Image Cover Sheet

CLASSIFICATION	SYSTEM NUMBER 516720
UNCLASSIFIED	

TITLE

CC130 pilot fatigue during re-supply missions to former Yugoslavia

System Number:

Patron Number:

Requester:

Notes:

DSIS Use only:

Deliver to: DK

This page is left blank

This page is left blank

VOLUME 72 NUMBER 11

NOVEMBER 2001

AVIATION, SPACE, AND ENVIRONMENTAL MEDICINE
DCIEM SL 1999-130

ORIGINAL RESEARCH

CC130 Pilot Fatigue during Re-supply Missions to Former Yugoslavia

M. A. PAUL, B.Sc, M.Sc., R. A. PIGEAU, M.A., PH.D., AND H. WEINBERG, M.A., PH.D.

PAUL MA, PIGEAU RA, WEINBERG H. *CC130 pilot fatigue during re-supply missions to former Yugoslavia.* **Aviat Space Environ Med 2001; 72:965–73.**

Purpose: Deployment of troops in foreign theaters requires a massive airlift capability. The fatigue encountered in such operations can be severe enough to pose a flight safety hazard. The current study documents sleep and the effect of fatigue on aircrew performance during re-supply missions in support of Canadian troops in Bosnia in 1996. ***Methods:*** Ten routine re-supply missions from Trenton, Canada, to Zagreb, Croatia, were studied and involved 9 pilots and 9 co-pilots. To document their sleep hygiene, all pilots wore wrist actigraphs from approximately 5 d prior to the mission, until the mission was completed. Psychomotor performance was tested during the actual flights. Three psychomotor trials during the outbound transatlantic leg (Trenton to Lyneham, UK) were employed, one trial on the Lyneham-Zagreb-Lyneham leg, and three trials on the return transatlantic leg from Lyneham to Trenton ***Results:*** The amount of daily sleep during the 3-d period prior to the mission steadily decreased from an average of 8 h 40 min per day to 6 h 30 min ($p < 0.001$). During the missions, the worst night of sleep occurred during the second night overseas. During both transatlantic legs, there were significant decrements in the subjective ratings of alertness ($p < 0.001$), and increases in physical ($p < 0.001$) and mental fatigue ($p < 0.001$). Performance on the logical reasoning task as well as the multitask showed probable fatigue effects during the outbound leg of the missions. ***Conclusions:*** Our transport pilots showed a pattern of progressively decreasing sleep. Self-rated scores for alertness, mental and physical fatigue, indicate a deterioration of alertness, and an increase in fatigue throughout the long transatlantic flights.
Keywords: air transport operations, fatigue, sleep hygiene, psychomotor performance.

FLYING REQUIRES a high degree of psychomotor skill, but stressors such as noise and vibration, long crew days, irregular work schedules, circadian disruptions, and inadequate sleep can interact to produce dangerous levels of pilot fatigue (9).

Deployment of troops in foreign theaters requires a massive airlift capability. Transport squadrons are called on 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 often push the upper limits of the maximum allowable 16-h. At the end of that crew day, crews get a minimum of about 14 h rest (during which they are often unable to sleep soundly because of circadian rhythm disruption) and then are required to repeat the process.

For example, in 1982, after Argentina invaded the Falkland Islands, a British task force was assembled to regain sovereignty of those Islands (Operation Corporate). While primarily a naval and army operation, the Royal Air Force (RAF) provided an air-bridge in support of this Operation. This air-bridge involved Hercules transports specially equipped with air-to-air refuelling (AAR) booms, supported AAR by VC10s operating between the UK and Ascension Island (about halfway between the UK and the Falkland Islands) and Victor K2s which refuelled the Hercules transports between Ascension Island and the Falkland Islands. Baird and Nicholson (4) present a graph showing the limit of flying hours over days calculated to indicate the maximum flying workload compatible with maintenance of an acceptable sleep pattern. Because this 'workload' curve was far exceeded, hypnotics were used to facilitate sleep for the aircrews during Operation Corporate. That there were no incidents or accidents attributable to human error is a solid testimonial to the management of Operation Corporate and to the calibre of the RAF aircrews.

In another example, Operation Alliance, a month-long Canadian Forces (CF) 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, Canada, to Split, Croatia with aircraft landing in theater every 4 h. During Operation Alliance, most crews attained the 120 h maximum allowable flying time per 30 d period, in as little as 10 to 12 d. There have been anecdotal comments from the aircrew, as to the severity of the fatigue levels they experienced while at the controls during these missions.

From the Defence and Civil Institute of Environmental Medicine, North York, Ontario, Canada
This manuscript was received for review in January 2000. It was revised in December 2000 and May 2001. It was accepted for publication in July 2001.
Address reprint requests to: Michel A. Paul, who is a Defence Scientist at Defence and Civil Institute of Environmental Medicine, 1133 Sheppard Avenue West, P.O. Box 2000, North York, Ontario, Canada M3 M 3B9, mpaul@dciem.dnd.ca.

TABLE I. DEMOGRAPHICS OF PARTICIPATING SUBJECTS.

	Aircraft Commanders	Co-Pilots
#/Gender	9 Males	8 Males/1 Female
Mean age (yr)	35.5	29.6
Age Range (yr)	29 to 42	24 to 32
Rank Range	Capt. to Lt. Col.	Lt. to Capt.

There are essential differences between Operation Corporate and Operation Alliance. During Operation Corporate the RAF crews flew about 7,000 mi across 4 times zones logging an average of 150 h in 24 d. During Operation Alliance the CF crews flew about 4,500 mi across 6 time zones logging an average of 120 h in as little as 10 to 12 d.

As desirable as it would have been to study Operation Alliance by monitoring aircrew performance and sleep hygiene throughout these missions, the airlift was completed before this could happen. However, there was the opportunity to study a follow-on operation.

After the main airlift was completed, three missions were flown each week to sustain Canadian troops in former Yugoslavia. We monitored one of these missions each week. Because these sustainment flights were flown in a less urgent manner, aircrew were given 32 h on the ground in the United Kingdom on arrival from Trenton, Canada (five time zones relative to home base), before proceeding to Zagreb, Croatia (one time zone relative to the United Kingdom). This is in contrast to only14 h that were given at this stage of the mission, during the original airlift.

