NASA/TM-2012-216053

Second Interim Report
NASA - easyJet Collaboration on the
Human Factors Monitoring Program (HFMP) Study

Ashok N. Srivistava
Ames Research Center
Moffett Field, California

Phil Barton
easyJet Airline Company, Ltd.
London Luton Airport, England

October 2012

NASA STI Program... in Profile

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA scientific and technical information (STI) program plays a key part in helping NASA maintain this important role.

The NASA STI program operates under the auspices of the Agency Chief Information Officer. It collects, organizes, provides for archiving, and disseminates NASA's STI. The NASA STI program provides access to the NASA Aeronautics and Space Database and its public interface, the NASA Technical Report Server, thus providing one of the largest collections of aeronautical and space science STI in the world. Results are published in both non-NASA channels and by NASA in the NASA STI Report Series, which includes the following report types:

- TECHNICAL PUBLICATION. Reports of completed research or a major significant phase of research that present the results of NASA Programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA counter part of peer-reviewed formal professional papers but has less stringent limitations on manuscript length and extent of graphic presentations.
- TECHNICAL MEMORANDUM. Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.
- CONTRACTOR REPORT. Scientific and technical findings by NASA-sponsored contractors and grantees.
- CONFERENCE PUBLICATION. Collects papers from scientific and technical conferences, synopsia, seminars, or other meetings sponsored or co-sponsored by NASA.
- SPECIAL PUBLICATION. Scientific, technical, or historical information from NASA programs, projects, and missions, often concerned with subjects having substantial public interest.
- TECHNICAL TRANSLATION. English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services also include creating custom thesauri, building customized databases, and organizing and publishing research results.

For more information about the NASA STI program, see the following:

- Access the NASA STI program home page at *http://www.sti.nasa.gov*
- E-mail your question via the Internet to help@sti.nasa.gov
- Fax your question to the NASA STI Help Desk at (301) 621-0134
- Phone the NASA STI Help Desk at (301) 621-0390
- Write to:
  NASA STI Help Desk
  NASA Center for AeroSpace Information
  7115 Standard Drive
  Hanover, MD 21076-1320

NASA/TM-2012-216053

Second Interim Report NASA - easyJet Collaboration on the Human Factors Monitoring Program (HFMP) Study

Ashok N. Srivistava
Ames Research Center
Moffett Field, California

Phil Barton
easyJet Airline Company, Ltd.
London Luton Airport, England

National Aeronautics and
Space Administration

Ames Research Center
Moffett Field, California, 94035-1000

October 2012

Available from:

NASA Center for Aerospace Information
7115 Standard Drive
Hanover, MD 21076-1320
443-757-5802

National Technical Information Service
5285 Port Royal Road
Springfield. VA 22161
703-487-4650

Ames Research Center

easyJet
Airline

Second Interim Report

NASA – easyJet Collaboration on the Human Factors Monitoring Program (HFMP) Study

Ashok N. Srivastava, PhD. Project Officer for NASA
Phil Barton, Project Officer for easyJet Airline Company, Ltd.

August XX, 2012

Prepared by:

NASA Ames Research Center, Code TH, Mountain View, CA
EasyJet Airline Company Ltd, London Luton Airport, England
Stinger Ghaffarian Technologies Inc., Mountain View, CA
San Jose State University, San Jose, C

Table of Contents

Table of Figures

Table of Tables

Executive Summary

This is the second interim report jointly prepared by NASA and easyJet on the work performed under the agreement to collaborate on a study of the factors entailed in flight- and cabin-crew fatigue and decreases in performance associated with fatigue. The objective of this Agreement is to generate reliable procedures that aid in understanding the levels and characteristics of flight- and cabin-crew fatigue factors, both latent and proximate, whose confluence will likely result in unacceptable crew performance. This study entails the analyses of numerical and textual data collected during operational flights. NASA and easyJet are both interested in assessing and testing NASA's automated capabilities for extracting operationally significant information from very large, diverse (textual and numerical) databases; much larger than can be handled practically by human experts.

This report is about the analyses performed on the data collected during a third study conducted from 1 July through 25 August 2011 in which the voluntary participants were 22 pilots. Since the issuance of the first report on this collaborative study, we have made significant progress toward the objectives of the Agreement.

This report presents descriptions of the factors on which data were collected and how each was measured during the July-August 2011 study period. We completed a review of all of the data and found them adequate in quality and quantity to perform reliable analyses. With a few exceptions, the pilots participating in this study were conscientious about recording the data as scheduled.

Our review of the data revealed significant variability among the 22 pilots and so, while some analyses of the group data were performed, our primary interest was in the data from each pilot. The results of the group data analysis and their implications to the objectives of this study are discussed. After examining the group data for understanding and guidance, we turned our attention to individual data so as to be able to create an individual profile based on the measures collected.

