone, R. M., and Stoliarov, N. Blinks, saccades and fixation pauses during vigilance task performance. (In press) Washington D.C.: Federal Aviation Administration, Office of Aviation Medicine Report.

Thackray, R. I., Bailey, J. P., and Touchstone, R. M. 1979. The effect of increased monitoring load on vigilance performance using a simulated radar display. *Ergonomics,* 22, 529-539.

Thackray, R. I. and Touchstone, R. M. 1981. *Age-related differences in complex monitoring performance.* Report No. DOT/FAA/AM-81-12. Washington, DC: Federal Aviation Administration, Office of Aviation Medicine.

Thackray, R. I. and Touchstone, R. M. 1983. *Rate of initial recovery and subsequent radar monitoring performance following a simulated emergency involving startle.* Report No. DOT/FAA/AM-83-13. Federal Aviation Administration, Office of Aviation Medicine.

Thackray, R. I. and Touchstone, R. M. 1985. *The effect of visual task load on critical flicker frequency (CFF) change during performance of a complex monitoring task.* Report No. DOT/FAA/AM-85-13. Federal Aviation Administration, Office of Aviation Medicine.

Thackray, R. I. and Touchstone, R. M. 1989. Effects of high visual task load on the behaviors involved in complex monitoring. *Ergonomics.* 32, 27-38.

Thackray, R. I. and Touchstone, R. M. 1991. Effects of monitoring under high and low task load on detection of flashing and colored radar targets. *Ergonomics,* 34, 1065-1081.

Webb, W. B. and Wherry, R. J., Jr. 1960. Vigilance in prolonged and repeated sessions. *Perceptual Motor Skills.* 10, 111-114.




★U.S. GOVERNMENT PRINTING OFFICE: 1994 - 560-063/80122