Image Cover Sheet

CLASSIFICATION	SYSTEM NUMBER
UNCLASSIFIED	508776

TITLE

HUMAN FACTORS OF CC-130 OPERATIONS. VOLUME 6: FATIGUE IN LONG-HAUL RE-SUPPLY MISSIONS

System Number:

Patron Number:

Requester:

Notes:

DSIS Use only:

Deliver to:

National Défense
Defence nationale

DCIEM Report No. 98-R-19

Human Factors of CC-130 Operations

Volume 6: Fatigue in long-haul re-supply missions

M.A. Paul
R.A. Pigeau
H. Weinberg

Defence and Civil
INSTITUTE OF ENVIRONMENTAL MEDICINE

Canada

February 1998 DCIEM No. 98-R-19

HUMAN FACTORS OF CC-130 OPERATIONS

VOLUME 6: Fatigue in long-haul re-supply missions

M. A. Paul
R. A. Pigeau
H. Weinberg

Defence and Civil Institute of Environmental Medicine
1133 Sheppard Avenue West, P.O. Box 2000
North York, Ontario
Canada M3M 3B9

© HER MAJESTY THE QUEEN IN RIGHT OF CANADA (1998)
as represented by the Minister of National Defence

© SA MAJESTE LA REINE EN DROIT DU CANADA (1998)
Défense Nationale Canada

DEPARTMENT OF NATIONAL DEFENCE – CANADA

CONTENTS

EXECUTIVE SUMMARY 3

INTRODUCTION 4

METHODS 6

RESULTS & DISCUSSION 12

CONCLUSIONS. 40

RECOMMENDATIONS 41

ACKNOWLEDGEMENTS 41

REFERENCES 42

EXECUTIVE SUMMARY

Introduction. Deployment of troops in foreign theatres requires a massive airlift capability. Transport squadrons are called upon to deliver personnel and materiel, day and night, around the clock, usually during long transmeridian flights. The relentless fatigue encountered in such operations can be severe enough to pose a flight safety hazard. During Operation Alliance, a month long airlift (January 1996) in support of Canadian troops in Bosnia, 18 Air Transport Group CC-130 Hercules carried out 86 missions from Trenton to Split with aircraft landing in theatre every four hours. Most crews attained the 120 hour maximum allowable flying time per 30 day period, in as little as two weeks. There were several reports of aircrew falling asleep at the controls. After the main airlift was completed, only three of these missions were flown each week. This frequency of re-supply missions was adequate to sustain Canadian troops in former Yugoslavia. Because these sustainment flights were flown in a more relaxed manner, aircrew were given 32 hours on the ground in the United Kingdom on airrival from Trenton, Canada (five time zones), before proceeding to Zagreb, Croatia (one time zone), instead of the 14 hours they were given at this stage of the mission, during the original airlift. The current study is an attempt to document to what extent fatigue (and time zone changes of five and six hours) can impact aircrew performance during routine re-supply missions. Any fatigue-related lapses in performance in this study, will only worsen when crews are asked to fly repeated missions back-to-back with minimum crew rest between mission legs as was done during Operation Alliance.

Methods. Ten routine re-supply missions from Trenton to Zagreb were studied and involved 53 aircrew subjects. In order to document their sleep hygiene, the aircrew were asked to wear wrist actigraphs from approximately five days prior to the mission, until the mission was completed. Aircrew psychomotor performance was tested during the actual flights on two different computer-based psychomotor batteries. One test battery consisted of a questionnaire designed to elicit subjective ratings of alertness, and fatigue, as well as three separate serial tests from the DCIEM SUSOPS battery (serial reaction time, logical reasoning, and serial subtraction). The other battery was a recently developed multi-task (analogous to a flying task) and incorporated four sub-tasks which had to be undertaken simultaneously. There were three psychomotor test sessions during the outbound transatlantic leg (Trenton to Lyneham, UK), one test session on the Lyneham-Zagreb-Lyneham leg, and three test sessions on the return transatlantic leg from Lyneham to Trenton.

Results. The amount of sleep the aircrews experienced during the days leading up to a mission steadily decreased from an average of 475 minutes per day to an average of 380 minutes per day, with 25% of the subjects getting less than six hours and some subjects obtaining as little as 250 minutes per day on the last pre-mission night at home. During the missions, the worst night of sleep occurred during the second night in the UK, no doubt because of having to sleep at inappropriate circadian times. Subjective ratings of alertness and fatigue became progressively worse during both transatlantic legs. Of the three SUSOPS psychomotor tests, only the logical reasoning task indicated any fatigue related lapses in performance during these missions, and only on the outbound leg. The multi-task data illustrates unequivocal fatigued-related lapses in performance on both transatlantic legs, with 13 of the 18 pilots tested, showing performance ranging from minimum to marked impairment.

Recommendations. 1) Repeat this study during a major airlift like Operation Alliance. 2) Given that the worst night for sleep occurred at home, the last night before the mission, delay departure by one to two hours if at all possible. 3) Determine whether or not melatonin can safely facilitate sleep at an inappropriate circadian time.

INTRODUCTION

While pilots must undertake their flying tasks with a high degree of psychomotor performance, stressors such as noise and vibration, long crew days, irregular work schedules, circadian disruptions, and inadequate sleep can interact to produce dangerous levels of fatigue (7). The correlation between increased fatigue and decreased performance has been reported by Barth et al (3), French et al (11,12), Mortimer (20), Perelli (21), and Shingledecker and Holding (25). Although the initial effects of fatigue can go unnoticed, eventually vigilance, judgement, situational awareness, and crew co-ordination may all be compromised (7). Belenky et al (4) found that for an artillery firing task, during continuous combat operations, when sleep is limited to less than six hours per day, performance deteriorates. Ritter (23) reviewed incident reports from the Aviation Safety Reporting System and concluded that sleep loss, circadian disruption and improper nutrition contributed to fatigue-related cognitive errors made by aircrews. In fact some of these incident reports demonstrated that aircrews routinely obtained less than six hours of sleep per night. McCauley (18) found that several aircraft mishaps have been attributed to pilot fatigue because of disrupted and inadequate sleep. Billings (5) states that more than half of all aviation accidents are probably the direct result of fatigue-related pilot inattentiveness.

Deployment of troops in foreign theatres requires a massive airlift capability. Transport squadrons are called upon to deliver personnel and materiel, day and/or night, around the clock, usually during long transmeridian flight. The relentless fatigue encountered in such operations can be severe enough to pose a flight safety hazard (26). Crew days are often pushing the upper limits of the maximum allowable 16 hour crew day. At the end of that crew day, crews get a minimum of about 14 hours rest (during which they are often unable to sleep soundly because of circadian rhythm disruption) and then are required to start all over again.

The key to minimizing the impact of such demanding missions is appropriate scheduling concessions to afford aircrews the opportunity to obtain adequate rest, both during and between missions. During the early to mid 1970s, several publications provided scheduling guidelines to limit the impact of long transmeridian flights on air crews in an effort to avoid compromising flight safety (6,15,19). One of these scheduling systems (6) was under consideration by what was then the Air Transport Command of the Canadian Forces. While routine transport missions can normally be scheduled with a view to providing ample opportunity for crew rest, the same cannot be said for contingency operations. The imperatives of contingency operations do not easily lend themselves to the notion of scheduling concessions, especially when crews are required to fly tactical transports designed for three to four hour long missions in a strategic role with crew days approaching the maximum allowable crew day, and sometimes, of necessity, exceeding it.

One of the most compelling series of statistics in the fatigue literature is provided by Coren (10) who examined Canadian Ministry of Transportation data on motor vehicle accidents to compare the number of accidents immediately before and after the shifts to and from daylight savings time for 1991 and 1992. He found that for every province except Saskatchewan (which does not shift to and from daylight savings time), on the Monday after the spring time change, traffic accidents increased by seven percent, and this effect disappeared within a week. Conversely, in the fall, when we gain

an extra hour of sleep, Coren found that this pattern is reversed such that there is a decrease of seven percent in the number of reported accidents but a week later the accident rate increases again, back to "normal". This data is even more compelling because it represents over 1.5 million accidents from all over Canada. If a one hour change in sleep time changes the motor vehicle accident rate by seven percent in a two-dimensional task like driving, then there is every reason to expect that when crews take off with a more significant sleep deficit than an hour, then fly (a more demanding 3-dimensional task) long transatlantic missions in slow aircraft and suffer from jet lag for a few days in Europe, before returning home via another long transatlantic flight, we should expect deleterious effects of fatigue on crew performance with the attendant potential to compromise flight safety.

In Operation Alliance, a month long airlift that took place during January 1996 in support of Canadian troops in Bosnia, 18 Air Transport Group (ATG) CC-130 Hercules carried out 86 missions from Trenton to Split with aircraft landing in theatre every four hours. During operation Alliance, most crews attained the 120 hour maximum allowable flying time per 30 day period, in as little as two weeks. There were several reports of aircrew falling asleep at the controls during these types of missions. Despite this anecdotal information, however, to date, little research has provided a detailed scientific documentation of the impact of fatigue upon aircrew performance.

After the main airlift was completed, only three of these missions were flown each week. This frequency of re-supply flights was adequate to sustain Canadian troops in former Yugoslavia. Because these sustainment flights were flown in a more relaxed manner, aircrew were given 32 hours on the ground in the United Kingdom on arrival from Trenton, Canada (five time zones), before proceeding to Zagreb, Croatia (one time zone). This is contrasted with the 14 hours they were given at this stage of the mission, during the original airlift. Our initial efforts were therefore to monitor the performance of the aircrew flying these sustainment missions. Later, we hope to be in a position to monitor aircrew performance during a major airlift.

The current study is an attempt to document to what extent fatigue (and time zone changes of five and six hours) can impact on aircrew performance during routine re-supply missions to former Yugoslavia. If we can document significant fatigue-related lapses in performance in this study, we can assume that the fatigue problem will only worsen when crews are asked to fly repeated missions back-to-back with minimum crew rest as was done during Operation Alliance.

While the original plan for this study called for the monitoring of psychomotor performance and sleep hygiene (via wrist actigraphs and activity logs) for 15 re-supply missions, only 10 missions were studied. This reduction was due to the fact that as the A-310 (CC-150) aircraft were returning to duty after being re-configured to carry cargo, they took on progressively more and more cargo missions to Yugoslavia, relieving the number of those missions flown by the CC-130 squadrons. When the data collection for this study began in late January 1997, there were three re-supply missions per week undertaken by CC-130 aircrew. By mid data collection (March 1997) there was on average only one mission per week flown by CC-130 aircrew. By the end of the data collection period (June 1997) it was a rare event to have a CC-130 fly to Croatia, and then probably only because the payload involved vehicles which the A-310s can not carry.

METHODS

Subjects

The MOC (military occupation code) demographics for the 53 subjects (10 crews) who participated in the study are illustrated in Table 1.

TABLE 1. Demographics of participating subjects

	Aircraft Commanders	Co-Pilots	Navigators	Flight Engineers	Loadmasters
#/gender	9 males	9 males/1female	8 males/2 females	10 males	15 males
mean age (yr)	35.5	29.6	32.5	40.4	34.5
age range	29 to 42	24 to 32	23 to 49	36 to 49	31 to 44
rank range	Capt to LCol	Lt to Capt	Lt to Maj	Sgt to MWO	M/Cpl to MWO

One limitation of this study was the difficulty of gaining access to the CC-130 aircrews for several days prior to the missions. Such access would have allowed for time to train them to a level of stable performance on the psychomotor tasks. Therefore the results will have to be interpreted carefully to factor out learning effects. In the future we hope to further address this limitation by using the non-fatigued aircrew subjects in a fatigue countermeasure study as controls. The psychomotor performance learning curves of the non-fatigued subjects will be compared to the learning curves of the aircrew on the missions to former Yugoslavia in order to separate out the effects of fatigue on the latter aircrew. We realize that this strategy is not optimal for controlling learning effects in our present study, but the exigencies of field studies often require such compromises.

Equipment and Data Collection Strategy

In order to establish the number of hours of sleep each aircrew member obtained during these missions, as well as in the period immediately prior to the commencement of the missions, all aircrew were asked to wear wrist actigraphs from approximately 5 days prior to the mission until the mission was completed. Further, all aircrew were asked to fill out sleep/activity/mood questionnaires in a paper log prior to the missions (on a pre-mission questionnaire), during the missions (on a mission questionnaire), and on a laptop computer, also during the mission. The questions relating to the amount and quality of sleep obtained each night were to be answered each morning, for five days prior to the mission, and just prior to the commencement of daily flying operations during a mission. The computer-based questions relating to their fatigue states were answered each time the aircrew performed a psychomotor test battery during the mission.

While it was anticipated that the crews on these missions would experience fatigue due to the length of the mission, they also likely experienced some fatigue due to the circadian stresses inherent in such long transmeridian flights. Thankfully, all of these missions departed from Trenton at approximately 0800 hrs local time. Therefore, we should not have any confounding of results due to different circadian rhythm stresses across crews.

Existing psychomotor test batteries, such as the DCIEM sustained operations

(SUSOPS) test battery, have a long history of well documented laboratory-based findings concerning the effects of fatigue on performance (1, 22). Thus, it was essential to include these tasks to tie any obtained findings to the existing literature. As valuable as these cognitive tasks are, one possible concern is the degree of applicability that lower level cognitive tasks have to a flying task. Therefore we also included a multi-tasking test (16, 27, 28) thought to possess greater context validity to flying, as well as requiring higher levels of cognitive activity to complete. Beyond providing a link to the existing fatigue literature, the present study represents one of the first attempts to fully validate this multitasking task. As a result, two different psychomotor performance batteries were run from a single lap-top computer;

- a subset of 3 tests (SRT (serial choice reaction time), LRT (logical reasoning task), and SST (serial subtraction task)) from the DCIEM SUSOPS battery which were performed, one at a time, over 10 minutes and

- a multi-tasking battery developed by Dr Hal Weinberg which took 15 minutes per data collection iteration (16, 27, 28).

The crews were only tested on the psychomotor performance batteries during actual flight in the following manner;

- three times on the outbound leg from Trenton (Canada) to Lyneham (U.K.)
- once during the leg from Lyneham to Zagreb (Croatia) and return, and
- three times on the inbound leg from Lyneham to Trenton.

All crew positions (pilots, navigator, flight engineer, and loadmasters) performed the serial iterations of the SUSOPS test battery, while only the pilots performed the multi-tasking battery (because it is essentially a flying task and the other crew positions would not have the pre-requisite training to give us meaningful data on this task). However, all crew positions were asked to take part in the wrist actigraphy/sleep logs and to complete the questionnaires relating to sleep quality/mood/alertness/fatigue ratings.

Task Descriptions

Susops tasks.

All of the SUSOPS tasks were performed on lap-top computers (Pentium 133 Mhz chips with 16 Mb RAM and 1 Mb of video memory) using an external 'bus' mouse as the input device. All subjects performed the tasks in the rear of the fuselage, in order to avoid distractions by cockpit instrumentation or interfering with the aircraft controls, and to perform the task in relative privacy. The subjects were seated in the standard collapsible web-seats used for troops and had a custom-contoured board (with an anti-skid surface) on their laps. The computers were placed on this 'lapboard' which was large enough to accommodate the computer and a mouse pad (approximately 59 cm x 42 cm). Subjects responded to the computer questionnaire and all three SUSOPS tasks by moving the mouse until the cursor was super-imposed over their chosen response, and then by clicking the mouse. Subjects were instructed to work as quickly as possible without sacrificing accuracy. The subjects wore headsets to defend against the high ambient noise but in order to avoid distractions by routine communications between crew members, the headsets were not plugged into the aircraft intercom while they were performing the psychomotor tests.

Computer Questionnaire.

The first item of any SUSOPS test battery iteration was the computer questionnaire which involved a brief response to the following four questions;

- the subject's assessment of his/her own state of alertness by selecting the most appropriate response from the 7 point Stanford Sleepiness Scale (17),
- the subject's assessment of his/her own mental fatigue state by selecting any number on a continuous scale from 1 (very mentally fresh) to 7 (very mentally fatigued),
- the subject's assessment of his/her own physical fatigue state on a continuous scale from 1 (very physically fresh) to 7 (very physically fatigued), and
- current aircraft position (latitude and longitude) in order to track where each testing session took place.

The aircraft position was obtained from the cockpit crew by the experimenter, via the intercom system.

Serial Reaction Time.

This task required the subjects to select which of four letters (A, B, C, or D, which were presented on the computer screen in a rectangular response grid) corresponded to the single stimulus letter (again A, B, C, or D) which was briefly presented on the computer screen (immediately above the response letter grid) and was subsequently replaced by single serial random presentations of any of the four stimulus letters immediately after each response.

Logical Reasoning.