We found that, on the average for the group of pilots, neither the levels of sleepiness nor the degradations of performance reached disturbing levels at any time during the study. An important conclusion from the analyses of the data for the individual pilots is that they experienced degradation of performance very infrequently in flight during the Duty Days as flown during this study. Since the Samn-Perelli Scores seldom exceed 5, the fairly low levels of sleepiness for most pilots during this study seem to have had little effect on their performance as measured by their Mean PVT Response Times. This might be explained by the ability of a well–trained professional to adapt to and overcome low levels of tiredness to perform his or her job acceptably. The evidence is that there was statistically insignificant potential for degraded vigilance performance due to fatigue during this study.

A finding important to achieving the objectives of this study is that the Samn-Perelli Scores (a subjective measure of fatigue) correlated significantly with the Mean PVT Response Times (an objective measure of cognitive performance) for the group and for 12 of 21 pilots.

Despite the good correlation between Samn-Perelli Scores and Mean PVT Response Times, we found significant differences between the two measures. For example, in contrast with the results for the Mean PVT Response Times, the Samn-Perelli Scores are strongly correlated with the time of day at which the test was recorded over the entire roster, over the Duty Days, and, especially, over the Rest Days. This is evidence of the strong influence of the circadian rhythm on the feeling of sleepiness that does not necessarily impact the performance of an assigned task by a professional.

We found that the self-perception of being ready to perform was invariably higher than the readiness to perform and especially in the evening hours.

There is a tendency among pilots to sleep less on a night prior to a day of early departure and more on a night following late finish. Humans are able to extend their sleep time in the morning more easily than they are able to start their sleep time early in the evening. Sleep cannot be scheduled at anytime of day and expected to be of equal quality.

Finally, and very importantly, we have found no evidence to say that any of the unwanted events experienced by the aircraft flown during this study were related to degraded vigilance performance of a crewmember due to sleepiness.

Although the small group of pilots who volunteered to participate in this study cannot be considered representative of the general population of easyJet pilots and although the individualistic differences are important, we believe that most of the results and conclusions we have reached so far apply to any group of similarly trained pilots flying the same FRV roster and comparable flights as those flown in this study.

We consider the analyses of correlations to be completed and we next will focus on building an individualistic model of fatigue for each pilot.

1. Introduction

Background

Fatigue is a physiological state in which there is a decreased capacity to perform tasks and an increased variability in performance. However, the primary concern in aviation is about the more subtle aspect of fatigue that is better denoted as ‘drowsiness’. Drowsiness may have little impact on the ability to perform trained motor skills, but can profoundly affect cognitive reaction time, vigilance, memory, and mistakes during even routine decision-making. Most importantly to aviation operations is that drowsiness increases the risk of pilot error. The ability to cope with unusual or unexpected situations deteriorates with drowsiness.

In this report, we will use *“fatigue”* and *“drowsiness”* interchangeably with the understanding that physical fatigue is seldom a factor in the concerns of this study.

A primary cause of drowsiness is, of course, attributed to periods of extended wakefulness in which ample recovery sleep is not obtained and the inherent 24-hour biological clock (called circadian rhythm) is disrupted. There are two natural circadian lows; one is a dip in the afternoon between 13:00 and 16:00, the other is a more pronounced low between 03:00 and 06:00. During the latter, known as the circadian nadir, mental and physical acuity are blunted. This is not the best time to be trying to cope with unexpected or emergency situations.

There are many factors in aviation operations that can cause sleep deprivation and disruption of circadian rhythm including irregular schedule, multiple flight legs, long duty days, reduced time off, and early report times. There are other factors that can contribute to drowsiness such as extended periods of vigilance, lighting conditions, the use of countermeasures, life style factors, unbalanced nutrition, job strain, and jet lag. Further, all of these factors are exacerbated by the increasingly complex operations that continue around-the-clock and the stressors associated with the adverse workplace environment of the flight deck. A good night’s sleep is the only panacea, but the timing, quality, and duration are important to its effectiveness. Sleep loss over several days can accumulate into sleep debt and desynchronosis from which it can take a long time to readjust. A roster that shifts rapidly between night duties and day duties may cause both desynchronosis and sleep loss.

Maximizing alertness and performance levels during aviation operations is critical to maintaining the continued safety of the air transportation system. The challenge is the difficulty of aligning the scientific findings regarding sleep deprivation, circadian principles, and working schedules and conditions with operational requirements. Moreover, there has been no clear evidence on which to establish the level of fatigue or drowsiness that causes a degradation of the pilot’s ability to perform his or her functions and operate the aircraft responsibly and safely. The difficulty of coping with this issue is exacerbated by the large differences among individuals in their reactions to the various causal and contributing factors of fatigue.