The current study, undertaken at the request of the Canadian Forces Air Command, was an attempt to document the extent to which fatigue (and time zone changes of 5 and 6 h) impact on aircrew performance during routine re-supply missions to former Yugoslavia. If fatigue-related lapses occur under these circumstances, it is probable that fatigue was much worse when crews were flying repeated missions, back-to-back, with minimum crew rest, both within and between missions, as occurred during Operation Alliance. Unlike Operation Alliance, each crew flew only a single mission. Psychomotor performance and sleep hygiene were monitored during 10 missions.

METHODS

Subjects

The demographics for the 18 Canadian Forces (CF) subjects (9 pilots and 9 co-pilots) who participated in the study are illustrated in **Table I.**

One limitation of this study was the difficulty of gaining access to the CC-130 aircrews for several days prior to the missions in order to allow sufficient time for training to a level of stable performance on the psychomotor tasks.

Equipment and Data Collection Strategy

All pilots wore wrist actigraphs (Precision Control Design, Fort Walton Beach, FL) from approximately 5 d prior to the mission until the mission was completed. This allowed periods of sleep and wakefulness to be obtained for each pilot during the missions as well as during the period immediately prior to the missions. Wrist actigraphs detect movements every 0.1 s, and the data can be integrated in user-selectable epochs (1-min epochs in this study). The data are then analyzed using software (Win Act, version 1.2, developed by Tim Elsmore, from Activity Research Services of San Diego, CA) that accumulated the frequency of such movements over time in order to quantify the number and duration of sleep episodes.

The following psychomotor test battery and a multitask were administered to the pilots from a single laptop computer: a subset of 3 tests [SRT (serial choice reaction time), LRT (logical reasoning task), and SST (serial subtraction task)] from the DCIEM SUSOPS (sustained operations) battery (2,3) were performed, successively, over 10 min; and a multitask taking 15 min per data collection iteration (27,29).

The pilots 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. For approximate test times and locations, see **Fig. 1**.

Task Descriptions

Psychomotor test batteries, such as the DCIEM sustained operations (SUSOPS) test battery have a long history of documenting laboratory-based findings concerning the effect of fatigue on performance (2,22). 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 multitasking test (27,28) believed to possess

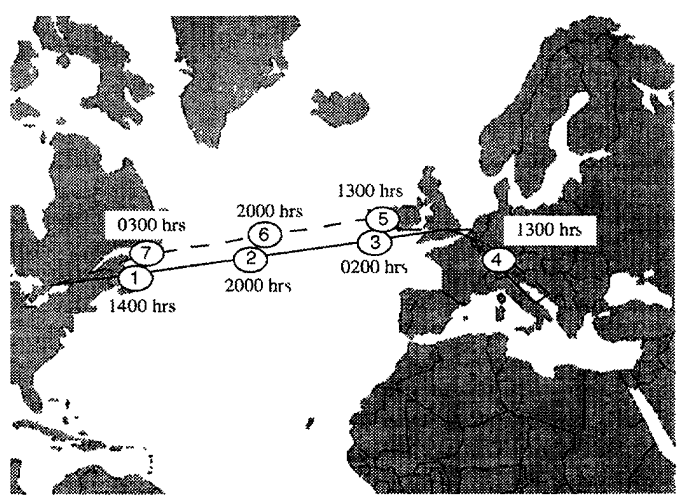

Fig. 1. Map illustrates approximate outbound (solid lines) and return (dashed lines) tracks, including approximate psychomotor performance testing locations and approximate testing times (GMT). Locations 1, 2, and 3 correspond to testing on the outbound leg. Location 4 averaged about 3 degrees East Longitude, and testing occurred 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.

greater context validity to flying, as well as requiring higher levels of cognitive activity to complete.

SUSOPS Tasks

All of the SUSOPS tasks were performed on laptop computers (Pentium 133 Mhz with 16 Mb RAM and 1 Mb of video memory) using an external mouse as the input device. All pilots performed the tasks in the rear of the fuselage, in order to avoid distractions by cockpit instrumentation or interference with the aircraft controls, and also to perform the task in relative privacy. The pilots were seated in the standard collapsible webseats 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). Pilots 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. They were instructed to work as quickly as possible without sacrificing accuracy. They wore headsets to defend against the high ambient noise, but in order to avoid distractions by routine communications between crewmembers, the headsets were not plugged into the aircraft intercom while they were performing the psychomotor tests.

Computer Questionnaire

The first item in each test session involved making the following three subjective assessments: 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); and 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).

The aircraft position (latitude and longitude) was obtained from the cockpit crew by the experimenter, in order to track where each testing session took place.

Serial Reaction Time (SRT)

This task required the subjects to select which of four letters (A, B, C, and D, 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). Immediately after each response, the stimulus letter was replaced by single serial random presentations of any of the four stimulus letters.

Serial Subtraction Task (SST)

This task is similar to a task described by Cook, Cohen, and Orne (10). 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.

Logical Reasoning Task (LRT)

This task was developed by Baddeley (3) and described by him as a measure of higher mental processes, and is based on grammatical transformation. The task involved understanding sentences of varying syntactic complexity. It consisted 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.

Multitask

This program was run on the same computer and "lapboard" as the above test battery. The multitask development resulted from approximately 8 mo of study and the observation of flight simulators, including the Airbus 310, the Boeing 747, and the Citation. The screen showed four separate displays presented in four quadrants of the screen representing four sub-tasks (**Fig. 2**) which were performed simultaneously. Three of these four tasks interacted. These tasks simulated flying an aircraft to specific targets or "waypoints" via a tracking task. There were vigilance sub-tasks with altitude assignment changes visible for only 5 s and the pilots also had to be vigilant in order to determine when the two "attitude indicators" disagreed with each other and then determine which of the attitude indicators accurately reflected the "aircraft attitude." A bar task (analagous to managing the power quadrant of a large multi-engine transport) did not interact with the other three sub-tasks. The measures of performance include scores related to error detection and selective attention, visuo-motor tracking and co-ordination, short-term memory, mental arithmetic, and scanning strategies. The raw output data file was merged with a computer reduction algorithm to yield a single final weighted composite score that reconciles correct responses and errors. This task is explained in more detail elsewhere (29).