This task was developed by Baddeley (2) and described by him as involving higher mental processes, based on grammatical transformation: the task involves understanding of sentences of varying syntactic complexity. It consists of presentations (on the computer screen) of one of 16 sentences (such as 'A is not preceded by B') followed by a pair of letters (either 'AB' or 'BA'). The subjects were required to indicate whether the sentence was a true or false description of the associated letter pair by selecting either the 'True' or 'False' response box on the computer screen. There were 32 possible combinations of sentence and letter pairs.

Serial Subtraction

This task is similar to a task described by Cook, Cohen, and Orne (9). Here the subjects were presented with a randomly chosen three-digit number between 500 and 999 (on the computer screen) (e.g., 763) and were also presented with a randomly chosen subtrahend between 5 and 9 (e.g., 9). The subjects were required to perform serial subtractions (e.g., 763-9=754, 754-9=745, 745-9=736, etc). The task also involved a short term memory component in that after the first subtraction was performed, all numbers disappeared from the computer screen, forcing the subjects to remember the last solution as well as the subtrahend.

Multitask.

General Description.

This programme was run on the same computer and 'lapboard' as the SUSOPS test battery. When the programme was running, the screen showed four separate displays (figure 1) representing four sub-tasks which must be performed simultaneously. Three of these four tasks interacted. These tasks simulated flying an aircraft to specific targets or "waypoints" as described below. The raw output data file was merged with a computer reduction algorithm to yield a single final weighted composite score.

Tracking Task.

The upper left display (tracking task) represented an aircraft control panel in which a small white box showed the action of the aircraft control column which was manipulated by the mouse. Left or right motions banked the aircraft to the left or right increasing or decreasing the rate of turn in degrees per second. Up and down mouse inputs pitched the aircraft up or down, increasing or decreasing the rate of change of altitude in feet per minute. These values were displayed on the panel along with the current aircraft heading and altitude which changed in response to inputs. In addition, there were flight director 'bars' (normally shown in orange) which helped guide the change of aircraft climb/descend rate and heading based on what has been entered into the 'flight computer'. Figure 1 shows the flight director bars (localizer and glidepath) intersecting to the right of, and just below, centre, which indicates a command to move the cursor (currently at dead centre which indicates straight and level flight) to the intersection of the localizer and glidepath bars, in order to 'capture' the current target heading and altitude. Thus these bars moved accordingly when the target values were changed by using the F1 (heading) or F2 (altitude) keys and typing in new values and indicated the flying pattern (inputs) needed to achieve the new heading and/or altitude. When the programme started, the aircraft was already on a heading of 0 degrees and at an altitude of 5,000 feet. The object of this task was to maintain the aircraft on the current heading and altitude (i.e., to try to bring the orange bars into the centre of the screen and maintain them in this position. However, the values that should be entered into the computer at any time ("target heading" and "target altitude") were dictated by the Waypoint task.

Waypoint task.

The upper right display (waypoint task) showed a map of the ground viewed from the aircraft. Straight up on the display represented the "front" of the aircraft (i.e., the current direction of the aircraft). The display showed a number of small triangles which represented waypoints, one of which was solid green in colour and represented the target waypoint. Of the 5 waypoints in shown in figure 1, the target waypoint is the one closest to the centre of the waypoint display (not shown as green in this figure because of printer limitations). As soon as the correct heading of the target, (relative to the current position of the aircraft), was determined by the subject, this value was entered into the flight computer as the "target heading". If the target waypoint was not visible the range of view of the "map" could be incrementally increased or decreased by a factor of 2 using the zoom-in and zoom-out keys (F8 and F9, respectively) until it was found. Similarly, the map could be zoomed in for a more detailed view of the target waypoint once it was in close range of the aircraft. The object of this task was to fly towards the target waypoint as efficiently as possible until it was intersected by the aircraft. Successful 'capture' of the waypoint was achieved if the centre of the triangle passed through the small circle in the centre of the screen. The subjects were told that is was important to minimise the amount of time that the target heading was different from that needed to intercept the waypoint. The subjects were also told that the score is reduced for the amount of time spent flying towards the target without changing the target heading to be as close to the necessary heading as possible.

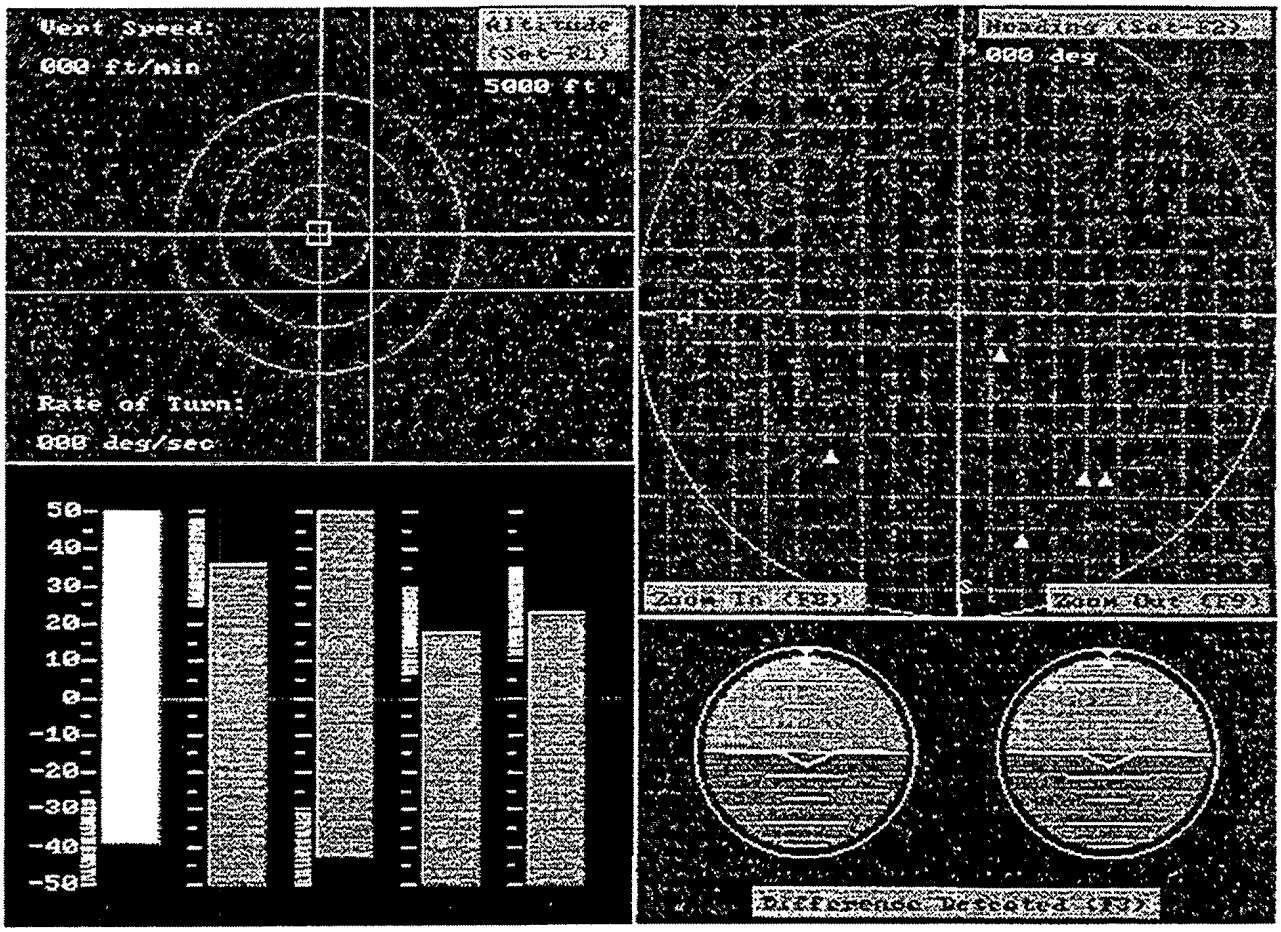

Figure 1. Computer screen 'capture' of multi-task display in grey scale instead of the colours which are present on the computer.

Once a waypoint was intercepted, it flashed briefly and a new triangle became the target. The subject attempted to capture as many waypoints as possible during the entire task.

A second task in this display was the instruction to change the target altitude to a new altitude (using the F1 button as described in the tracking task). This occurred periodically and consisted of the new target altitude being presented for 5 seconds at the upper left corner of the 'map' accompanied by a short beeping sound, however because of the high ambient noise in the CC130, this beep could not be heard. This auditory limitation caused this particular sub-task to become a visual vigilance task rather than the designed combination of auditory and/or visual vigilance which would have been possible in a more quiet aircraft. It was important to note and respond to these requested altitude changes immediately - once the display disappeared from the screen it could not be retrieved. The overall score was also influenced by the time the target heading was set to the correct value.

Instrument task.

The lower right display (instrument task) showed two attitude indicators (artificial horizon) which reflected the current attitude of the aircraft as determined by the rate of turn and climb/descent shown in the tracking task display. In addition, the deviation of

the current aircraft heading from the target heading was shown by a small arrow around the outside of the attitude indicator. This arrow came into line with the top of the display as the target heading was reached.

Periodically, there was a discrepancy in the information (e.g., angle of bank or pitch) being displayed by these two instruments where only one instrument showed the correct information. The object of this visual vigilance sub-task was to depress the "difference detected" button (F3) as soon as such a difference is perceived. Subsequently, 2 buttons appeared below each instrument indicating which key (F4 or F5) to press to indicate the instrument which was showing the correct information, after which the two instruments again displayed the same information. Points were deducted for failure to notice these differences in sufficient time as well as 'false alarms' and incorrect choice of the instrument showing the 'correct' or 'true' information.

Bar Task.

The lower left display (Bar Task) showed five vertical bars which changed in their length. Beside each bar there was a red line indicating an acceptable 'target zone' of the bar. The object of this task was to keep the end of the bars within this range at all times. The bars changed their length periodically and target zones could also change. The length of the bars could be changed in the following way. The left and right arrow keys were used to select a bar. The active bar was indicated by a colour change from green to yellow. The active bar in figure 1 is the left most bar which is presented in white. The up and down arrow keys were used to select up or down movement of the bar with subsequent presses of the arrow keys changing the rate of movement of the bar in single steps up to a maximum rate. When the target zone was reached the arrow keys could be used to slow down and/or reverse the change in length of the bar to keep it within the target zone. This task did not interact with any of the other tasks.

Experimental Design Considerations

Because of the requirement to collect aircrew performance data without interfering with their flying duties, we did not use rigorously defined time intervals over which we collected data. However, pilots and navigators from ATG HQ identified the following possible data collection opportunities.

1) Trenton (Canada) to Lyneham, (United Kingdom) leg

Upon reaching top of climb (TOC) after departure from Trenton, the first testing cycle was commenced with the concurrence of the aircraft commander (AC) and, one at a time, each crew member rotated through the test site, located in the after end of the fuselage. For all such testing, the aircrew wore their head-sets to defend against the high ambient noise level but they were not plugged into the intercom system, so as to avoid the possibility of distractions from normal crew communications. Approximately half way across the Atlantic (about 6 hours elapsed time after take-off), all crew positions were again asked to perform the second iteration of thetest batteries. Approximately 2.5 hours out from Lyneham, all crew members were asked to perform a third iteration, such that all data collection for this leg was completed by the time the aircraft was about one hour out from Lyneham, in order not to interfere with the relatively busy approach to landing in Lyneham through congested European airspace.

2) Lyneham to Zagreb (Croatia) & return

Given that this entire leg was flown in very congested airspace, each crew member was tested only once on this part of the mission, either in-bound to Zagreb, or during the return leg to Lyneham. Often, it was most convenient to test one or both of the pilots on the ground in Zagreb.

3) Lyneham to Trenton leg

The data collection protocol on this final leg was similar to the first leg of the mission (i.e. from Trenton to Lyneham) except that there were fewer constraints toward the end of this leg, given that Canadian airspace was not as congested as the airspace over Europe. Further, the average flying time on this return leg was two hours longer than for the outbound leg, because of the normal westerly headwinds. The approximate psychomotor testing locations for all mission legs are illustrated in figure 2.

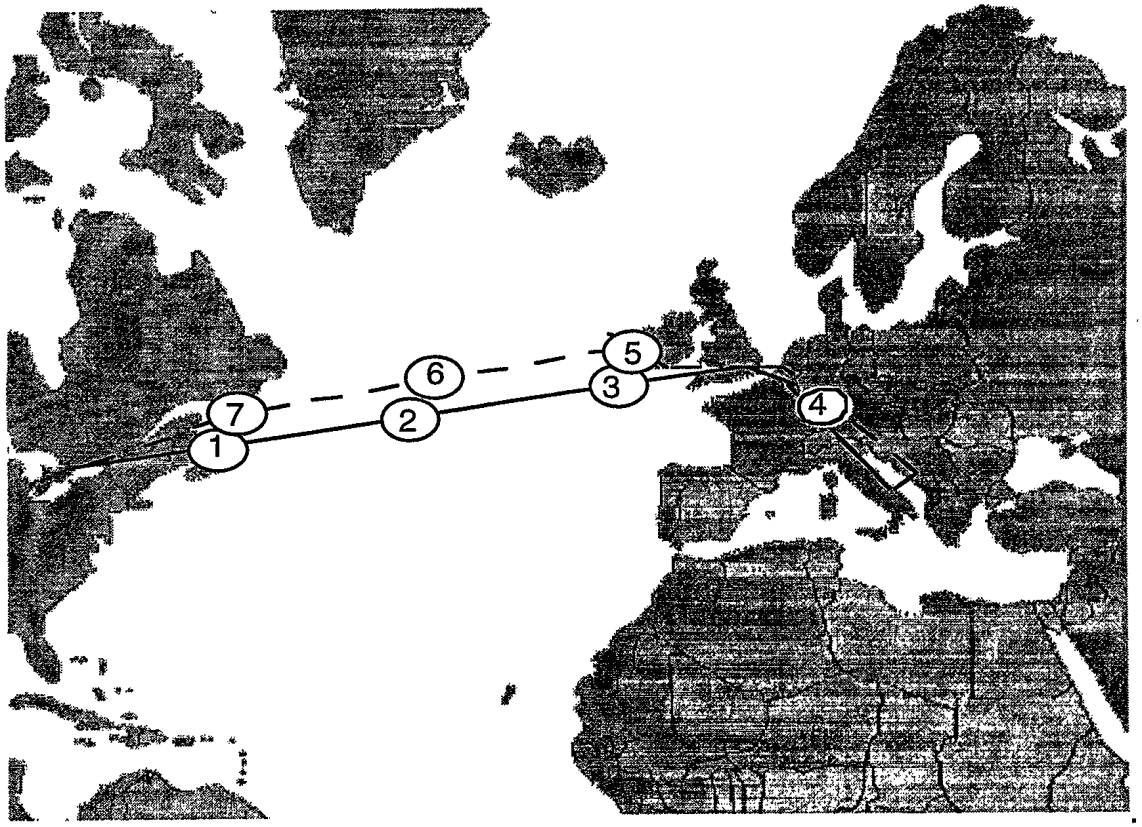

Figure 2. Map illustrating approximate outbound (solid lines) and return (dashed lines) tracks including approximate psychomotor performance testing locations. Locations 1, 2, and 3 correspond to testing locations for the outbound transatlantic leg, while location 4 averaged about 3 degrees East Longitude on either the outbound leg from Lyneham to Zagreb or several hours later on the return leg from Zagreb to Lyneham. Locations 5, 6, and 7 represent the test locations during the return transatlantic leg from Lyneham to Trenton.

Statistical Analysis

The tests results from the SUSOPS battery, the multi-tasking battery, the wrist actigraphs were submitted to one way analysis of variance with "MOC" (military occupation code) as the between factor and days or trials as the within factor. We used the Duncan Multiple Range Test to analyse any simple main effects or interactions. Because of low compliance with the request to maintain the paper logs, the pre-mission and mission sleep/activity/mood questionnaire data was not analysed.

RESULTS & DISCUSSION

Wrist Actigraphy

Wrist actigraphs measure every tenth of a second, whether a movement was detected from the wearer. Software (a programme called Win ACT, version 1.2, developed by Tim Elsmore from Activity Research Services of San Diego, California) looks at the frequency of such movements over time in order to quantify the number and duration of sleep episodes. The number of minutes spent asleep in a 24 hour period demonstrated a highly significant main effect of days ($p<.0001$) as illustrated in figure 3 where "daily sleep in minutes" are plotted over pre-mission and mission 'days'. The results of post hoc analyses are shown as 'p' values between adjacent cells in figures 3 and 5. The amount of sleep the subjects experienced during the days leading up to a mission steadily decreased from 475 minutes per day to 380 per day (figure 3). Given that the planned take-off time was 0800 hrs local time and that the crews reported to operations two hours before take-off, having arisen at their respective homes anywhere from one to two hours prior to reporting at operations, this result is not surprising. This is only 20 minutes more sleep than Belenky (4) found as the threshold minimal sleep which produced performance impairment in continuous combat operations. The 380 minutes (6 hrs and 20 minutes) of sleep during the last night in Trenton before the start of the mission is an average value across all aircrew. In fact, 23% of the aircrew received less than 6 hours of sleep with one individual receiving only 248 minutes (4 hours and 8 minutes) for that last pre-mission night.