A requirement of Safety Management Systems across the international aviation industry is to monitor and predict fatigue of flight crews. The response to the need for a scientifically valid fatigue-management approach that will lead to continuous safety enhancements by identifying and addressing both physiological and operational fatigue factors across time and changing circumstances is the institution of a Fatigue Risk Management System (FRMS). A FRMS is a data-driven, scientifically based process that provides for continuous monitoring and management of safety risks associated with fatigue-related error. This process leads to continuous safety enhancements by identifying and addressing fatigue factors across time and changing physiological and operational circumstances. Key components of the FRMS approach are: 1) access to fatigue related data; 2) fatigue analysis methods; 3) identification and management of fatigue drivers, and 4) application of fatigue mitigation procedures.

Project overview

EasyJet Airline Company Ltd. has initiated the Human Factors Monitoring Program (HFMP) in support of its FRMS to better understand how both latent and proximate causal fatigue factors potentially contribute to impaired flight- and cabin-crew performance. Most of the research related to human fatigue in aviation has either been conducted in the laboratory or has focused on military operations or commercial long-haul pilots, but pilots of the new non-legacy air carriers like easyJet may also experience elevated levels of fatigue. An objective of easyJet's FRMS is to establish a roster design for their operations that minimizes the potential of incurring levels of crew fatigue that can contribute to decrements in aircraft performance. It will be necessary to identify the simplest reliable measurement system for monitoring fatigue, crew performance, and aircraft performance and to validate a predictive model to achieve the objectives of the SMS and the FRMS.

NASA is collaborating with easyJet on the HFMP studies by providing technologies and methodologies to enable a data-driven and scientifically based process to monitor and manage safety risks associated with fatigue-related error.

The initial report on the NASA-easyJet collaboration on HFMP published October 1, 2010 was the NASA publication titled "First Annual Report: NASA–easyJet Collaboration on the Human Factors Monitoring Program (HFMP) Study" co-authored by the two Project Officers, Ashok N. Srivastava for NASA and Phil Barton for easyJet Airline Company, Ltd.

The link to this report is: http://ti.arc.nasa.gov/publications/3491/download/

At NASA, this work is incorporated within the Data Mining and Knowledge Discovery Theme and the Human Systems Solutions Theme of NASA's System-wide Safety and Assurance Technologies (SSAT) Program. At easyJet, the HFMP is under the auspices of their Fatigue Risk Management System, which has been incorporated as part of easyJet's Safety Management System.

The HFMP entails acquiring, processing, integrating, and interpreting large quantities of diverse numerical and textual data collected from flight-crew and cabin-crew participants during easyJet's normal operations. Under the terms of a Space Act Agreement (SAA) with easyJet, NASA has been granted access to data on aircraft performance and on crew performance. NASA's Intelligent Systems Division (Code TI) is analyzing the aircraft-performance data and NASA's Human Systems Integration Division (Code TH) is analyzing the crew-performance data. The results of these two efforts will be combined to examine the relationship between aircraft performance and crewmember performance.

EasyJet is seeking answers to the following specific questions:

1. Is there reliable evidence that levels of fatigue can be correlated with scheduling strategy **AND** is there reliable evidence that work pattern (scheduling strategy) is a causal factor of performance-degrading levels of fatigue?
2. What is the minimum set of measures to reliably indicate that identified aircraft performance decrements were probably related to fatigue?
3. What levels of crew fatigue are likely to cause degraded aircraft performance?
4. Pragmatically, which measures can be implemented during normal operations to monitor for levels of human fatigue that could affect performance?
5. What are the data sources that provide reliable information on the consequences of performance-degrading levels of fatigue?
6. What are the data sources that provide reliable information on the latent and proximate causal and contributing factors of human fatigue?
7. What are the fatigue profiles of operators based on individual measures over the course of flights? Are these indicators convergent?
8. Is there a predictive model that can be used reliably to design interventions?

NASA proposes to find answers to these challenging questions by using the linked, time-stamped data to look for:

a) causal correlation of an event indicative of a decrement in aircraft performance with an identified decrement in crew performance
b) causal correlation of a decrement in crew performance with fatigue (drowsiness), and
c) the causal and contributing factors of fatigue.

A goal is to identify the simplest reliable measurement system for monitoring and relating fatigue, crew performance, and aircraft performance. The challenge is two-fold: how to measure drowsiness and how to establish that drowsiness was a causal factor of an operational event.

This is a report of work in progress on the analysis of the data collected during the third study in the conduct of the HFMP.

The data we collected

During July and August 2011, easyJet collected data for the HFMP study from 22 volunteers of easyJet's pilots. All the subjects of this study flew the Flexible Roster Variation (FRV) schedule diagramed in Figure 1. FRV allows for some variability in assigned flights depending on base of operations. By design, each participating pilot returned to his or her home base at the end of every Duty Day during this study.