Experimental Design Considerations

Although it was anticipated that the pilots on these missions would experience fatigue due to the length of the missions, it is also possible that they experienced some fatigue due to the circadian stress inherent in such long transmeridian flights. Fortunately, all missions departed from Trenton at approximately 0800 h local time, minimizing any confounding of results due to different circadian rhythm stresses across crews.

Because of the requirement to collect pilot performance data without interfering with their flying duties,

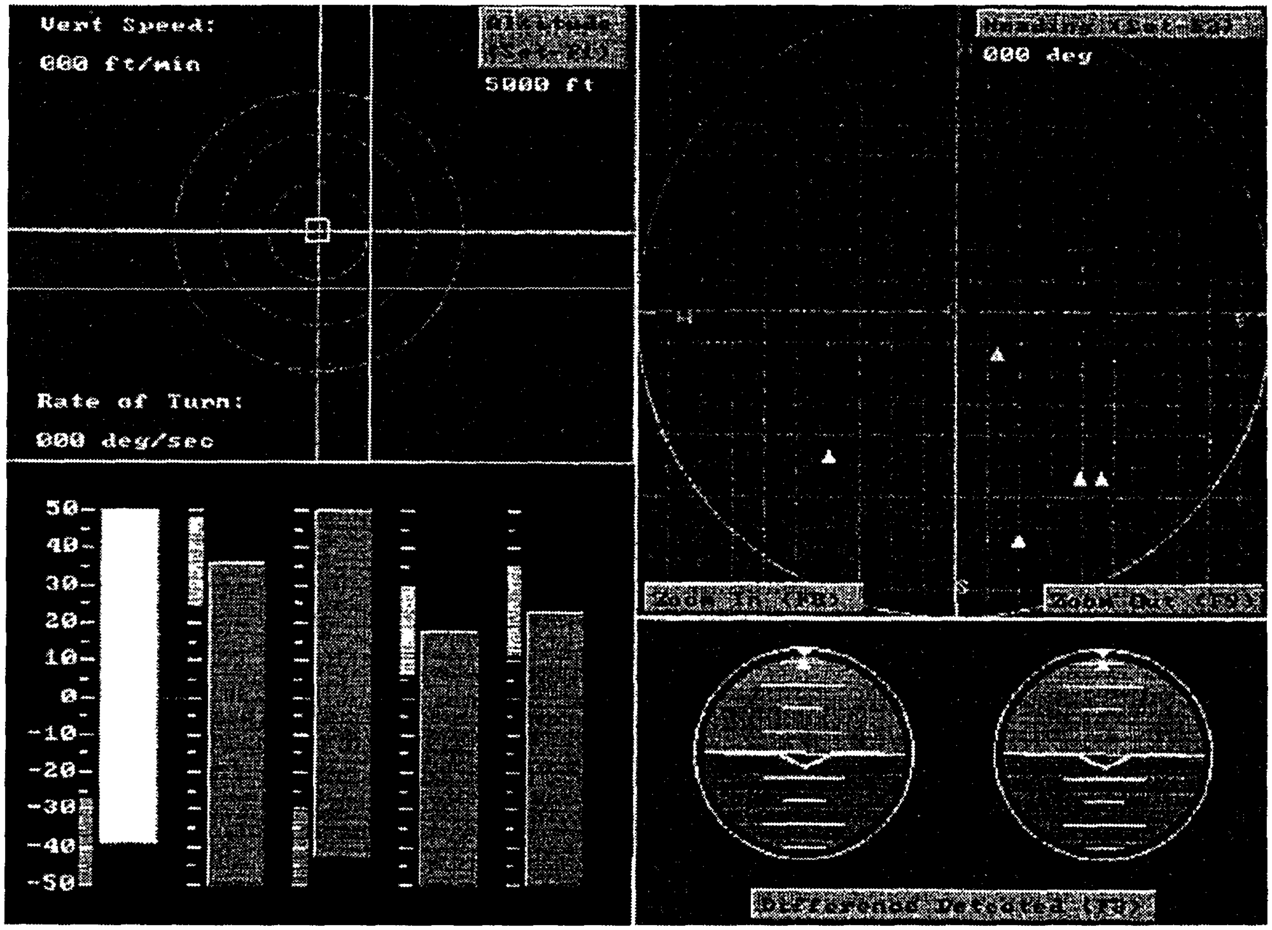

Fig. 2. Computer screen "capture" of multi-task display in gray scale instead of colors which are present on the computer.

we could not rigorously define the time-intervals over which we could collect data. However, pilots and navigators from HQ identified the following data collection opportunities.

1) Outbound leg—Trenton, Canada, to Lyneham, United Kingdom: On reaching top of climb (TOC) after departure from Trenton (approximately 1400 h GMT), the first testing cycle began with the concurrence of the aircraft commander (AC) and, one at a time, each pilot rotated through the test site, located in the after-end of the fuselage. Approximately half way across the Atlantic (about 6 h elapsed time after take-off, or approximately 2000 h GMT), both pilots were asked to perform the second iteration of the test batteries. Approximately 2 h out from Lyneham (approximately 0200 h GMT), both pilots were asked to perform a third iteration. All data collection for this leg was completed by the time the aircraft was about 1 h out from Lyneham, in order not to interfere with the relatively busy period of approach to landing in Lyneham through congested European airspace.

2) Lyneham to Zagreb, Croatia, and return: This leg was flown in very congested airspace, so each pilot was tested only once for 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 (approximately 1300 h GMT).

3) Inbound (return) leg—Lyneham to Trenton: 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, since Canadian airspace was not as congested as that over Europe. Further, the average flying time on this return leg was 2 h longer than for the outbound leg, because of normal westerly headwinds.