Total sleep time tended to recover somewhat during the first day in Lyneham, then drop significantly again during the second night in Lyneham, before recovering over the last night in Lyneham. This can be explained by the fact that on arrival in England after a long crew day and carrying a sleep deficit from their last night at home, the crews were so tired that they were able to sleep relatively easily. However, after obtaining some restorative sleep upon arrival, by the second night in Lyneham, circadian rhythm disruption tended to make sleep more difficult. Here, 15% of the aircrew received less than 6 hours of sleep with one individual receiving only 259 minutes (4 hours and 19 minutes). By the time the crews returned from Zagreb for their last night in Lyneham their total sleep time was significantly increased and back to the levels evident two or three nights before the mission.

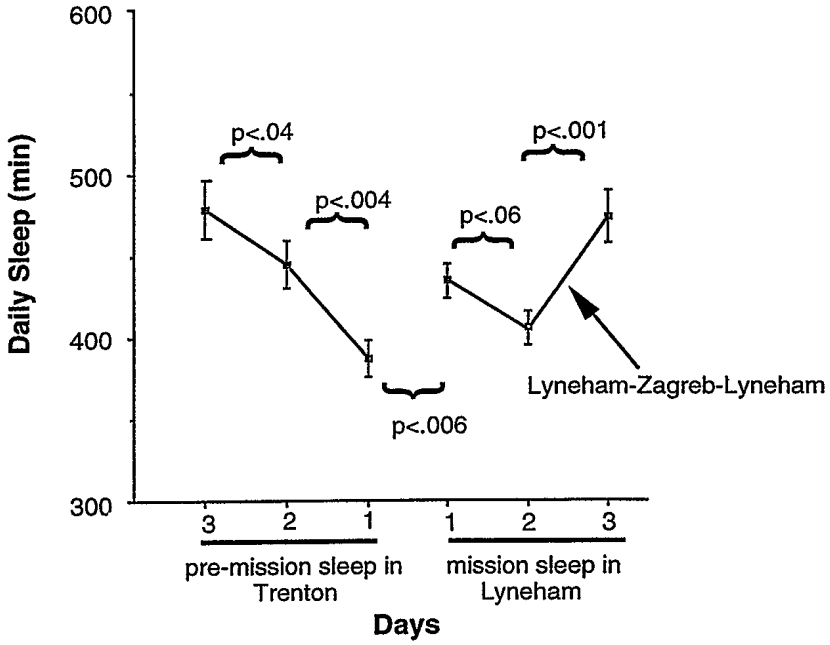

Figure 3. *Plot of Daily Sleep Minutes over pre-mission and mission days, collapsed over MOCs.*

The main effect of MOC was not significant however there was a significant interaction between MOC and days ($p<.0073$) which is shown in figure 4. The differences in total sleep time between MOCs shows that the loadmasters get less sleep than the other crew members and that they seem to obtain the same minimum levels of sleep from prior to, and throughout the mission. Further the Co-Pilots and the flight engineers get more sleep than other crew members, three days prior to the mission and the last night of the mission. During the last night at home before the mission and during the second night in Lyneham (when the circadian stress is most evident) all MOCs tend to obtain similar limited sleep.

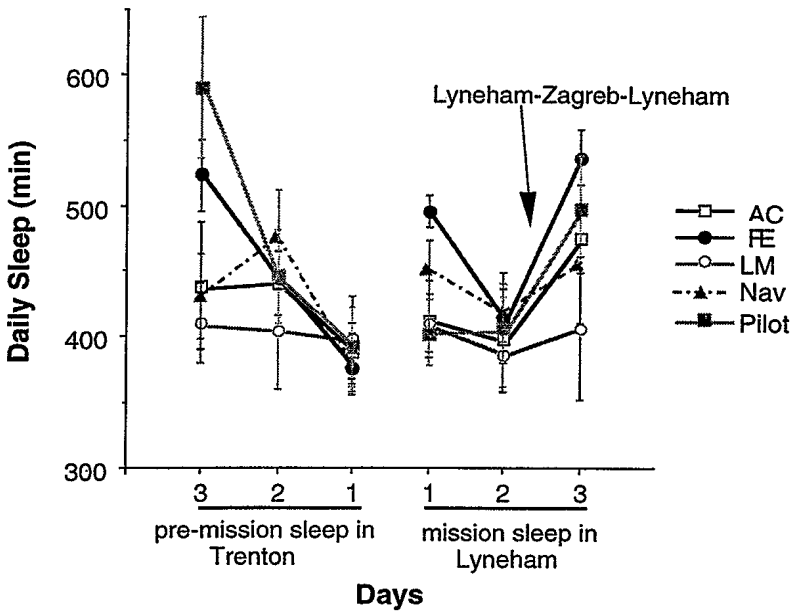

Figure 4. *Plot of Daily Sleep Minutes over pre-mission and mission days, for each MOC.*

For "number of sleep episodes/day" there was a significant main effect of days ($p<.0001$) which is illustrated in figure 5 and shows that the aircrews tended to have a greater number of sleep episodes per day while sleeping at home prior to the mission than they experienced while sleeping in Lyneham. The main effect of MOC was not significant, nor was the interaction between MOC and days.

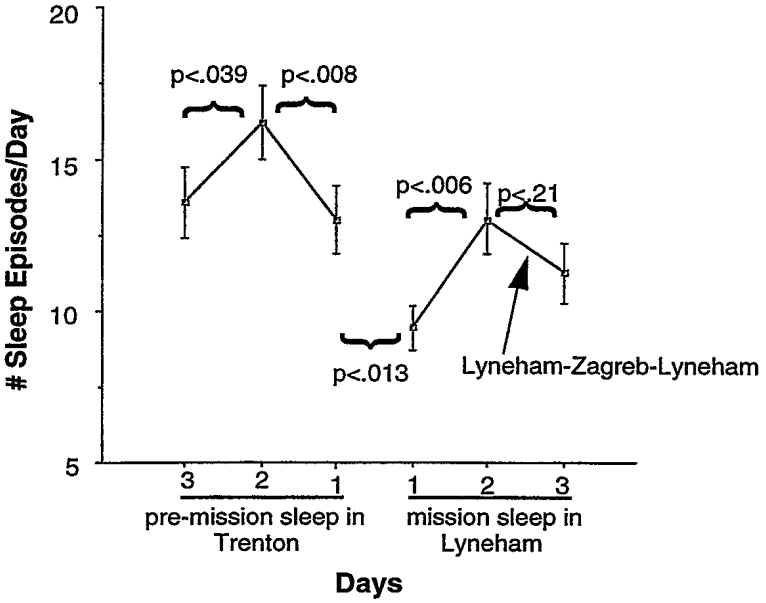

Figure 5. *Plot of # daily sleep episodes over pre-mission and mission days, collapsed over MOCs.*

For "duration of the daily mean sleep episode in minutes" there was a significant main effect of days ($p<.0104$) which is plotted in figure 6 and shows that the average sleep episode is shorter while at home than while overseas. Here also, the main effect of MOC was not significant, nor was the interaction between MOC and days. These data (figures 5 and 6) indicate that while the crews get less sleep overseas than they do at home, their overseas sleep architecture is altered such that they have fewer but longer sleep episodes which might suggest a more restorative type of sleep, perhaps compensating somewhat for less sleep.

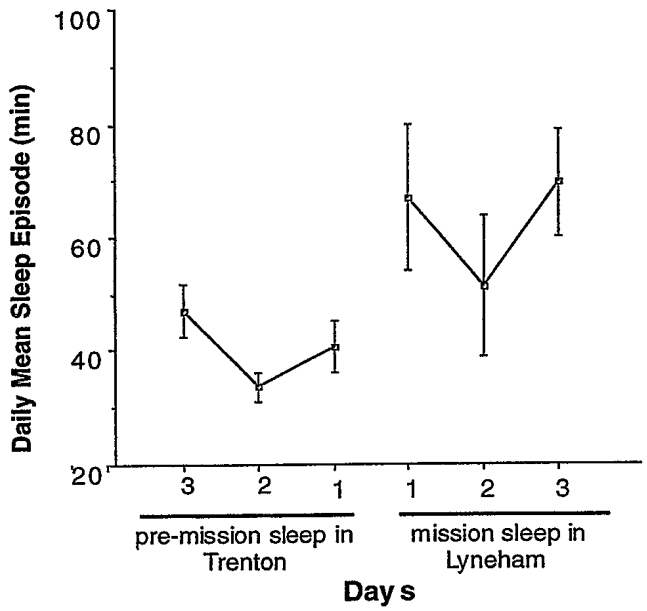

Figure 6. *Plot of Daily Mean Sleep Episode (in minutes) over pre-mission and mission days and collapsed over MOCs.*

Susops Questionnaire Data

Recall that the psychomotor task data (SUSOPS and multi-task) were collected only during the flights themselves. In order to directly compare performance between the two long and fatiguing transatlantic legs, these data (figures 7 to 13) were submitted to a two-factor repeated measures analysis of variance with one factor ('legs') having two levels (one outbound and one return transatlantic leg) and the other factor ('trials') having three levels (three trials per leg). The requirement to have a balanced analytical design dictated dropping the performance trial done on the Lyneham-Zagreb-Lyneham leg. Dropping this trial also facilitated the illustration of the results of Figures 7 to 13 by allowing the super-imposition of the 'leg' curves for the transatlantic legs.

For the Stanford Sleepiness Scale, the main effect of trials was significant ($p<.0001$) and is illustrated in figure 7. Neither the main effect of MOC nor the interaction of MOC and trials was significant. This indicates that with respect to the aircrews self rating of alertness (Stanford Sleepiness Scale), alertness decreases progressively (and in a similar manner across MOCs) during both transatlantic legs with the outbound leg (Trenton to Lyneham) tending to be more stressful in this regard than the return leg (Lyneham to Trenton) but not significantly so.

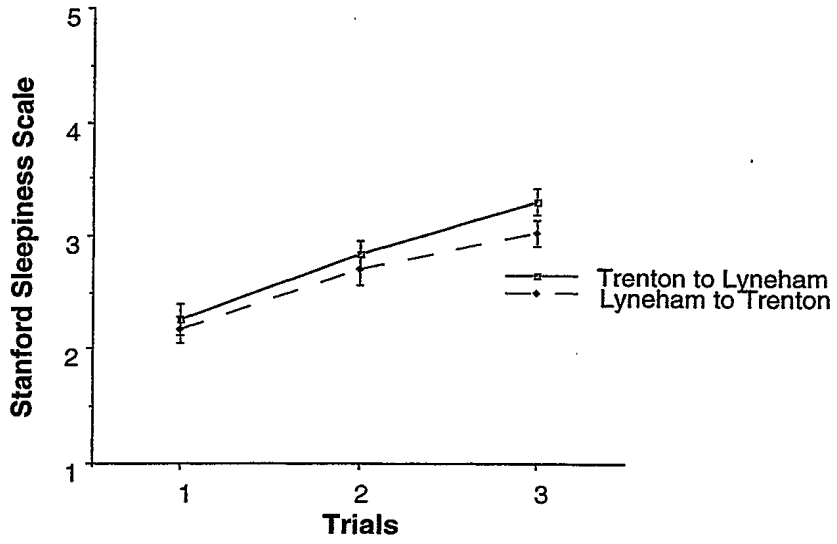

Figure 7. *Stanford Sleepiness Scale plotted over trials for departure (Trenton to Lyneham) and return (Lyneham to Trenton) legs.*

For "mental fatigue rating" the main effect of trials was significant ($p<.0001$) and is plotted in figure 8. The main effect of MOC was not significant, nor was the interaction of MOC and trials.

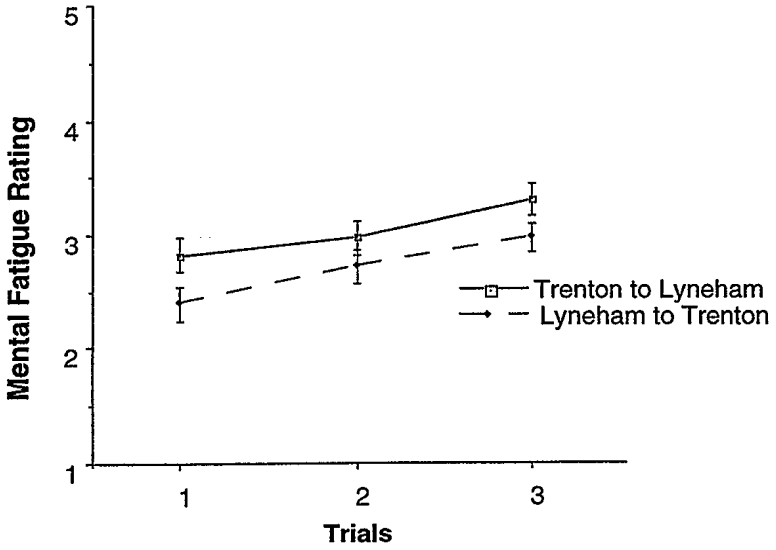

Figure 8. *Mental Fatigue Scale plotted over trials for departure (Trenton to Lyneham) and return (Lyneham to Trenton) legs.*

For "Physical Fatigue Rating" the main effect of trials was significant ($p<.0001$) and is plotted in figure 9. The main effect of MOC was not significant, nor was the interaction between MOC and trials.

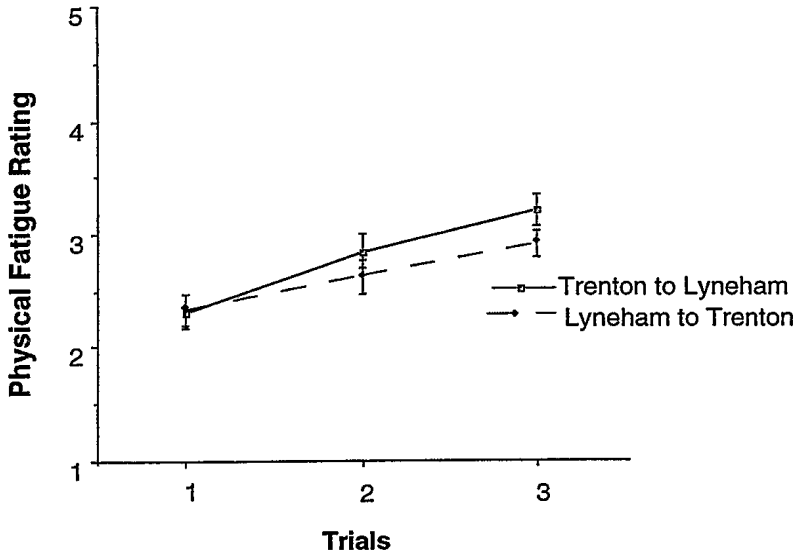

Figure 9. *Physical Fatigue Rating plotted over trials for departure (Trenton to Lyneham) and return (Lyneham to Trenton) legs.*

Essentially, self ratings of mental and physical fatigue tend to change over the transatlantic legs in a manner similar to the self ratings for alertness. This is to say that both mental and physical fatigue progressively increase over both transatlantic legs (in a similar manner across MOCs) with the outbound leg tending to be more stressful than the return leg, but not significantly so.

Susops Test Data

Serial Reaction Time (SRT)

For the number of correct responses to the SRT task, the main effect of trials was significant ($p<.0001$), the main effect of legs (departure from Trenton to Lyneham vs return from Lyneham to Trenton) was significant ($p<.0001$), and the interaction between trials and legs was also significant ($p<.0001$). Both of these main effects and this interaction are illustrated in figure 10. Further, trend analysis confirms that the curve for the return leg is quadratic ($p<.002$). The significant interaction between 'trials' and 'legs' indicates that even if performance is better on the return leg, the learning rate on the return leg is reduced, relative to the steeper slope of the outbound learning curve. The significant quadratic function for the return curve indicates that the curve tends to become flat between trials two and three of the return leg, suggesting that performance has almost attained asymptote (fully trained, optimum performance). However, for this task, there are no obvious performance decrements over time (trials). This suggests that the serial reaction time task is not sensitive to the level of fatigue encountered on these missions. Even though the subjects reported subjective levels of fatigue, they may have been able to rally their cognitive resources to compensate it. If this task was performed for a longer duration (rather than 3 minutes) perhaps fatigue effects would become apparent.

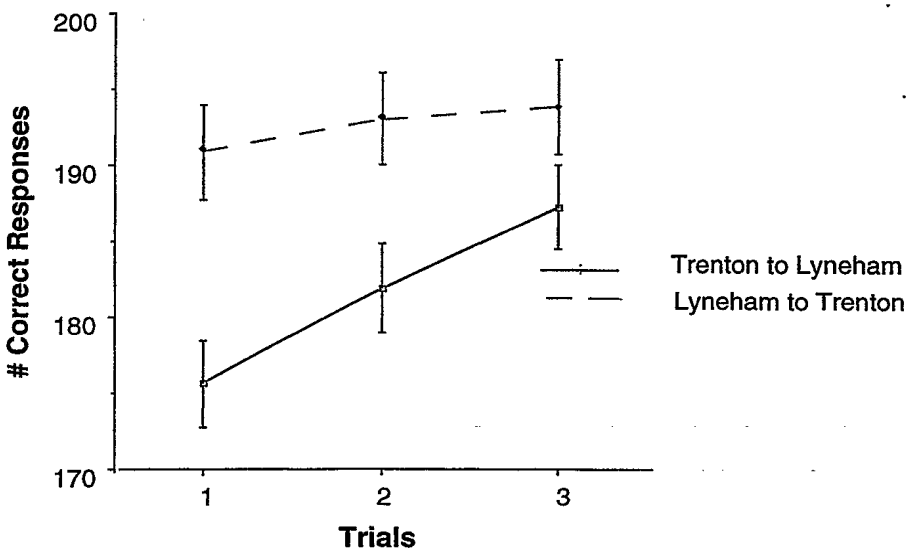

Figure 10. *Number of correct responses to Serial Reaction Time task plotted over trials for departure (Trenton to Lyneham) and return (Lyneham to Trenton) legs.*

Logical Reasoning task (LRT)

For the number of correct responses to the LRT task , the main effect of trials was significant ($p<.0001$), the difference between the departure and return legs is significant $p<.0001$), and the interaction between the trial effect and the leg effect is also significant ($p<.028$) These two main effects and the interaction between them are illustrated in figure 11. Polynomial analysis reveals that the trends representing each of the departure and return legs are singificant quadratic functions ($p<.016$).

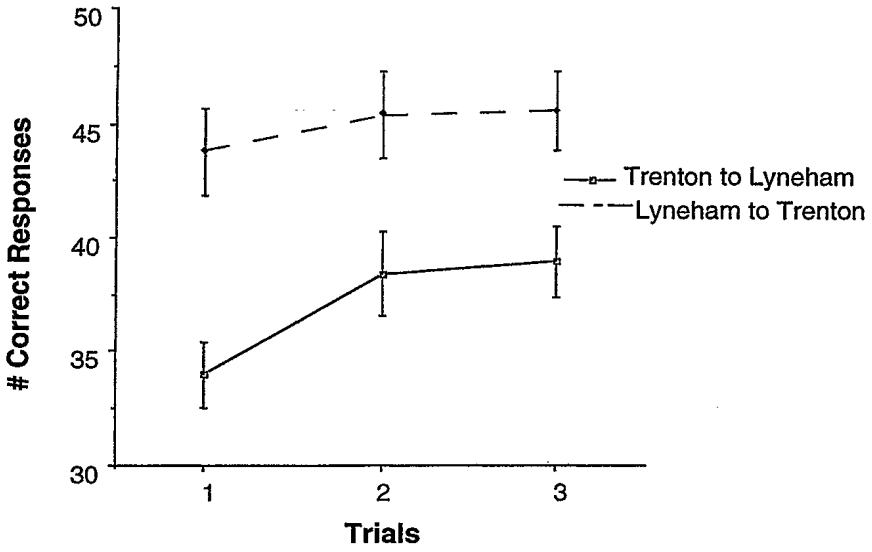

Figure 11. *Number of correct responses to the Logical Reasoning Task plotted over trials for departure (Trenton to Lyneham) and return (Lyneham to Trenton) legs.*

The significant interaction between trials and legs indicates that the learning rate for this task, like that for the reaction time task, is higher during the outbound leg than during the return leg. However, trend analysis of this particular data reveals that, unlike the serial reaction time performance (where only the return curve was quadratic) the curves for each of the outbound and return legs are quadratic. This effect can be seen in that for each of these curves the performance falls off after the second trial. For the outbound leg this is arguably an effect on higher level thinking due to the impact of fatigue. For example, if fatigue was not disrupting performance during the outbound leg, one would expect a linear increase in performance throughout the three trials of that leg, similar to the linear outbound curve for the serial reaction time performance shown in figure 10. Here the average value for trial 3 of the outbound leg (figure 10) is quite similar to the average value for trial 1 of the return leg. For figure 11, the quadratic curve for the outbound leg, illustrates a fatigue-induced attenuation of learning in the very early part of the learning curve when one when expect facile learning, barring the presence of some stressor, such as fatigue. For the return leg, although fatigue may be a factor here as well, it is also possible that the flattening of the curve could be suggestive of approaching asymptote.

Serial Subtraction task (SST)

For the number of correct responses to the SST task, the main effect of trials was significant (p<.0001), the difference between the departure and return legs is significant (p< .0001), and the interaction between the trial effect and the leg effect is alsosignificant (p<.0.0001) These two main effects and the interaction between them are illustrated in figure 12. Trend analysis indicates that only the return curve is quadratic (p<.002). The significant interaction between trials and legs indicates that the learning rate is higher during the outbound leg than during the inbound leg (again similar to the performance of the serial reaction time and logical reasoning tasks). Trend analysis reveals that only the return curve for this task is quadratic. Although it is possible that this could be due to fatigue, the most likely explanation for this drop in performance after the second trial of the return leg is that the subjects are approaching asymptote performance. For this task, there appears to be no compelling evidence of an impact of fatigue on performance.

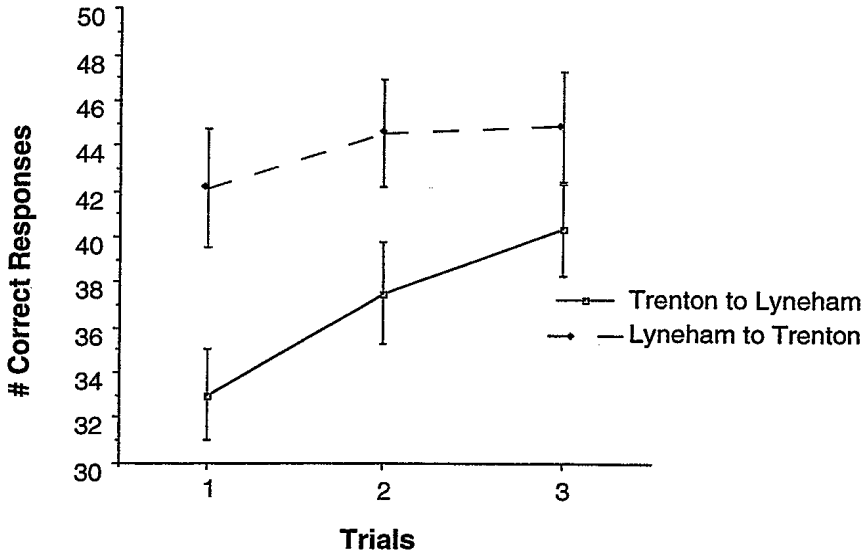

Figure 12. *Number of correct responses to Serial Subtraction task plotted over trials for departure (Trenton to Lyneham) and return (Lyneham to Trenton) legs.*

Of the three SUSOPS tasks (serial reaction time, logical reasoning, and serial subtraction) the only task whose performance is sensitive to the level of fatigue encountered during these missions is the logical reasoning task.

Multi-task Data

With respect to the multi-task "score", the main effect of trials was significant (p<.0003), the difference between the departure and return legs is significant (p<.0001), while the interaction between the trial effect and the leg effect falls short of significance (p<.0.094) These two main effects and the interaction between them are illustrated in

figure 13. Trend analysis indicates that each of the departure and return curves are quadratic ($p<.026$).

The significant trials effect illustrates the expected learning effect for the multi-task, while the significant leg effect indicates the expected higher level of performance on the return leg relative to the outbound leg (figure 13). The significant quadratic function for each of the outbound and return legs indicates a major impact of fatigue at trial three of the outbound leg and at trial two of the return leg which recovers to the levels of trial one of the same leg, before arriving home at the end of the mission. These data would lend themselves to the following interpretation: On the long transatlantic outbound leg, after a night of minimal sleep, the level of fatigue among the subjects performing this task (Aircraft Commanders and Co-Pilots) is such that there is a very demonstrable and deleterious impact on performance. Here, further learning of the task (even at this early part of the learning curve where the expectation is that learning will occur relatively easily) is interrupted and performance actually falls. For the return transatlantic leg, after a better sleep than they obtained the night before they started the outbound leg, learning is again interrupted, this time at the second trial, and performance recovers before landing at home in Trenton, probably because of the psychological boost from the knowledge that the mission is almost over, which allows them to marshal "extra resources".

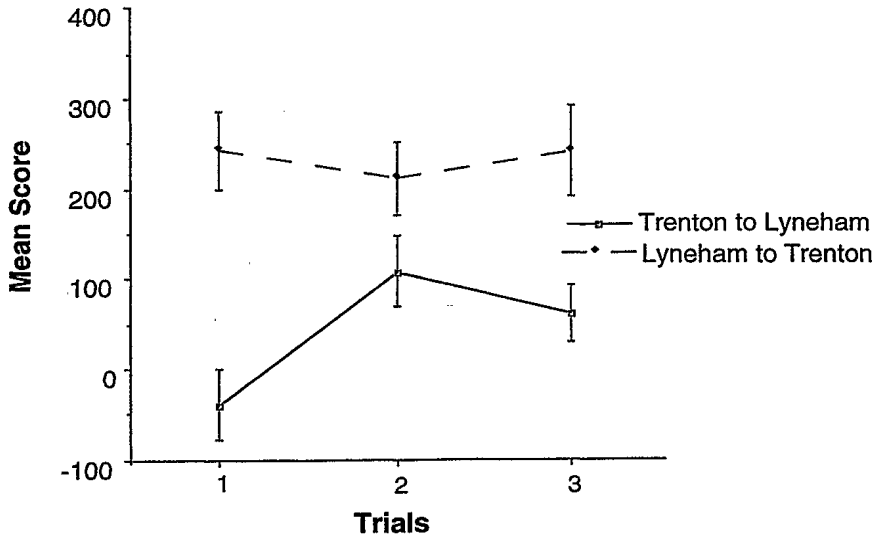

Figure 13. *Multi-task scores plotted over trials for departure (Trenton to Lyneham) and return (Lyneham to Trenton) legs.*

While the main effect of 'position' (pilot or co-pilot) is not statistically significant ($p<.125$), none-the-less, the relatively large differences in performance between the aircraft commanders and their co-pilots is illustrated in figure 14. While this effect of MOC falls short of statistical significance, the performance differences between

Aircraft Commanders and their Co-Pilots are quite large and would possibly have been statistically significant if the data had been a little less variable, within and between subjects.

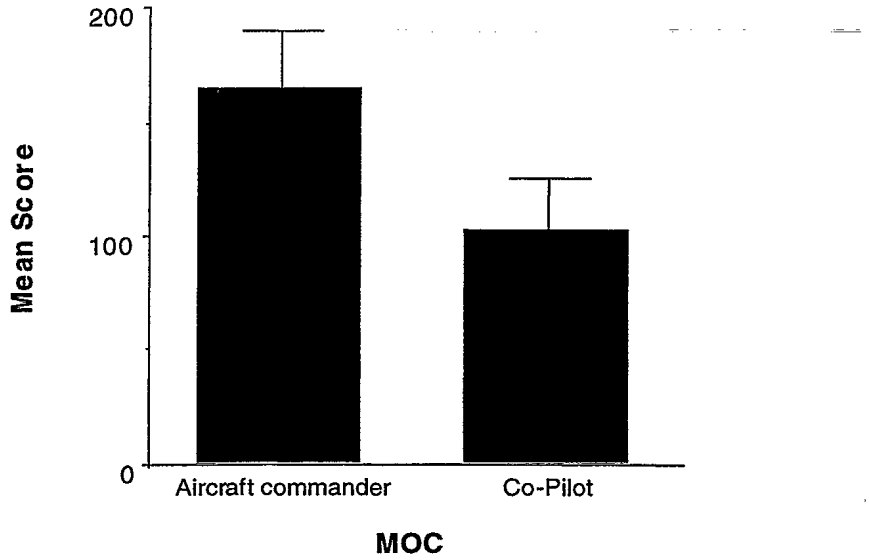

Figure 14. *Multi-task scores of air craft commanders vs co-pilots collapsed over all trials.*

The interaction between MOC and trials is almost significant ($p<.068$) and these differences are plotted in figure 15. Post hoc t-test comparisons for this interaction reveal that the multi-task performance difference between the aircraft commanders and the co-pilots is significant for the 2nd trial of the outbound leg ($p<.04$) and almost significant for the first trial of the return leg ($p<.14$). This interaction illustrates that the Aircraft Commanders performed much better than their Co-Pilots at trial two of the outbound leg but by trial three of that leg, their performance was reduced to a level similar to their Co-Pilots. For the return leg the Aircraft Commanders appeared to start out at a higher level of performance which by the second trial decreased to the levels of their Co-Pilots. Although the Co-Pilots tended to have a lower overall rate of learning than the Aircraft Commanders, their lower rate of learning was constant over the mission such that their learning rate did not appear to be affected by fatigue. One could perhaps suggest that while the Aircraft Commanders tended to perform better than their Co-Pilots, their higher rate of learning was impacted by fatigue whereas the lower learning rate of the Co-Pilots appears to be robust to their level of fatigue. Perhaps the "weight of command" on the Aircraft Commanders interacted with their fatigue level to produce this effect.

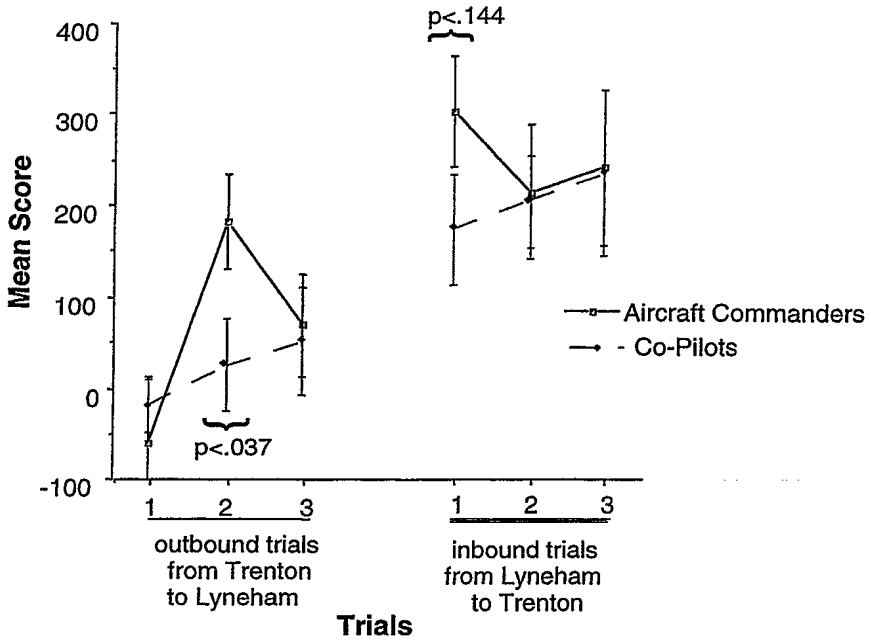

Figure 15. *Multi-task scores for each of Aircraft Commanders and their Co-Pilots plotted over trials for departure (Trenton to Lyneham) and return (Lyneham to Trenton) legs.*

As evidenced by the relatively large standard error bars on the multi-task graphs (figures 13, 14, and 15), there was a large variability in the performance of this task, without which the main effect of MOC, and the interaction between MOC and trials would probably have attained statistical significance. To illustrate some of the large individual differences and therefore the relatively large subject variability (both within and between subjects) in multi-task performance, data from individual aircraft commanders and co-pilots are super-imposed on the main effect of trials for figures 16 through 33. The large variability both within and between subjects probably reflects that, as is well know, some individuals are very sensitive to fatigue while others are relatively fatigue resistant (8, 13, 14, 23). The data from each trial for individual subjects superimposed on the group trial average data show that some individuals within our subject population are sensitive to the fatigue encountered on these missions while some of their counterparts are not.

Subject 1 (figure 16) was tested only once on the outbound leg to Lyneham and produced a score just slightly above the average score for this first trial. The next opportunity for testing this subject was on the Lyneham-Zagreb-Lyneham leg and his performance was remarkably poor (even considering that this was only his second trial). Clearly something affected his performance (likely fatigue and jet-lag) on this trial and his performance improved on the return leg from Lyneham to Trenton, but even this level of performance was much lower than the average performance for these trials (again, even considering that these were only his third and fourth trials).