Statistical Analysis

Sleep data from the wrist actigraphs were submitted to one-way repeated measures analysis of variance with "days" as the within factor in order to test any differences in daily sleep quantity over the monitored period. The subjective questionnaire data were submitted to two-way repeated measures analysis of variance with "trials" and "legs" (outbound leg from Trenton to Lyneham and inbound leg from Lyneham to Trenton) as the within factors. This ANOVA design would allow us to determine whether fatigue increased over the long transatlantic legs and whether there was a difference in fatigue between the outbound and return legs. The psychomotor data for the three trials of each of the departure and return legs were submitted to the same repeated measures analysis. We used the Least Significant Difference

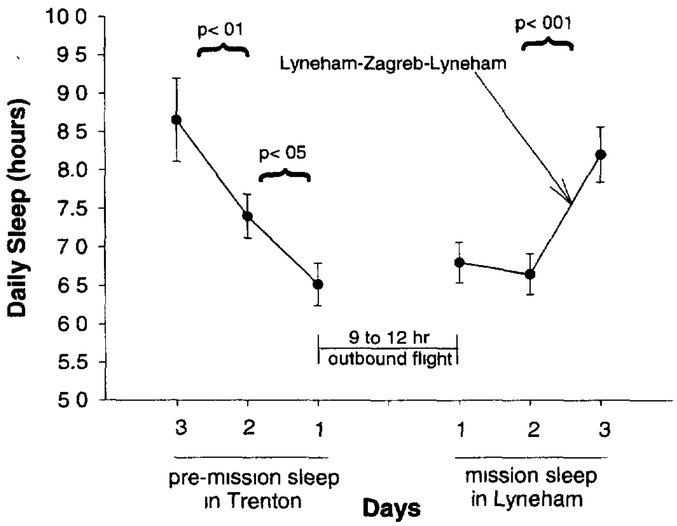

Fig. 3. Plot of Daily Sleep Hours (mean values ± s.e.m.) over pre-mission and mission days collapsed over pilots and co-pilots.

Test to analyze any simple main effects or interactions with the significance level set at 0.05.

RESULTS

Wrist Actigraphy

The number of minutes spent asleep during the pre-mission and mission "days" are illustrated in **Fig. 3**. There was a significant main effect of days $F(5,70) = 7.85$, $p < 0.001$ with the average amount of sleep the subjects experienced during the final days leading up to a mission steadily decreasing from 8 h and 40 min to 6 h and 30 min per day. This result is not surprising since the planned take-off time was 0800 h local time and the crews reported to operations 2 h before take-off.

Total daily sleep time did not recover (to the levels of 3 and 2 d before the mission) until the third night overseas. These results can be explained by the fact that on arrival in England, after a long crew day and a sleep deficit, the crews were very fatigued and were able to sleep relatively easily (24). However, the crews deliberately limited their sleep opportunity during the first sleep in England. Having gone to bed around 0500 h to 0600 h (local time in England), they awakened around noon in an effort to adapt to local time and thus attempt to limit the difficulty in obtaining sleep later that night. In spite of this effort, after obtaining some restorative sleep on arrival, by the second night in Lyneham, the disruption of circadian rhythm tended to make sleep more difficult (15,24). When the crews returned from Zagreb for their last night in Lyneham, their total sleep time significantly increased and returned to the levels evident two and three nights before the mission.

SUSOPS Questionnaire Data

For the Stanford Sleepiness Scale, the main effect of trials was significant $F(2,34) = 37.34$, $p < 0.001$ and is illustrated in **Fig. 4a**. This indicates that with respect to the pilots' self- rating (Stanford Sleepiness Scale), sleepiness increases progressively during both transatlantic legs.

For both "mental fatigue rating" $F(2,34) = 12.99$, $p < 0.001$ and "physical fatigue rating" $F(2,34) = 10.33$, $p < 0.001$, the main effects of trials were also significant and are plotted in **Fig. 4b and 4c**.