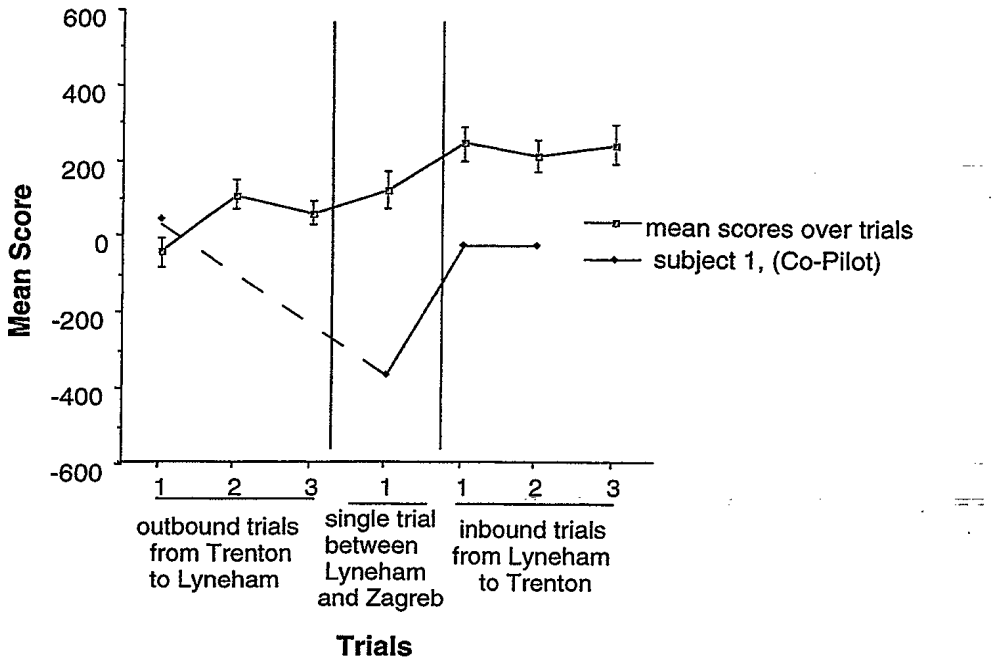

Figure 16. *Illustration of multi-task scores for subject 1 (Co-Pilot) super-imposed on the mean group scores (Aircraft commanders & Co-Pilots) over trials. Dotted line indicates missing Subject 1data for outbound trials 2 and 3.*

Subject 2 (figure 17) seems to be somewhat sensitive to fatigue in that he started out performing a little above average on the first trial but by the second was lower than average and by the third and final trial of the long outbound leg his performance was much lower than the average for outbound trial three. His performance on the single trial during the Lyneham-Zagreb-Lyneham leg was even worse. However, for the return leg to Trenton, while he started at a slightly lower level of performance for trial 1 of that leg, by the end of that leg his performance was above average. Overall, an interpretation of this subject's performance might be that the minimum sleep we documented for the crews the night before departure from Trenton might be the reason for his performance on the outbound leg, and the changing sleep patterns during the first and second nights in England might account for his poor performance on the Zagreb day. The last night of sleep in England (figure 2) (which produced the longest sleep times since three nights prior to commencement of the mission) seemed to be restorative to the point where performance begins to improve and continues to improve during the long flight home.

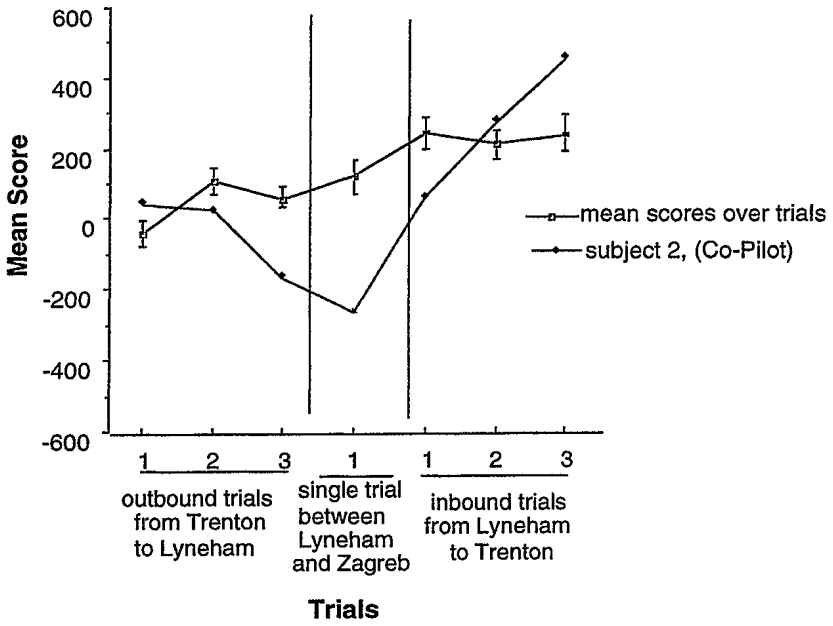

Figure 17. *Illustration of multi-task scores for subject 2 (Co-Pilot) super-imposed on the mean group scores (Aircraft commanders & Co-Pilots) over trials.*

The multi-task performance of subject 3 (figure 18) seems to suggest that he is fatigue resistant. On the outbound leg, his performance improved continuously over the three trials, unlike the group average which showed an initial improvement from trial 1 to trial 2 and then a drop in performance for trial three. Even though this subject was not tested on the Lyneham-Zagreb-Lyneham leg, his performance on the return leg (Lyneham to Trenton) mirrored the group average performance for the first 2 trials of that leg but by the last trial, his performance was much higher than the group average. This increase in performance during the last trial of the return leg is evident for nine of the eighteen pilots and could be an indication that they their appear to be able to "marshal extra resources" for the push home in that they become "pumped" as they approach the end of the mission.

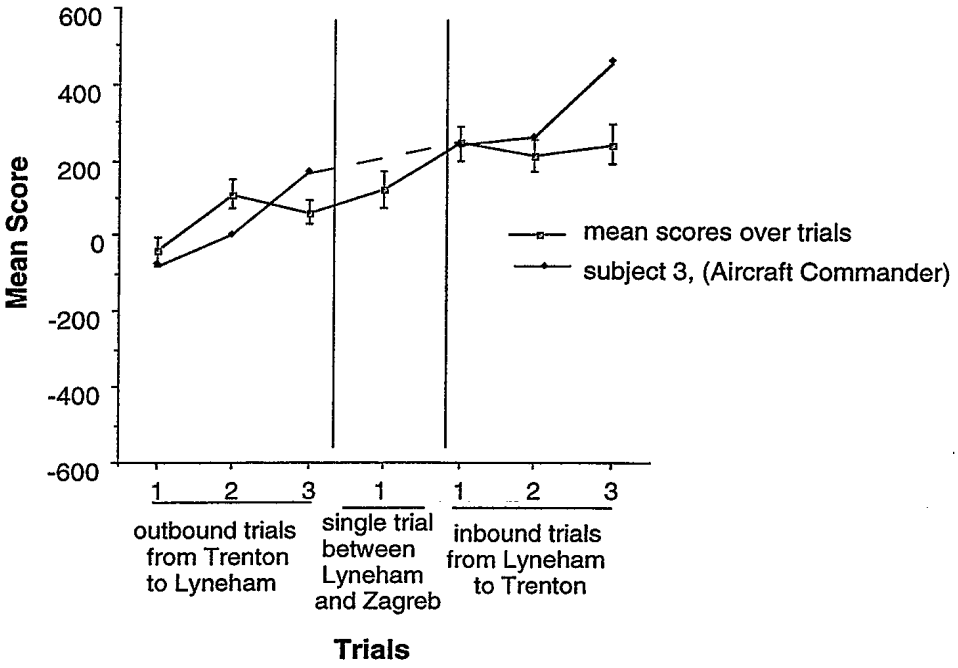

Figure 18. *Illustration of multi-task scores for subject 3 (Aircraft Commander) super-imposed on the mean group scores (Aircraft commanders & Co-Pilots) over trials. Dashed line illustrates missing subject 3 data on Lyneham-Zagreb-Lyneham leg.*

The performance of subject 4 (figure 19) over the outbound and Lyneham-Zagreb-Lyneham legs is alternately slightly above and slightly below but none-the-less similar to the group average for those trials. By the return leg, this subject's performance was clearly above the group average over all three trials. He is one of the nine subjects who show an increase in performance for trial three (of the return leg) after a long crew day.

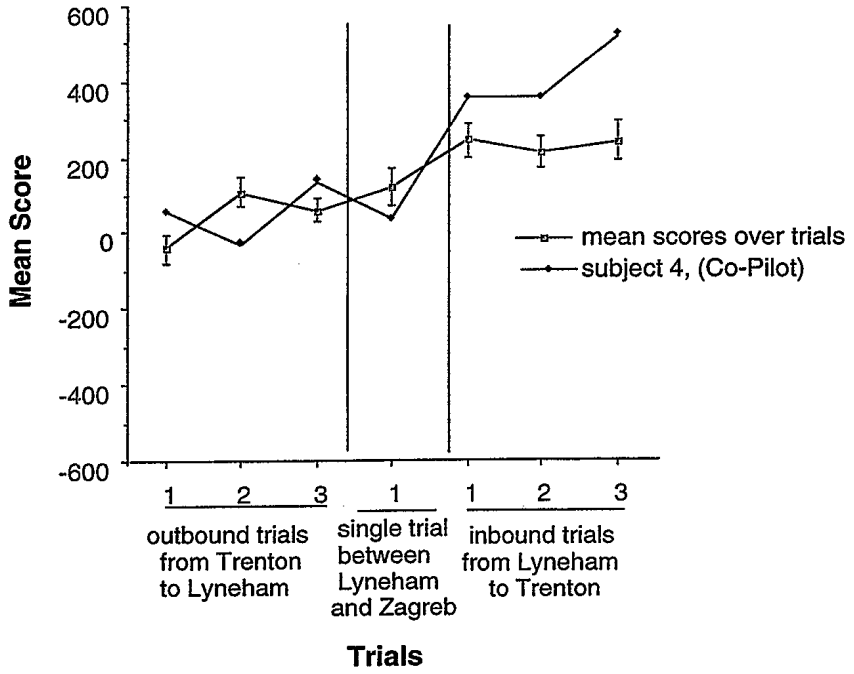

Figure 19. *Illustration of multi-task scores for subject 4 (Co-Pilot) super-imposed on the mean group scores (Aircraft commanders & Co-Pilots) over trials.*

Subject 5 (figure 20) performed at a level similar to the group average over the first 2 trials of the outbound leg but was not tested again until the return leg. His performance on the return leg started out much lower than the group average, (and at the level he obtained in his first outbound trial) but he improved continuously over that leg such that by the time he completed the last trial, his performance was almost at the level of the group average.

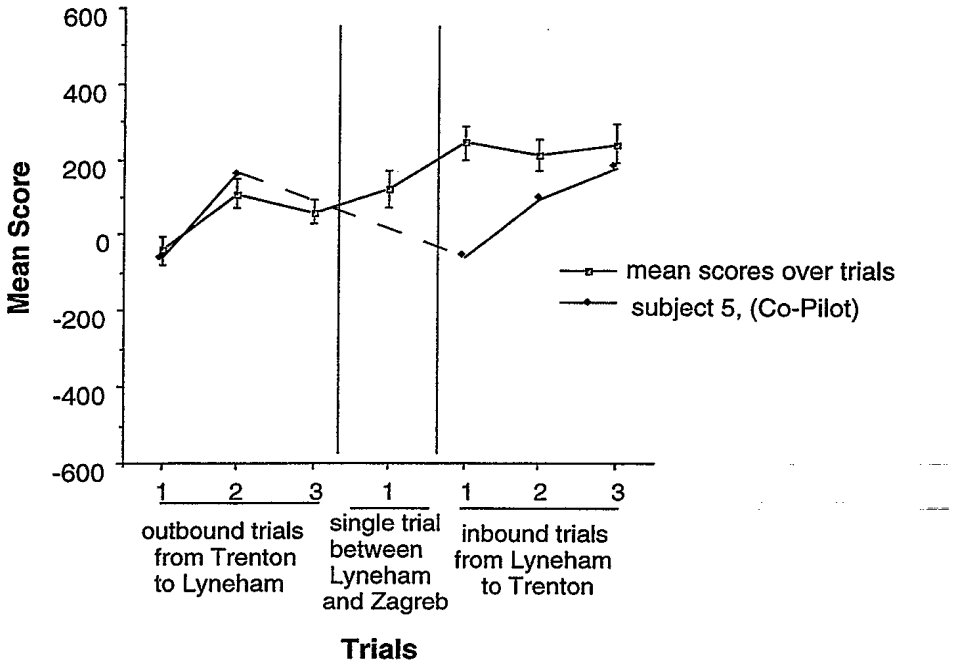

Figure 20. *Illustration of multi-task scores for subject 5 (Co-Pilot) super-imposed on the mean group scores (Aircraft commanders & Co-Pilots) over trials. Dotted line indicates missing Subject 5 data for outbound trial 3 and Zagreb trial.*

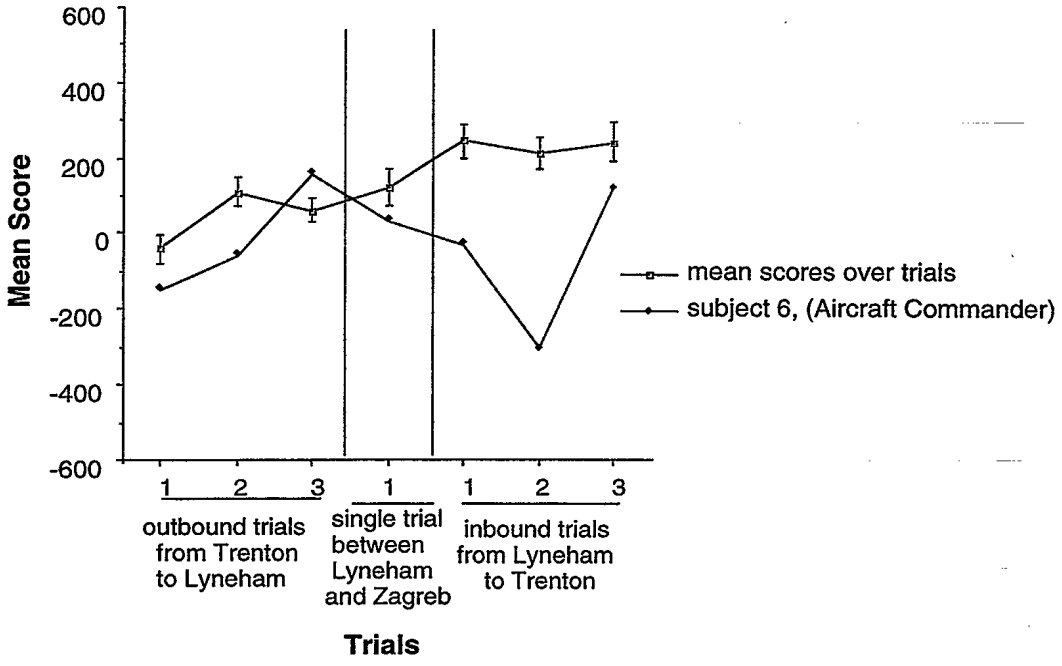

Figure 21. *Illustration of multi-task scores for subject 6 (Aircraft Commander) super-imposed on the mean group scores (Aircraft commanders & Co-Pilots) over trials.*

Subject 6 (figure 21) performed below the level of the group average except for the third trial of the outbound leg. On the Lyneham-Zagreb-Lyneham leg his performance dropped relative to his last trial on the outbound leg and dropped even further at trial one of the return leg, and then dropped precipitously at trial two of that leg. However he rallied remarkably for the last trial such that his performance was approaching the group average for that trial.