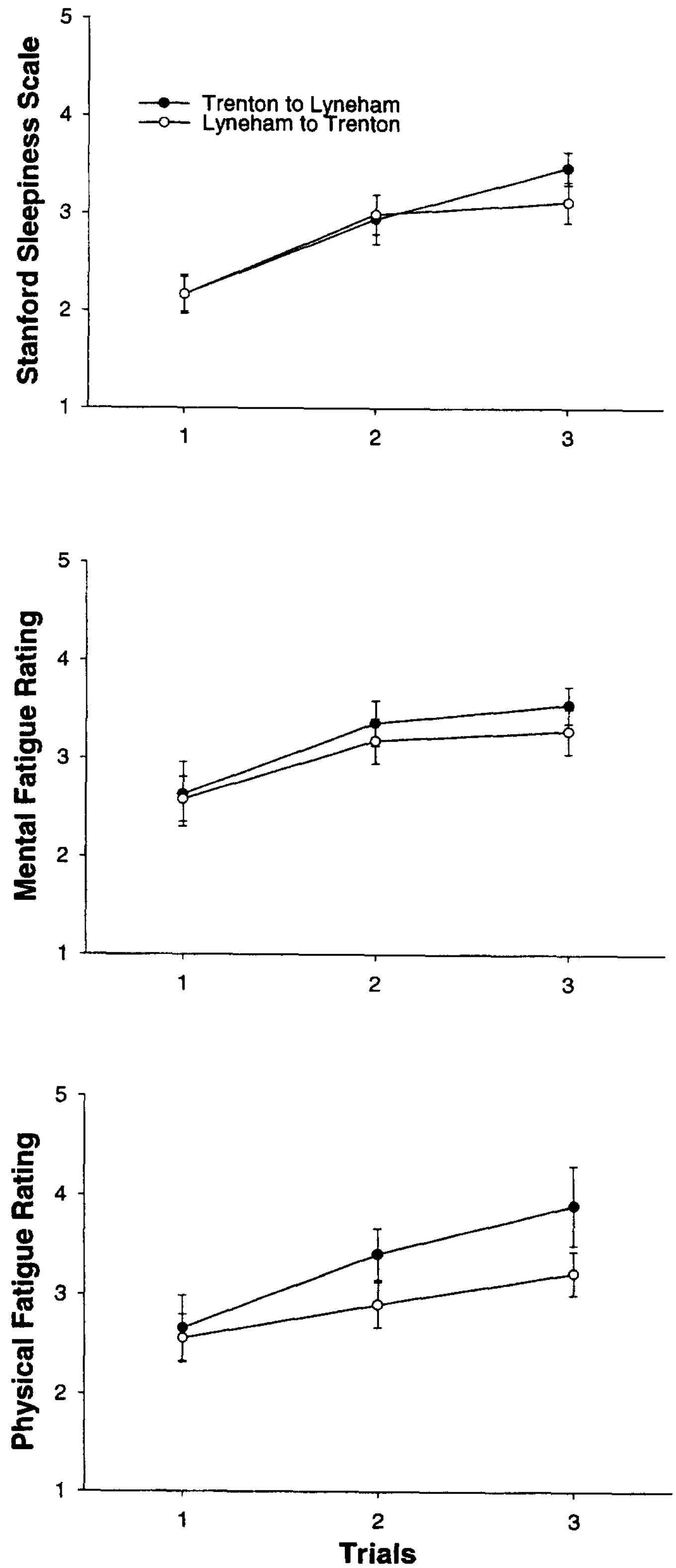

Fig. 4. Subjective assessments of sleepiness and fatigue (mean values ± s.e m.) plotted over trials for departure (Trenton to Lyneham) and return (Lyneham to Trenton) legs.

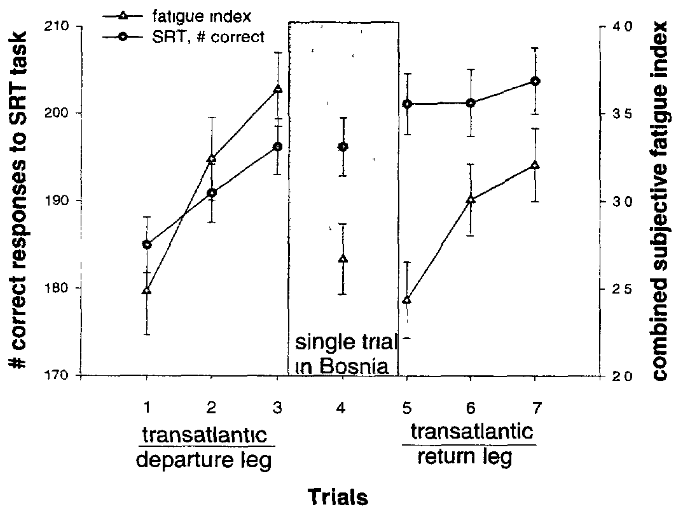

Fig. 5. Number of correct responses to Serial Reaction Time task plotted over trials for departure (Trenton to Lyneham), in-theatre (Lyneham-Zagreb-Lyneham), and return (Lyneham to Trenton) legs, and super-imposed over combined subjective fatigue index (all values are mean ± s.e.m.).

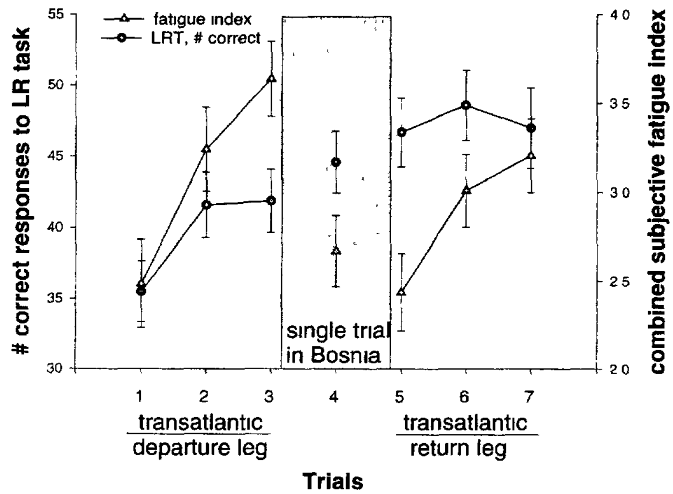

Fig. 7. Number of correct responses to Logical Reasoning task plotted over trials for departure (Trenton to Lyneham), in-theatre (Lyneham-Zagreb-Lyneham), and return (Lyneham to Trenton) legs, and superimposed on combined subjective fatigue index (all values are mean ± s.e.m.).

Self-ratings of mental and physical fatigue tend to change during the transatlantic legs in a manner similar to the self-ratings for sleepiness, in that subjective estimates of both mental and physical fatigue progressively increase over each transatlantic leg.

Psychomotor Data

Psychomotor performance data (SUSOPS and Multitask) were collected only during the flights themselves. The number of correct responses to each of the serial reaction time (SRT), serial subtraction (SS) tasks, logical reasoning (LR), and the multitask scores are superimposed under a combined subjective fatigue index and plotted in **Fig. 5, 6, 7,** and **8**, respectively. A combined subjective fatigue index was generated by averaging the subjective fatigue data (from the Stanford Sleepiness Scale, as well as the mental and physical fatigue scales) shown in Fig. 4. By superimposing an estimate of subjective fatigue onto the plots for cognitive tasks, it is possible to judge the relative effect on performance. The three psychomotor trials of the outbound leg (from Trenton to Lyneham) as well as those of the return leg (from Lyneham to Trenton) were performed serially approximately every 3 h. The interval between trial 3 of the outbound leg and trial 4 (the in-theater trial) was approximately 36 h while the interval between trials 4 and 5 (the first transatlantic trial of the return leg) was approximately 24 h.

Serial Reaction Time

The SRT performance (Fig. 5) shows a linear increase in performance during the three trials of the transatlantic departure leg, in spite of dramatically increasing levels of fatigue on that leg. Obviously, the subjects

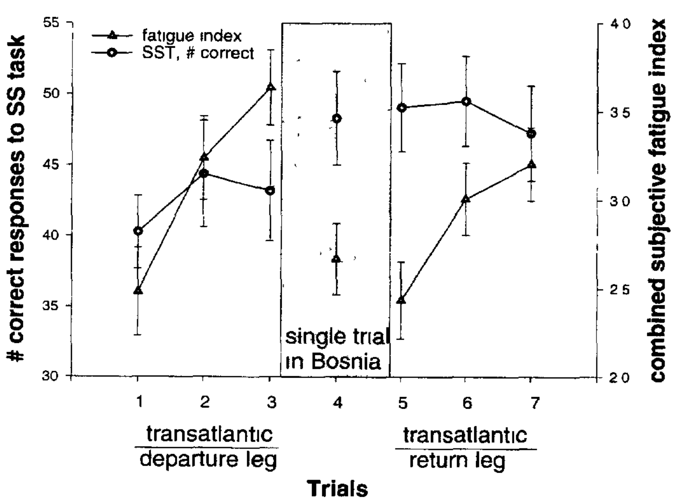

Fig. 6. Number of correct responses to Serial Subtraction task plotted over trials for departure (Trenton to Lyneham), in-theatre (Lyneham-Zagreb-Lyneham), and return (Lyneham to Trenton) legs, and superimposed on combined subjective fatigue index (all values are mean ± s.e.m.).

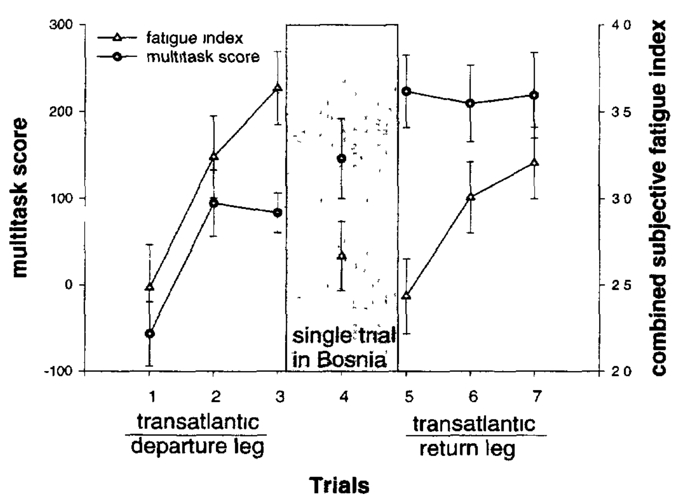

Fig. 8. Multitask score plotted over trials for departure (Trenton to Lyneham), in-theatre (Lyneham-Zagreb-Lyneham), and return (Lyneham to Trenton) legs, and super-imposed over combined subjective fatigue index (all values are mean ± s.e.m.).

continued to improve on this task despite rising levels of subjective fatigue. During trial 4 in Zagreb, 36 h after arrival in Lyneham, subjective fatigue levels are significantly reduced and SRT performance is similar to that of trial 3. At the beginning of the return transatlantic leg (24 h later), after finally attaining a full night of sleep (Fig. 3), fatigue levels are lowest of the entire mission period. While fatigue again increases during this long transatlantic leg, the corresponding SRT performance appears to plateau suggesting a learning asymptote has been attained.

These results suggests that while the subjects are still learning strategies to improve their performance, the serial reaction time task is not sensitive to the level of fatigue encountered on these missions. Furthermore, even though the subjects reported subjective levels of fatigue on the return leg, they may have been able to rally their cognitive resources to compensate, thus showing no change in performance.

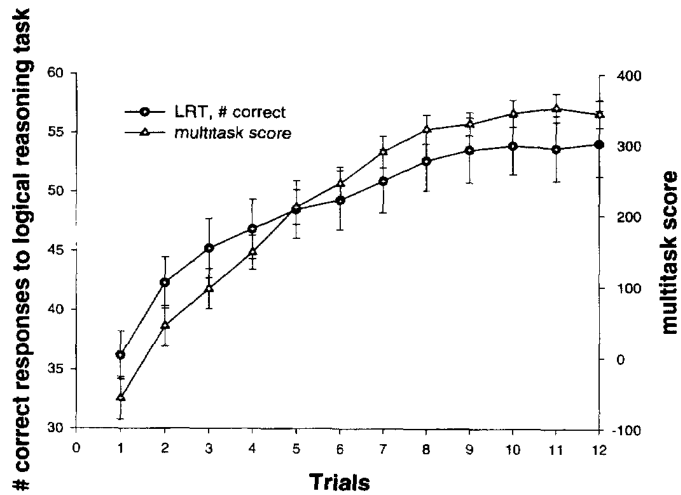

Fig. 9. Normal learning curves for each of logical reasoning and multitask tests, (all values are mean ± s.e.m.)

Serial Subtraction Task

Throughout the mission, the performance curve for the SST is very similar to that of the LRT. Although this curve suggests a fatigue-induced impact on learning at trial 3 and that a learning asymptote has been reached during the return transatlantic leg, high variability in performance has precluded any significant effects for the SST over either transatlantic leg.

Logical Reasoning Task

Unlike SRT during the transatlantic departure leg, LRT learning is not linear (Fig. 7). While the results reveal an improvement in performance between trial 1 and trial 2 ($p < 0.002$), no such improvement in performance occurs between trial 2 and trial 3 ($p < 0.819$). Performance on the LRT continues to improve later on in the mission (see trials 5 and 6). This suggests that the high levels of subjective fatigue experienced during trial 3 of the departure leg may have disrupted learning of the LRT task. In Zagreb, 36 h later, when subjective fatigue is much less, performance continues to improve. Then 24 h after leaving Zagreb, at trial 5 performance improves again eventually levelling off, suggesting either a training asymptote has been reached or another fatigue-induced impact on logical reasoning performance.

Multitask

The performance curve for the multitask is similar to that of the LRT in that learning appears to be disrupted during the last trial of the transatlantic departure leg (Fig. 8) when the subjects were experiencing significant subjective fatigue and where the expected improvement in multitask performance between trial 1 and trial 2 is evident ($p < 0.005$) but no further improvement occurs between trial 2 and trial 3 ($p < 0.833$). Similarly, on the return transatlantic leg, performance appears to have reached a learning asymptote.

As mentioned earlier, we were not given the opportunity to train the aircrew subjects on the psychomotor tasks prior to this study. However, in order to evaluate our current results, we are illustrating the normal learning curves for the logical reasoning and multitasks (**Fig. 9**). This data is from our laboratory and involves 45 subjects in training prior to participation in three different studies. Fig. 9 shows no interruption of learning at trial 3 for either the logical reasoning or the multitask learning curves.

When data were looked at individually, 13 of the 18 pilots and co-pilots demonstrated a drop in multitask performance during the transatlantic legs of these missions, apparently due to fatigue. Of these "fatigue sensitive" individuals, 12 showed a performance decrement in the outbound transatlantic leg, while 9 showed a performance decrement during the return transatlantic leg, and 8 showed a fatigue-induced performance decrement on both transatlantic legs.