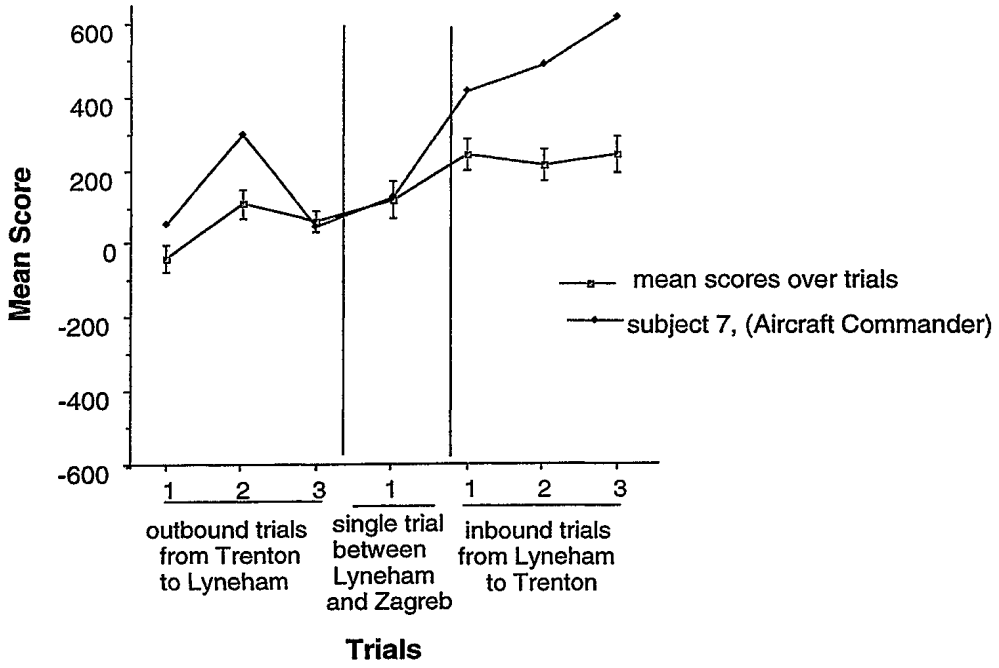

Figure 22. *Illustration of multi-task scores for subject 7 (Aircraft Commander) super-imposed on the mean group scores (Aircraft commanders & Co-Pilots) over trials.*

The performance of subject 7 (figure 22) over the first two outbound trials is better than the group average but by the last trial of that leg his performance is even lower than it was for the first trial, indicating a significant impact of fatigue towards the end of a long crew day with limited sleep the previous evening. For the Lyneham-Zagreb-Lyneham leg his performance is essentially identical to the group average. His performance during the return leg is higher than the group averages for those three trials and improves with each successive trial.

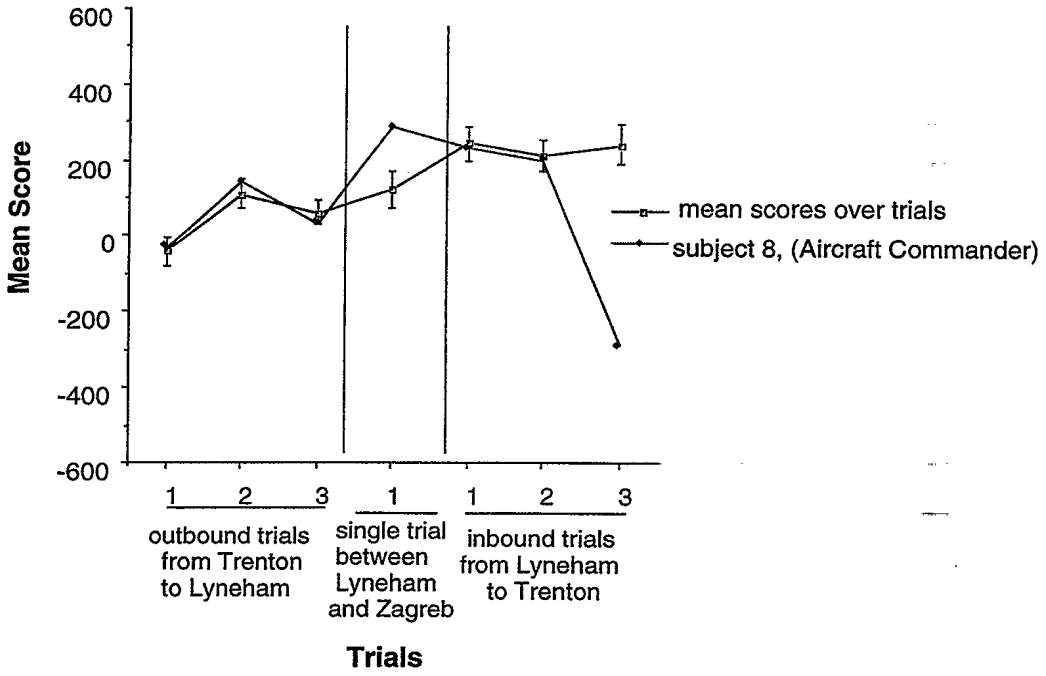

Figure 23. *Illustration of multi-task scores for subject 8 (Aircraft Commander) super-imposed on the mean group scores (Aircraft commanders & Co-Pilots) over trials.*

During the outbound leg, the performance of subject 8 (figure 23) is the same as the group average over all three trials, thus showing the drop in performance at the third trial (which also occurs towards the end of the long crew day after limited sleep the previous night). During the Lyneham-Zagreb-Lyneham leg his performance is somewhat higher than the group average for that particular trial. While on the return leg, his performance closely mirrors the group average for the first two trials of that leg and then falls drastically at the third trial which suggests that this subject's performance is quite susceptible to fatigue.

On the outbound leg subject 9 (figure 24) performs below the group average at trial one, exceeds the group average at trial two and then at trial three, like seven of his colleagues, he shows a marked drop in performance, again arguably because he is approaching the end of a long crew day after a night a minimal sleep. This subject's performance then improves to a level higher than the group average on the Lyneham-Zagreb-Lyneham leg. While his performance at trial one of the return leg is the same as the group average, he improves to well above the group average by trial two and then shows a sensitivity to fatigue by his drop in performance at trial three of this return leg.

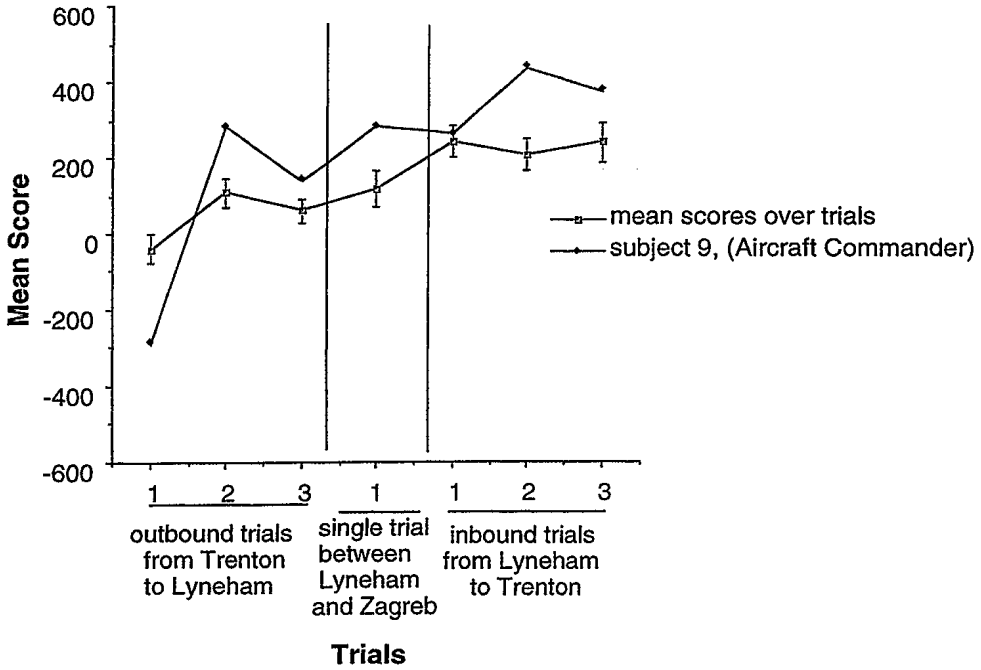

Figure 24. *Illustration of multi-task scores for subject 9 (Aircraft Commander) super-imposed on the mean group scores (Aircraft commanders & Co-Pilots) over trials.*

While subject 10 (figure 25) performs at the group average for trial one of the outbound leg he barely improves at the next trial and by the third trial his performance is lower than his first trial which suggests that fatigue has impacted on him such that his learning of this novel task is totally disrupted. This subject's performance recovers to somewhat better than the group average during the Lyneham-Zagreb-Lyneham leg. On the return leg his performance starts out at a level much lower than the group average for trial one, rises to almost the group average by trial two, and then he shows a marked sensitivity to fatigue by a precipitous drop in performance (to the level of his first ever trial on this task on departure from Trenton).

On the outbound leg subject 11 (figure 26) performs at the group average for trial one, improves to well above the group average at trial two and then his performance falls back down to the level of the group average by trial three. On the Lyneham-Zagreb-Lyneham leg his performance is somewhat higher than the group average. His performance on the return leg is very similar to the group average, including the small increase in performance at trial three of that leg which, as mentioned earlier, we feel indicates a mood/motivation modified improvement in performance given that the end of the mission is imminent.

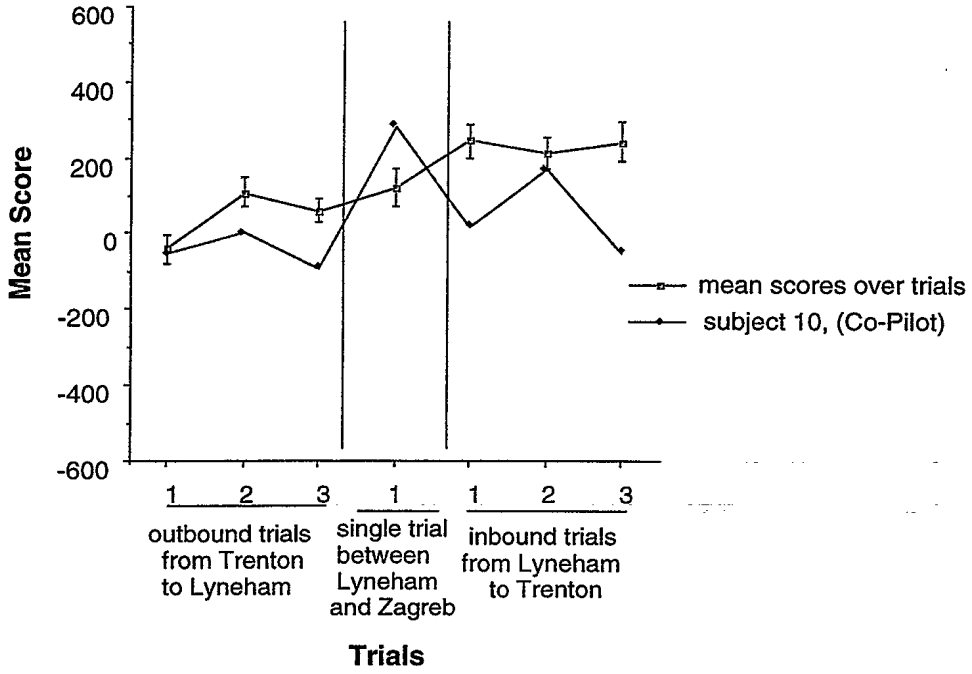

Figure 25. *Illustration of multi-task scores for subject 10 (Co-Pilot) super-imposed on the mean group scores (Aircraft commanders & Co-Pilots) over trials.*

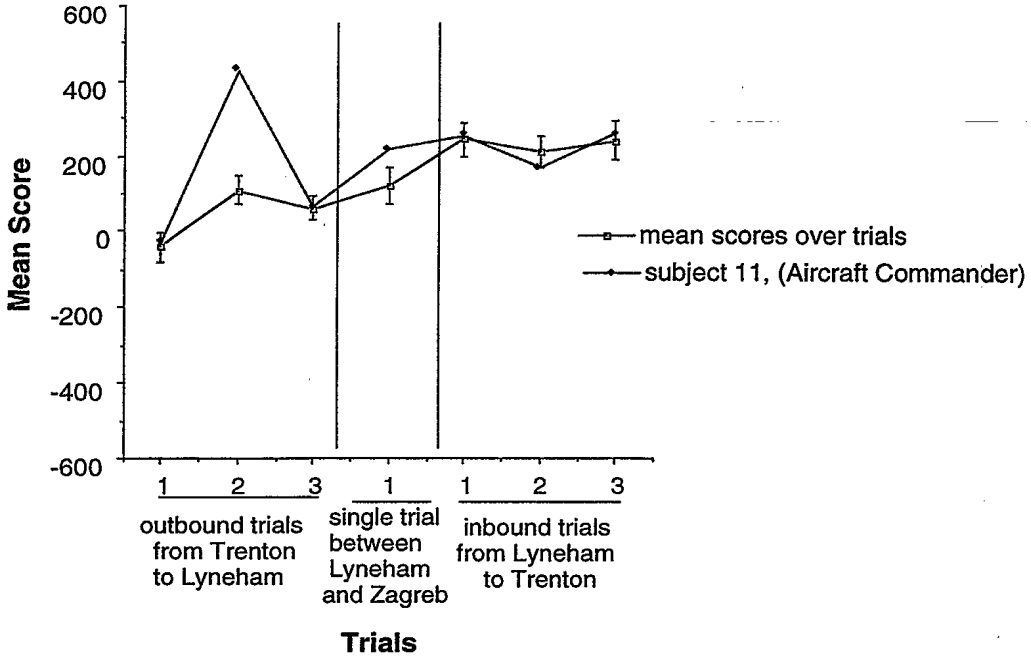

Figure 26. *Illustration of multi-task scores for subject 11 (Aircraft Commander) super-imposed on the mean group scores (Aircraft commanders & Co-Pilots) over trials.*

While subject 12 (figure 27) performs at a level similar to the group average for trials one and two of the outbound leg, his performance doesn't drop at trial three to the same level as the group average. However, at this trial his skill acquisition is definitely attenuated, probably by fatigue, in this very early part of the learning curve when increases in performance are normally very apparent, (at least in the absence of stressors such as fatigue). During the Lyneham-Zagreb-Lyneham leg his performance is slightly above the group average. On the return leg his performance starts out at the group average for trial one, increases above the group average for trial two, and then drops precipitously, no doubt in response to fatigue at trial three.

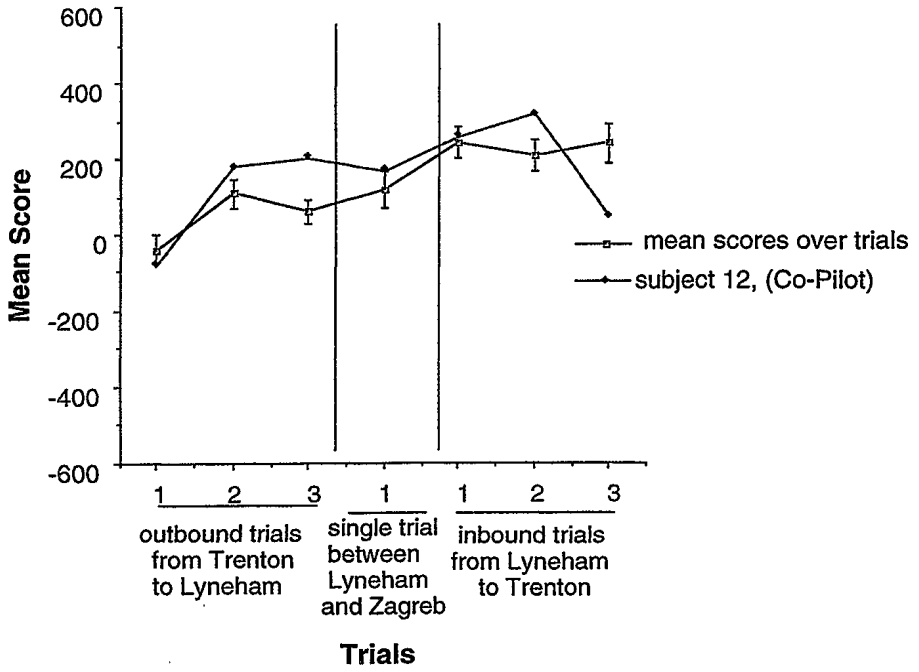

Figure 27. *Illustration of multi-task scores for subject 12 (Co-Pilot) super-imposed on the mean group scores (Aircraft commanders & Co-Pilots) over trials.*

Subject 13 (figure 28) is only tested twice on the outbound leg, not tested during the Lyneham-Zagreb-Lyneham leg, but completes the three trials of the return leg. For the departure leg his trial one performance is well above the group average, and falls to below the group average by trial two. On the return leg his performance is at the level of the group average at trial one, drops markedly at trial two, and rallies somewhat at trial four.