DISCUSSION

The psychomotor performance results can be interpreted from the perspective of an interruption of a "learning" paradigm where post hoc testing involving the three trials of the outbound leg (for the LRT and multitask performances, illustrated in Figs. 7 and 8, respectively) suggests a possible impact of fatigue on learning. When Figs. 7 and 8 are compared with the normal learning curves, which have no interruptions in learning (Fig. 9), the possibility of a fatigue-induced impact on learning of the LRT and multitask is quite credible. This conclusion is supported by the dramatic and linear rise in subjective levels of fatigue during the long transatlantic legs (Fig. 5, 6, 7, and 8), especially during the outbound leg, which began with a relative sleep deficit (Fig. 3). The impact of fatigue on the LRT and multitask occurs during trial 3 of the outbound transatlantic leg in that learning of this task is interrupted in response to the highest levels of subjective fatigue of the entire mission (Fig. 7). However, after the pilots have rested, performance on these tasks continues to improve.

Prior to the mission there is a progressive deterioration in sleep time during the last three nights of home-based sleep with the least amount of sleep (6 h and 30 min) occurring during the last night before commencement of the mission (Fig. 3). This is in sharp contrast to the 8 h and 12 min of sleep these pilots received during their last sleep in Lyneham before returning to Canada (Fig. 3). This return leg was flown over a similar route as the outbound leg, but due to the restorative sleep obtained during the last night in Lyneham, fatigue is lower during the return leg than during the outbound leg. Consequently psychomotor performance appears to be free of any impact from fatigue during that leg.

Performance on any given psychomotor task is a function of the ability to rally one's attentional resources to accomplish the task in question. Further, the greater the fatigue level and the more difficult the task, the more performance suffers (1). The SUSOPS batteries were developed to look at performance during sustained operations studies, some of which lasted as long as 64 h. The sensitivity of the SUSOPS tasks have proven to be more than adequate measures of performance degradation in sustained operations research. They were not designed for measuring more subtle effects of fatigue, such as for the current study, even though the LRT task (a measure of higher order thinking) was sensitive enough to measure what we feel is a fatigue-induced performance decrement. 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.

The multitask showed fatigue effects for group averaged data in a similar pattern as for the LRT; i.e., at trial 3 of the outbound leg (Fig. 8). Further, multitask performance for 13 of the 18 pilots demonstrated varying degrees of fatigue-induced impact, ranging from minimal to marked impairment.

The correlation between increased fatigue and decreased performance has been reported by Barth et al. (5) French et al. (13,14), Mortimer (20), Perelli (21), and Shingledecker and Holding (25). Although the initial effects of fatigue can go unnoticed, eventually vigilance, judgment, situational awareness, and crew coordination may all be compromised (9). 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 reports demonstrated that aircrews routinely obtained less than 6 h of sleep per night. Almost 25% of the participating pilots in this study received less than 6 h sleep during their last night at home, immediately prior to the start of these missions. Less than 6 h has been found to produce significant decrements in performance (6). McCauley (18) found that several aircraft mishaps have been attributed to pilot fatigue because of disrupted and inadequate sleep. More than half of all aviation accidents are probably the direct result of human errors (7) which are probably due to fatigue-related pilot in-attentiveness (12).

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 1970's, several publications provided scheduling guidelines to limit the impact of long transmeridian flights on aircrews in an effort to avoid compromising flight safety (8,16,19). One of these scheduling systems (8) was under consideration by what was then the Air Transport Command of the Canadian Forces. Although 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 scheduling concessions, especially when crews are required to fly tactical transports designed for 3- to 4-h missions in a strategic role when crew days approach the maximum allowable crew-day, and sometimes, of necessity, exceed it.

One of the most compelling series of statistics in the fatigue literature is provided by Coren (11) who examined Canadian Ministry of Transportation data on motor vehicle accidents. He compared the number of accidents immediately before and after 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 7%, 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 resulting in a decrease of 7% in the number of reported accidents. However a week later, the accident rate increases again, to "normal." This data is compelling because it represents over 1.5 million accidents from all over Canada. If a 1-h change in sleep time affects the motor vehicle accident rate by 7% 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 1 h, fly long transatlantic missions in slow aircraft (a more demanding 3-dimensional task), encounter jet lag for several days in Europe, and return home via another long transatlantic flight, then deleterious effects of fatigue on crew performance would be expected, along with an attendant potential to compromise flight safety.

With respect to the limitations of sleep hygiene for the pilots and crews flying the missions monitored in this study, for the last night at home before the mission, departure should be delayed by 2 h. For the second night in the England, when the mission schedule dictates sleep at an inappropriately early circadian body clock time, the crews have difficulty sleeping. In an effort to facilitate such mission-driven early circadian sleep, we will undertake a study to determine whether or not melatonin or zopiclone can facilitate early circadian sleep, and if so, whether or not there is an attendant psychomotor performance liability. A psychomotor performance liability would preclude the use of either of these medications in aircrew.

The purpose of this study was to document the impact of fatigue on psychomotor performance and to assess the corresponding implications for flight safety. 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 theater of operations). Therefore, the results of this study can be considered a baseline with respect to the genesis of fatigue on routine transatlantic air transport operations. Should transport crews be called on to perform another airlift similar to Operation Alliance, with minimum opportunity for crew rest both during and between missions, aircrew fatigue would be expected to be much more severe than the fatigue we were able to measure on these routine re-supply missions.

CONCLUSIONS

Overall, our transport pilots showed a linear pattern of decreasing sleep over the last 3 d before embarking on these re-supply missions, with the last night at home in Trenton providing the least sleep of the entire mission period.

The self-rated scores for sleepiness, mental, and physical fatigue, indicate increases in sleepiness and fatigue throughout the long transatlantic flights.

The Logical Reasoning Task and Multitask performance indicates a possible impact of fatigue toward the end of the outbound transatlantic leg.

ACKNOWLEDGMENTS

We are indebted to the 18 pilots 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.