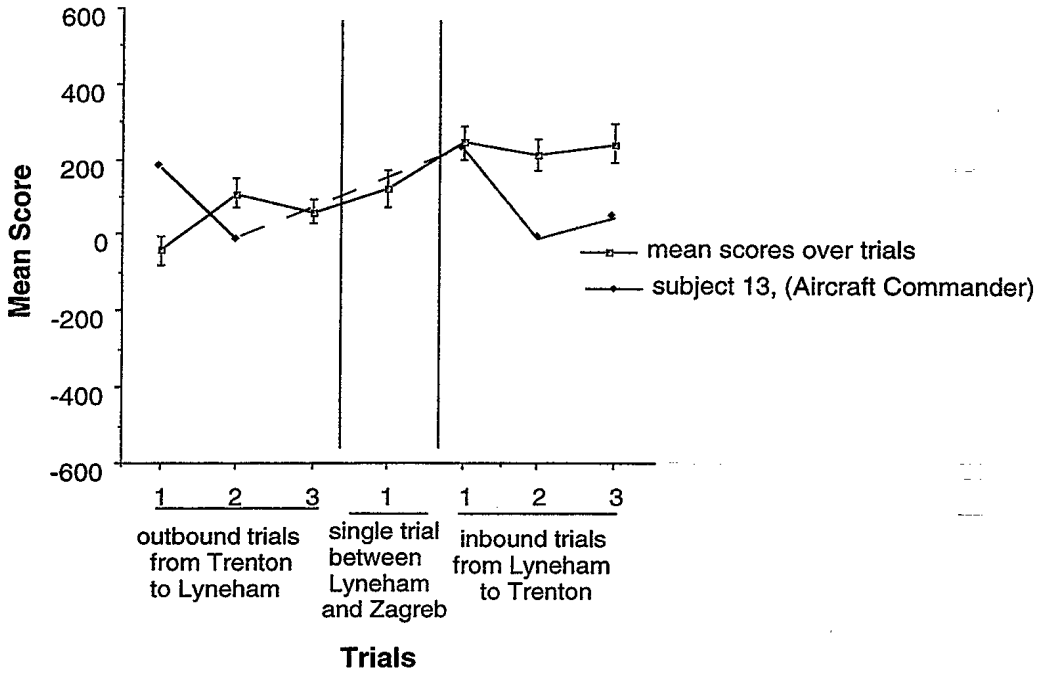

Figure 28. *Illustration of multi-task scores for subject 13 (Aircraft Commander) super-imposed on the mean group scores (Aircraft commanders & Co-Pilots) over trials. Dotted line indicates missing subject 13 data for outbound trial 3 and Zagreb trial.*

While the performance of subject 14 (figure 29) doesn't drop at trial three of the outbound leg, he demonstrates the same attenuation of learning at this trial as that exhibited by subject 12 (figure 26). His performance during the single trial of the Lyneham-Zagreb-Lyneham leg is well above the group average. On the return leg his performance at trial one is much higher than the group average, and falls continuously over the last two trials, reaching the level of the group average by the last trial, arguably in response to fatigue.

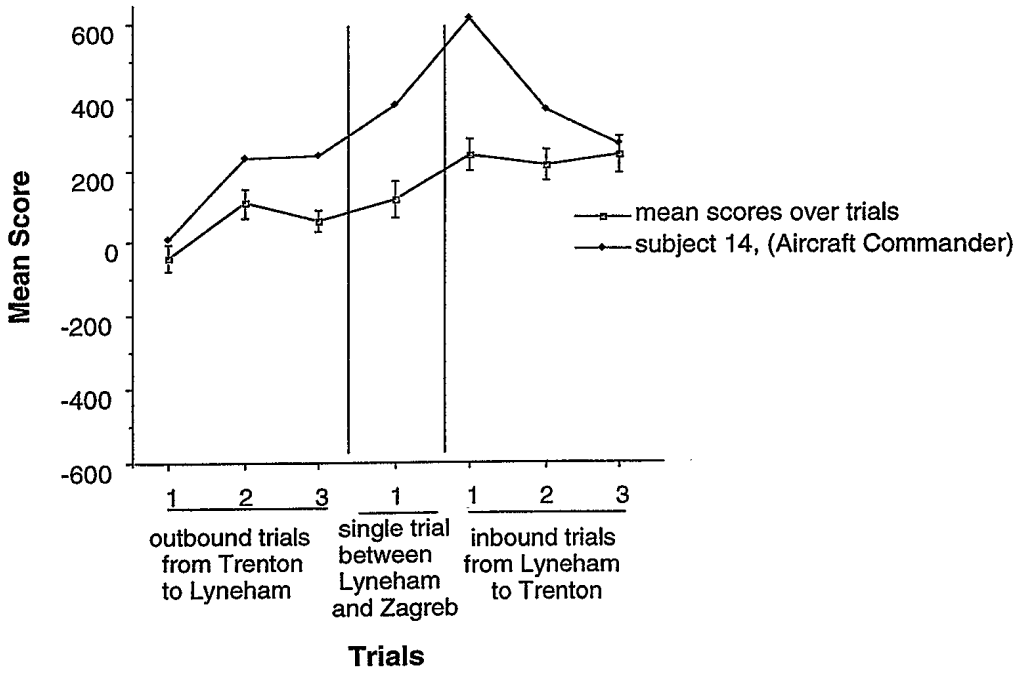

Figure 29. *Illustration of multi-task scores for subject 14 (Aircraft Commander) super-imposed on the mean group scores (Aircraft commanders & Co-Pilots) over trials.*

Subject 15 (figure 30) has a multi-task performance slightly above the group average at trial one of the outbound leg which drops significantly at trial two and rebounds slightly above the group average by trial three. His performance on the Lyneham-Zagreb-Lyneham leg is well above the group average. While his performance on the return leg started out higher than the group average at trial one of that leg, not much more can be said of his return leg performance because he was not tested after the first trial of the that leg.

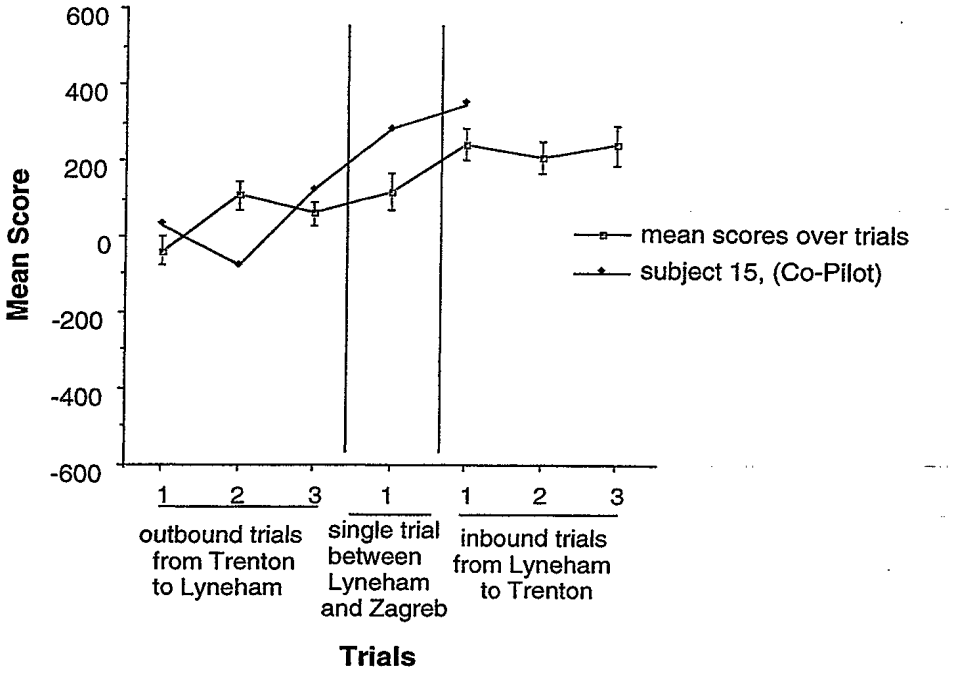

Figure 30. *Illustration of multi-task scores for subject 15 (Co-Pilot) super-imposed on the mean group scores (Aircraft commanders & Co-Pilots) over trials.*

The performance of subject 16 (figure 31) is noteworthy because he was the only subject from this study who flew two missions, separated by approximately seven weeks. On his first mission, his performance on the first trial of the outbound leg is very poor, and by the second trial he is performing slightly above the group average. However, his performance drops by trial three, which is arguably a fatigue-induced drop in performance (even though it is a slightly better performance than the group average for this trial) at the end of the long outbound leg. His performance on the Lyneham-Zagreb-Lyneham leg is better than the group average. On the return leg his performance starts at the same level as the group average for trial one, then increases to a level above the group average at trial 2 and shows a slight drop in performance at trial three which could possibly be interpreted as a training asymptote, except that on his second mission he performs at a much higher level during the first and second outbound trials and the first and third return trials. Further, for the second mission, his first outbound trial performance is at a level slightly higher than the last trial of his first mission. His performance then improves somewhat at the second outbound trial and drops very precipitously on the last trial of that day. A check of this subject's sleep actigraph data reveals that for each of the last nights in Trenton before the start of each mission he received exactly the same amount of sleep (390 minutes). In spite of this fact, there is a very large decrement on performance at the third trial of the outbound leg (when any impact of fatigue would be most obvious) for mission two but not for mission one. There was obviously some stressor other than fatigue that was present on the second mission but not on the first mission. For Lyneham-Zagreb-Lyneham leg of the second mission, his performance was still very poor. By the return leg of this second mission, performance was very high at trial one, dropped down to the group average at trial two, and rallied again to a high level for the last trial. Unfortunately the subject's actigraph data for the sleep periods in England is missing for both missions so it is impossible to

illuminate these performance changes on the basis of his sleep history.

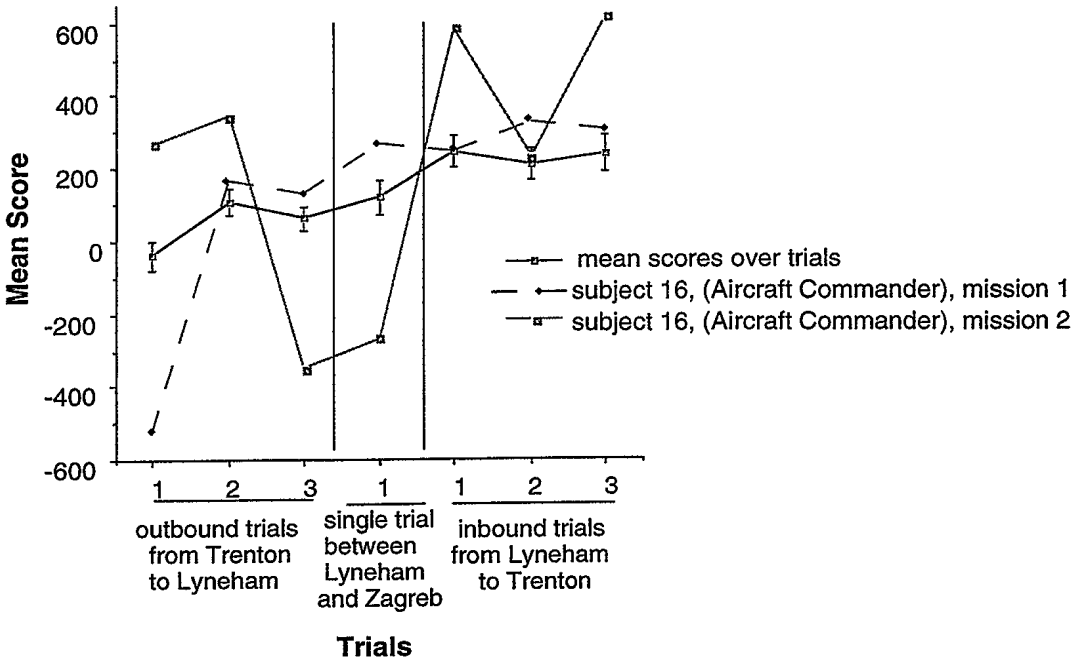

Figure 31. *Illustration of multi-task scores for subject 16 (Aircraft Commander) super-imposed on the mean group scores (Aircraft commanders & Co-Pilots) over trials. This subject flew 2 of these missions separated by 7 weeks.*

The performance of subject 17 (figure 32) is similar to that of subject 15 (figure 29) in that for the outbound leg trial one score is higher than the group average but drops markedly at trial two, and recovers to the group average by trial three. Like subject 15, his performance is also better than the group average for the Lyneham-Zagreb-Lyneham leg. On the return leg, he performed at a level very similar to the group average for the three trials of that leg.

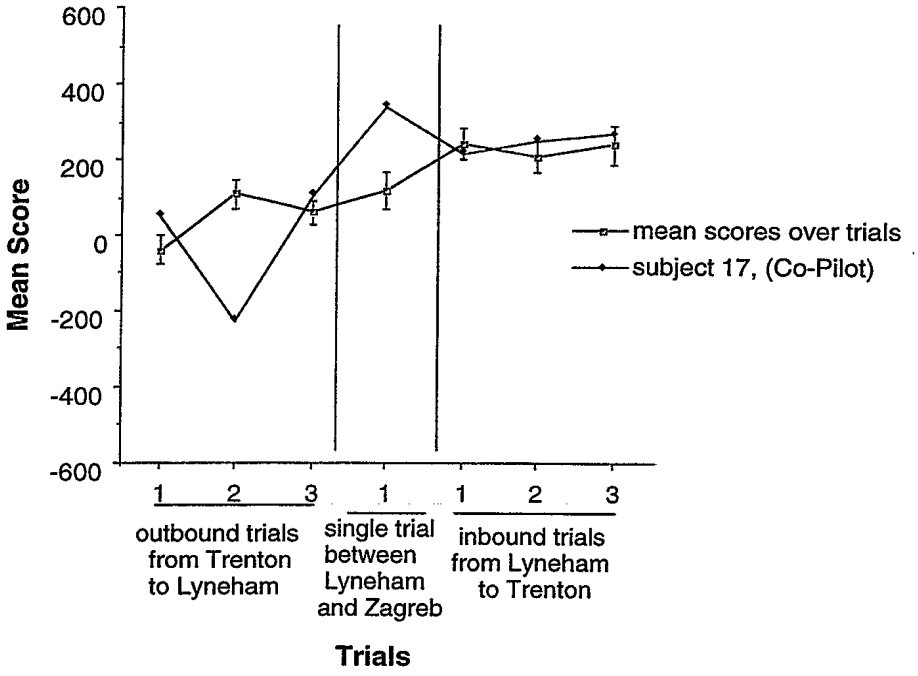

Figure 32. *Illustration of multi-task scores for subject 17 (Co-Pilot) super-imposed on the mean group scores (Aircraft commanders & Co-Pilots) over trials.*

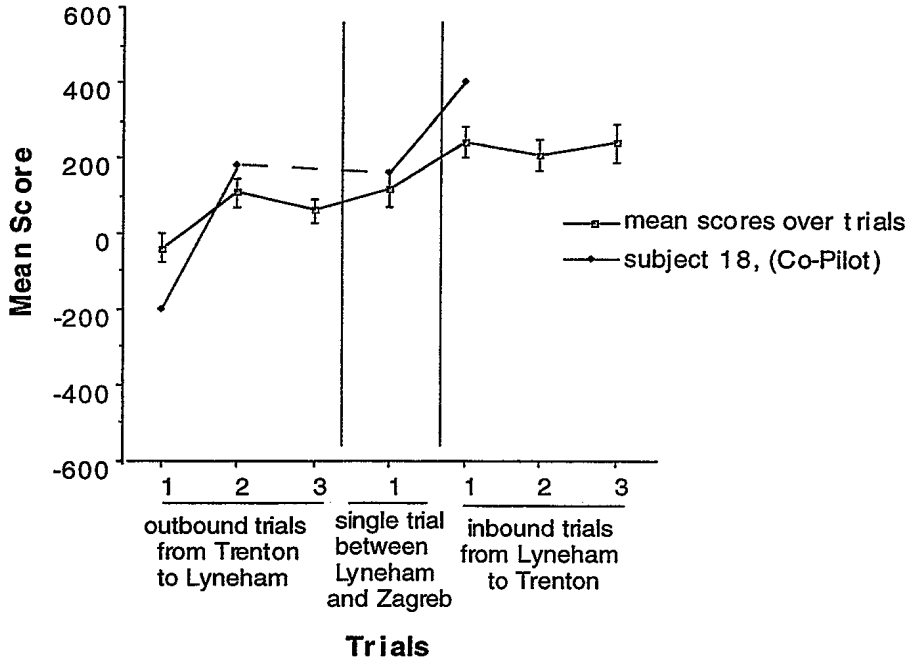

Figure 33. *Illustration of multi-task scores for subject 18 (Co-Pilot) super-imposed on the mean group scores (Aircraft commanders & Co-Pilots) over trials. Dotted line indicates missing Subject 18 data for outbound trial 3. This subjects data is also missing for inbound trials 2 and 3.*

The multi-task data for subject 18 (figure 33) is somewhat spotty in that he was not tested at trial three of the outbound leg nor at trials two or three of the return leg. The limited data from this subject appears to be quite similar to the group averages for the trials he performed.

Thirteen of the eighteen pilots and co-pilots show what is probably a fatigue-induced drop in performance during the transatlantic legs of these missions. Of these "fatigue sensitive" individuals, twelve show a performance decrement in the outbound transatlantic leg, while nine show a performance decrement during the return transatlantic leg, and eight show a fatigue-induced performance decrement on both transatlantic legs.

It might be instructive to recall that the purpose of the study was to document the impact of fatigue (if any) on psychomotor performance and to assess the corresponding implications for flight safety. Given that the scheduling demands on the aircrew undertaking these re-supply missions in support of our troops in Bosnia were minimal in comparison with the huge effort of Operation Alliance (the initial airlift to deliver heavy vehicles and equipment to the theatre of operations), the results of this study can be considered a baseline with respect to the genesis of fatigue on routine transatlantic air transport operations. Should our transport crews be called upon to perform another airlift similar to Operation Alliance with minimum opportunity for crew rest both during and between missions (they flew three missions back-to-back before being "time expired" or removed from the flow of the airlift because they flew the maximum allowable number of hours per month in about two weeks) aircrew fatigue would be much more severe than the significant fatigue we were able to measure on these routine re-supply missions, no doubt with serious implications for flight safety.

CONCLUSIONS

1) Overall, our transport aircrews showed a linear pattern of decreasing sleep over the last three days before embarking on these re-supply missions with the last night at home in Trenton providing the least sleep of the entire mission period. Almost 25% of the participating aircrew received less than six hours sleep during their last night at home, immediately prior to the start of these missions. As indicated earlier, less than 6 hours has been found to produce significant decrements in performance (4). The Loadmasters appeared to get the least amount of sleep of all crew MOCs such that their sleep patterns do not change appreciably during either the pre-mission or mission periods. Once the mission commences, the worst night for sleep hygiene (for all MOCs) during the mission occurs during the second night in England. During this second night in England, 15% of the aircrew received less than 6 hours of sleep.

2) The self-rated scores for alertness, mental, and physical fatigue across all the aircrew MOCs indicate a linear deterioration of alertness and a linear increase in fatigue throughout the long transatlantic flights with the outbound flights to England tending to be slightly more fatiguing than the return flights to Trenton.

3) While there are some decrements in performance of individual crew members for each of the SUSOPS tasks, the logical reasoning task is the only SUSOPS task which clearly suggested the presence of fatigue (for group averaged data) during these missions. Performance on any given psychomotor task is a function of the ability to rally ones attentional resources to accomplish the task in question. In a fatigue study, this is easy to

do when one is fresh (before fatigue develops). Further, the greater the fatigue level and the more difficult the task, the more performance suffers. The SUSOPS batteries were developed to look at performance during sustained operations studies, some of which lasted as long as 64 hours. The sensitivity of the SUSOPS tasks have proven to be more than adequate measures of performance degradation in susained operations research. They were not designed for measuring more subtle effects of fatigue such as for the current study. However, these tasks were included in order to tie into the existing fatigue literature. We would venture to suggest that if the SUSOPS tests were run during a major airlift like Operation Alliance, they would have produced more significant fatigue effects than were evident in this study of routine re-supply missions in support of our troops in Bosnia. Alternatively, if the SUSOPS tasks were given for a longer duration in the present study, fatigue effects might have been more evident. Both of these are empirical issues.

4) The multi-task data shows clear-cut fatigue effects for group averaged data with thirteen of the eighteen pilots and co-pilots we tested on this task, demonstrating varying degrees of performance impairment, ranging from minimal to marked impairment. The multi-task is therefore quite sensitive to the level of fatigue encountered on these re-supply misisons.

RECOMMENDATIONS

1) The next time our transport crews are called upon to undertake a major airlift with a schedule as demanding as that of Operation Alliance, we need to repeat this study in order to document what will no doubt be a much more deleterious effect on performance. With data from both extremes of scheduling demands (airlift vs sustainment flights) we should be in a position to determine to what extent airlift schedules may impact on crew performance.

2) Given that almost 25% of the participating aircrew received less than six hours sleep during their last night at home, immediately prior to the start of these missions, departure for these types of missions should be delayed by an hour or two, if at all possible.

3) Further, given the limited sleep obtained by our crews during their second night in England (no doubt due to having to sleep at an inappropriate circadian time), we should undertake a study of melatonin with a view to determining whether or not melatonin is an effective countermeasure which improves sleep without any performance liability. To this end an experimental protocol has already been approved by the DCIEM Human Ethics Committee. It is anticipated that this study will be completed over the fall of 1998.

Acknowledgements

We are indebted to Dr Jonathon French of the USAF Armstrong Labs in San Antonio for bringing us "up to speed" in this domain of research and for lending us wrist actigraphs to collect our sleep data.

programmed all of our wrist actigraphs and downloaded the sleep data from the actigraphs to our computers.

We are also indebted to the 53 air crew who participated in this study. Sometimes, when fatigued during stressful sections of these missions, the last thing they really wanted to do was undergo yet another session of psychomotor performance testing. Also, many of these individuals were not very happy with the notion of wearing what they called "a box lunch on a wrist" but they did so anyway. "Thank-you all. We could not have done this work without your support."

The data collection efforts of the following people are gratefully acknowledged; Capt L. Donati, Mr. W.D. Fraser, Mr. I. Mack, Ms P. Odell, WO Y. Robichaud, and Sgt L. White.

References

1) Angus, RG, Pigeau RA, Heslegrave R. 1992. "Human performance and sleep research: from the field to the laboratory; in C. Stampi (editor), Why we Nap. pp 217-241. Boston: Birkhauser.

2) Baddeley AD, A 3 min reasoning test based on grammatical transformation. Psychonomic Science. 1968; 10:341-2.

3) Barth JL, Holding JH, and Stamford BA. Risk versus effort in the assessment of motor fatigue. Journal of Motor Behaviour, 1976. 8, 189-194.

4) Belenky G, Penetar DM, Thorne D, et al. The effects of sleep deprivation on performance during continuous combat operations: In: Staff. Food components to enhance performance. Washington, DC: National Academy Press, 1994.

5)Billings CD, Reynard WD. Human factors in aircraft incidents: results of a 7-year study. Aviat Space Environ Med 1984; 55: 960-65.

6) Buley, LE. Experience with a Physiologically-based Formula for Determining Rest Periods on Long-Distance Air Travel. 1970. Aerospace Medicine. 1970;41:680-683.

7) Caldwell JA. Fatigue in the Aviation Environment: An Overview of the Causes and the Effects As Well As Recommended Countermeasures. Aviat Space Environ Med. 1997; 68:932-8.

8) Chidester TR. Trends and Individual Differences in Response to short-haul flight operations. Aviat, Space Environ Med. 1990; 61:132-37.

9) Cook MR, Cohen H, & Orne MT. Recovery from Fatigue. (Tech. Rep No. 55). Fort Detrick MD: 1972. Army Medical Research and Development Command.

10) Coren S. Sleep Thieves; An Eye-opening Exploration Into the Science & Mysteries of Sleep. The Free Press, New York, NY. 1996.

11) French, J. Bisson, RU, Storm WF, Boll PA, Martinez J. Mitchida J, Neville KJ, Vogel E, and Whitmore J. Crew Performance and fatigue during sustained B1-B sorties: 1993; Project Futility Magellan (Report AL/CF-TR-1004-0014). Brooks Air Force Base, TX: Armstrong Laboratory, Crew Systems Division.

12) French J, Hannon P, and Brainard GC. Effects of bright illuminance on body temperature and human performance. 1990; Annual Review of Chronopharmacology, 7, 45-49.

13) Gander PA, Myhre G Graeber RC, Andersen HT, and Lauber K. Adjustment of sleep and the circadian temperature rhythm after flights across nine time zones. Aviat, Space Envviron Med. 1989; 60: 733-42.

14) Gander PH, DeNguyer BD, Rosekind MR, and Connel LJ. Age, circadian rhythms, and sleep loss in flight crews. Aviat, Space Environ Med. 1993; 64: 189-95.

15) Gerathewohl, SJ. Simple Calculator for Determining the Physiological Rest Period after Jet Flights Involving Time Zone Shifts. Aerospace Medicine. 1974;45:449-450.

16) Heron RM, Weinberg H, Robertson A, Jantzen KJ. Development of a multitasking device for the measurement of air traffic controller skills. 1995. Proceedings of Eighth International Symposium on Aviation Psychology, Columbus, OH: Ohio State Press.

17) Hoddes E, Zarcone V, Smythe H, Phillips R, and Dement WC. Quantification of Sleepiness: A new approach. Psychopohysiology, 1973, 10: 431-436.

18) McCauley SF. Aircrew fatigue countermeasures. In: AGARD Conference Proceedings, 495. Symposium on Progress in Military Airlift, 1990 May 23-31; Lisbon, Portugal. Neuilly sur Seine, France: Advisory Group for Aerospace Research and Development 1991.

19) Mohler, SR. Physiological Index as an Aid in Developing Airline Pilot Scheduling Patterns. Aviat. Space Environ. Med. 1976;47:238-247.

20) Mortimer RG. Illness, drugs, fatigue, and stress in the cockpit reported by pilots. In: R.S. Jensen and D. Neumeister (Eds), Proceedings of the Seventh International Symposium on Aviation Psychology (pp. 650-653). Colombus: Ohio State University, Dept. of Aviation.

21) Perelli LP. Fatigue Stressors in simulated long-duration flight: Effects on performance, information processing, subjective fatigue, and physiological cost (Report SAM-TR-80-49). Brooks Air Force Base, TX: School of Aerospace Medicine, 1980.

22) Pigeau R, Naitoh P, Buguet A, McCann C, Baranksi J, Taylor M, Thompson M, and Mack,I. (1995). Modafinil, d-amphetamine, and placebo during 64 hours of sustained mental work: Effects of mood, fatigue, cognitive performance, and body temperature. Journal of Sleep Research, 4(4), 212-228.

23) Ritter RD. "And we were tired" fatigue and aircrew errors. IEEE AES Systems Magazine; March 21-26, 1993.

24) Samel A, Wegman HM, Summa W, and Naumann M. Sleep Patterns in Aircrew Operating on the Polar Route Between Germany and East Asia. Aviat, Space Environ Med. 1991; 62: 661-669.

25) Shingledecker CA, and Holding DH. Risk and effort measures of fatigue. Journal of Motor Behaviour, 1974; 6, 17-25.

26) HM and Klein KE. Jet-lag and aircrew scheduling. In: S. Folkard and T.H. Monk (eds.), Hours of work. New York: Wiley, 1985.

27) Weinberg H. The assessment and classification of complex information-processing in multitasking environments. International congress of Ergonomics, Toronto, August 1994.

28) Weinberg H, Cheyne D, Jantzen KJ. The assessment of information processing in complex multitasking environments. 1996. AGARD NATO, Bonn, Germany.

DOCUMENT CONTROL DATA

(Security classification of title, body of abstract and indexing annotation must be entered when the overall document is classified)

1. ORIGINATOR (the name and address of the organization preparing the document. Organizations for whom the document was prepared, e.g., Establishment sponsoring a contractor's report, or tasking agency, are entered in section 12.)
DCIEM
PO Box 2000 North York, ON
CANADA

2. DOCUMENT SECURITY CLASSIFICATION (overall security classification of the document including special warning terms if applicable)
Unclassified

3. DOCUMENT TITLE (the complete document title as indicated on the title page. Its classification should be indicated be the appropriate abbreviation (S,C,R or U) in parentheses after the title.)
Human factors of CC-130 operations. Volume 6: Fatigue in long-haul re-supply missions

4. DESCRIPTIVE NOTES (the category of the document, e.g., technical report, technical note or memorandum. If appropriate, enter the type of report, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.)
Final Report

5. AUTHOR(S) (Last name, first name, middle initial. If military, show rank, e.g. Burns, Maj. Frank E.)
M.A. Paul, R.A. Pigeau, and H. Weinberg

6. DOCUMENT DATE (month and year of publication of document)	7.a. NO. OF PAGES (total containing information. Include Annexes, Appendices, etc.)	7.b. NO. OF REFS. (total cited in document)
February 1998	43	27

8.a. PROJECT OR GRANT NO. (if appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant)	8.b. CONTRACT NO. (if appropriate, the applicable number under which the document was written)
Project 3fa11	
9.a. ORIGINATOR'S DOCUMENT NUMBER (the official document number by which the document is identified by the originating activity. This number must be unique to this document.)	**9.b. OTHER DOCUMENT NO.(S) (any other numbers which may be assigned this document either by the originator or by the sponsor.)**
DCIEM RR-98-R-19	

10. DOCUMENT AVAILABILITY (any limitation on further dissemination of the document, other than those imposed by security classification)

X	Unlimited distribution
	Distribution limited to defence departments and defence contractors; further distribution only as approved
	Distribution limited to defence departments and Canadian defence contractors; further distribution only as approved
	Distribution limited to government departments and agencies; further distribution only as approved
	Distribution limited to defence departments; further distribution only as approved
	Other

11. ANNOUNCEMENT AVAILABILITY (any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availability (10.) However, where further distribution (beyond the audience specified in 10) is possible, a wider announcement audience may be selected.)
Unlimited

12. SPONSORING ACTIVITY (the name of the department project office or laboratory sponsoring the research and development. Include the address.)
DCIEM
PO Box 2000, North York, ON
CANADA

DSIS DCD03
HFD 09/94

13. ABSTRACT (a brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), (R), or (U). It is not necessary to include here abstracts in both official languages unless the text is bilingual).

Intoduction. Deployment of troops in foreign theatres requires a massive airlift capability. Transport squadrons are called upon to deliver men and materiel, day and night, around the clock, usually during long transmeridian flight. The relentless fatigue encountered in such operations can be severe enough to pose a flight safety hazard. During Operation Alliance, a month long airlift (January 1996) in support of Canadian tropps in Bosnia, 18 Air Transport Group CC-130 Hercules carried out 86 missions from Trenton to Split with aircraft landing in theatre every 4 hours. Most crews attained the 120 hour maximum allowable flying time per 30 day period, in as little as 2 weeks. There were several reports of aircrew falling asleep at the controls. After the main airlift was completed, only three of these missions were flown each week. This frequency of re-supply missions was adequate to sustain Canadian troops in former Yugoslavia. Because these sustainment flights were flown in a more relaxed manner, aircrew were given 32 hours on the ground in the United Kingdom on airrival from Trenton, Canada (5 time zones), before proceeding to Zagreb, Croatia (1 time zone), instead of the 14 hours they were given at this stage of the mission, during the original airlift. The current study is an attempt to document to what extent fatigue (and time zone changes of 5 and 6 hours) can impact of aircrew performance during routine re-supply missions. Any fatigue-related lapses in performance in this study, will only worsen when crews are asked to fly repeated missions back-to-back with minimum crew rest between mission legs as was done during operation alliance. **Methods.** Ten routine re-supply missions from Trenton to Zagreb were studied and involved 53 aircrew subjects. In order to document their sleep hygiene, the aircrew were asked to were wrist actigraphs from approximately 5 days prior to the mission, until the mission was completed. Aircrew psychomotor performance was tested during the actual flights on two different computer-based psychomotor batteries. One test battery consisted of a questionnaire designed to elicit subjective ratings of alertness, and fatigue, as well as three separate serial tests from the DCIEM SUSOPS battery (serial reaction time, logical reasoning, and serial subtraction). The other battery was a recently developed multi-task (analogous to a flying task) and incorporated four sub-tasks which had to be undertaken simultaneously. There were three psychomotor test sessions during the outbound transatlantic leg (Trenton to Lyneham, UK), one test session on the Lyneham-Zagreb-Lyneham leg, and three test sessions on the return transatlantic leg from Lyneham to Trenton. **Results.** The amount of sleep the aircrews experienced during the days leading up to a mission steadily decreased from an average of 475 minutes per day to an average of 380 minutes per day, with 25% of the subjects getting less than 6 hours and some subjects obtaining as little as 250 minutes per day on the last pre-mission night at home, prior to starting the mission. During the missions, the worst night of sleep occurred during the second night in the UK, no doubt because of having to sleep at inappropriate circadian times. Subjective ratings of alertness and fatigue became progressively worse during both transatlantic legs. Of the three SUSOPS psychomotor tests, only the logical reasoning task indicated any fatigue related lapses in performance during these missions, and only on the outbound leg. The multi-task data illustrates unequivocal fatigued-related lapses in performance on both transatlantic legs, with 13 of the 18 pilots tested, showing performance ranging from minimum to marked impairment. **Recommendations.** Repeat this study during a major airlift like Operation Alliance. 2) Given that the worst night for sleep occurred at home, the last night before the mission, delay departure by one to two hours if at all possible. 3) Determine whether or not melatonin can safely facilitate sleep at an inappropriate circadian time.

Future work (Sept 98 - Nov 98) will involve a laboratory-base study to determine whether or not melatonin can facilitate crew rest at an inappropriate circadian time. Should we find melatonin to be efficacious in this regard, we would be prepared to conduct melatonin trials during real overseas operations. While the implementation phase builds logically on the current study, this work is beyond the scope of the original tasking and will be reported separately and at a later date.

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible, keywords should be selected from a published thesaurus, e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

aviation safety
fatigue
long haul transport missions
transmeridian

#588776