We are indebted to Dr. Jonathon French of the USAF Armstrong Labs in San Antonio, TX, for advice in this domain of research, and for lending us wrist actigraphs to collect our sleep data. We are also indebted to Dr. Manny Radomski for advice on the study and for reviewing this manuscript. The excellent work of Ms. Andrea Hawton is gratefully acknowledged She programmed all of our wrist actigraphs and downloaded the sleep data from the actigraphs to our computers. The data collection efforts of the following people are gratefully acknowledged: Capt. L. Donati, Mr. W. D. Fraser, Mr I. Mack, Ms. P. Odell, Wing Officer Y. Robichaud, and Sgt. L White.

REFERENCES

1 Angus RG, Heslegrave RJ. Effects of sleep loss on sustained cognitive performance during a command and control simulation. Behav Res Methods Inst Comp 1985; 17:55–67.
2. Angus RG, Pigeau RA, Heslegrave R Why we nap. In· Stampi C, (ed.) Human performance and sleep research. from the field to the laboratory Boston: Birkhauser, 1992.
3 Baddley AD. A 3 min reasoning test based on grammatical transformation. Psychonomic Sci 1968; 10:341–2.
4. Baird JA, Nicholson AN. Human factors of air operations in the South Atlantic Campaign. In: AGARD Conference Proceedings 338. Sustained Intensive Air Operations: Physiological and Performance Aspects, April 18–22, 1983, Paris, France. Neuilly sur Seine, France: Advisory Group for Aerospace Research and Development, November, 1983.
5 Barth JL, Holding JH, Stamford BA. Risk versus effort in the assessment of motor fatigue. J Mot Behav 1976, 8:189–94.
6. Belenky G, Penetar DM, Thorne D, et al. The effects of sleep deprivation on performance during continuous combat operations In. Marriott BM, ed. Food components to enhance performance. Washington, DC· National Academy Press, 1994.
7. Billings CD, Reynard WD. Human factors in aircraft incidents: results of a 7-year study. Aviat Space Environ Med 1984; 55:960–5.
8. Buley LE. Experience with a physiologically-based formula for determining rest periods on long-distance air travel. Aerosp Med 1970, 41:680–3.
9. Caldwell JA Fatigue in the aviation environment: an overview of the causes and effects as well as recommended countermeasures. Aviat Space Environ Med 1997; 68:932–8.
10. Cook MR, Cohen H, Orne MT. Recovery from fatigue. Fort Detrick, MD Army Research and Development Command, 1972; Tech Rep No 55.
11 Coren S. Sleep thieves: an eye-opening exploration into the science and mysteries of sleep. New York, NY· The Free Press, 1996; 304
12 Dinges DF, Graeber RC. Crew fatigue monitoring. Flight Safety Dig 1989; (Suppl)· 65–75.
13. French J, Bisson RU, Storm WF, Boll PA, Martinez J, Mitchida J, Neville KJ, Vogel E, Whitmore J Crew Performance and fatigue during sustained B1-B sorties. Brooks Air Force Base, TX: Armstrong Laboratory, Crew Systems Division, 1993, Project Futility Magellan (Report AL/CF-TR-1004–0014.
14 French J, Hannon P, Brainard GC. Effects of bright illuminance on body temperature and human performance. Annu Rev Chronopharmacol 1990; 7:45–9
15 Gander PA, Myhre G, Graeber RC, Andersen HT, Lauber K Adjustment of sleep and the circadian temperature rhythm after flights across nine time zones. Aviat Space Environ Med 1989; 60.733–42
16. Gerathewohl SJ. Simple calculator for determining the physiological rest period after jet flights involving time zone shifts Aerosp Med 1974; 45:449–50.
17 Hoddes E, Zarcone V, Smythe H, et al. Quantification of sleepiness: a new approach. Psychophysiology 1973, 10:431–6.
18. McCauley SF Aircrew fatigue countermeasures. In· AGARD Conference Proceedings, 495. Symposium on Progress in Military Airlift, May 23–31, 1990, 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. 238–47
20. Mortimer RG. Illness, drugs, fatigue, and stress in the cockpit reported by pilots In: Jensen RS, Neumeister D, eds. Proceedings of the Seventh International Symposium on Aviation Psychology. Colombus: Ohio State University, Dept. of Aviation, 1993, 650–3
21 Perelli LP. Fatigue stressors in simulated long-duration flight. effects on performance, information processing, subjective fatigue, and physiological cost. Brooks Air Force Base, TX: School of Aerospace Medicine, 1980, Report SAM-TR-80–49.
22. Pigeau R, Naitoh P, Buguet A, et al Modafinil, d-amphetamine, and placebo during 64 hours of sustained mental work. Effects on mood, fatigue, cognitive performance, and body temperature. J Sleep Res 1995, 4·212–28.
23 Ritter RD. "And we were tired" fatigue and aircrew errors. IEEE AES Systems Magazine; March 1993, 21–26.
24. Samel A, Wegman HM, Summa W, Naumann M. Sleep patterns in aircrew operating on the polar route between Germany and East Asia Aviat Space Environ Med 1991, 62·661–9.
25. Shingledecker CA, Holding DH. Risk, and effort measures of fatigue. J Mot Behav 1974; 6·17–25.
26 Wegman HM, Klein KE. Jet-lag and aircrew scheduling. In. Folkard S, Monk TH, eds. Hours of work. New York. Wiley, 1985.
27. Weinberg H, Heron R, Svoboda J, et al. The assessment and classification of complex information-processing in multitasking environments. In: McFadden S, Innes L, Hill M, eds. 12th Triennial Congress of the International Ergonomics Association. Mississauga, Canada: Human Factors Association of Canada, 1994:243–5.
28. Weinberg H, Carson P, Joly R, et al. Measurement, and monitoring of the effects of work schedule, and jet lag on the information processing capacity of individual pilots In: Jensen RS, et al., eds Tenth International Symposium on Aviation Psychology. Columbus, OH. The Ohio State University, 1999
29. Weinberg H, Jantzen KJ, Cheyne D, Carson P, Joly RJetlag, Gamma Activity, and Managing Pilot Fatigue In. Posluns H. (ed.) Transport Canada Conference Proceedings TP 13375. Fatigue in Transportation Workshop: Multimodal Issues and Solutions October 15–16, 1998. Ottawa· Transport Canada

#516720

CA020067