D/O=Day Off; E=Early Departure; L=Late Arrival

Rest Days			Duty Days-Block A					Rest Days			Duty Days-Block B				
D/O	D/O	D/O	E1	E2	E3	L1	L2	D/O	D/O	D/O	E1	E2	E3	L1	L2

Rest Days		Duty Days-Block C					Rest Days		
D/O	D/O	L1	L2	L3	L4	L5	D/O	D/O	D/O

Figure 1-FRV Schedule of 26 Consecutive Duty and Rest Days[1]

By easyJet's definitions, for the purpose of this study, an early start (E) is a duty day that commences between 03:00 and 09:29 local time inclusive. A Normal (No) duty day starts and finishes between 09:30 and 17:59 local time inclusive. A duty day is considered to have a late finish (L) when the duty day ends between 18:00 and 03:00 local time inclusive. A Night (Ni) duty day is one that commences between 18:00 and 02:59 local time inclusive and finishes between 03:00 and 17:59 local time inclusive.

There were variations within the roster shown in Figure 1 among the participating pilots, but, for most pilots, Block A of the Duty Days included 3 early-start days and 2 late-start days, which allowed for collection of sleep and performance data during the first schedule transition. Following Block A, 3 Rest Days were scheduled. The second duty block (Block B) contained one further transition change and the duty sequence closely reflected timings and workload of Block A, for comparison purposes. Another two Rest Days were provided following Block B. Data were then collected for another 5-day duty period (Block C) in order to compare performance levels following two days off to that following three days off.

Table 1 shows the data that were collected, the days and times for each, and the methods of recording. All data were annotated with a common time-stamp (GMT+1) to enable their linkage. Details of these measures are described later in this report. The line item identified as "FOQA" in Table 1 is the in-flight-recorded data. FOQA means Flight Operational Quality Assurance, which is the name given to the program in the U.S. and is equivalent to the European Flight Data Monitoring (FDM) program. The FOQA and FDM programs entail continuously recording during flight and analyzing data on hundreds of flight parameters.

---

[1] The roster used in the study on which this report is based was different than the FRV roster defined in the first report. Block C for this study included only late-finish Duty Days for all pilots.

Table 1 – HFMP Study Measures

	Training Day	**Rest Days**	**Duty Days (Blocks A, B & C)**	**Non-operated duty (standby)**	**Recovery Days**	**Method of Data collection**
Samn Perelli-subjective alertness scale	Instructed how to use	On waking + 20 min. Morning = 1-2 hrs after waking Mid-day = 8-9 hrs after waking Evening = 1-2 hrs before sleeping On lights out at bedtime	On waking + 20 min. Pre - flight Top of Descent each sector Post flight (duty log-off) On lights out at bedtime	On waking + 20 min. Morning Mid-day Evening On lights out at bedtime	On waking + 20 min. Morning Mid-day Evening On lights out at bedtime	Paper and pencil and also entered in Actiwatch
PVT-psychomotor vigilance task PDA	Instructed how to use	Morning = 1-2 hrs after waking Mid-day = 8-9 hrs after waking Evening = 1-2 hrs before sleep	Top of Descent each sector Post flight (duty log-off)	Morning Mid-day Evening	Morning Mid-day Evening	PDA
Use of Fatigue Countermeasures*	Instructed how to use	Not collected	Turnarounds, throughout roster Rosters vary by individual – refer to master spreadsheet for duty timing.	Not collected	Not collected	Paper and pencil
NASA TLX	Instructed how to use	Not collected	Turnarounds	Not collected	Not collected	Paper and pencil
Hassle factors	Instructed how to use	Not collected	Turnarounds	Not collected	Not collected	Paper and pencil
Mood/Alertness Scale	Instructed how to use	Not collected	Prior to commute to work Post work commute home on Bio Harness download	Not collected	Not collected	Laptop
Sleep diary	Instructed how to use	Pre-sleep Post-sleep (20 minutes after waking)	Pre-sleep Post-sleep (20 minutes after waking)	Pre-sleep Post-sleep (20 minutes after waking)	Pre-sleep Post-sleep (20 minutes after waking)	Paper and pencil
Actigraphy	Instructed how to use	Continuous	Continuous	Continuous	Continuous	Actigraph
BioHarness-Physiological variables	Instructed how to use	Not collected)	< 16 hrs (daytime; during duty)	Not collected)	Not collected))	Bio harness
FOQA	N/A	Not collected)	Collected per flight/individual	Not collected	Not collected	A/C FDM system
			Questionnaires			
CIS-checklist of individual strength	Completed training day only	Not collected	Not collected	Not collected	Not collected	Paper and pencil

MES-morning evening scale	Completed training day only	Not collected	Not collected	Not collected	Not collected	Paper and pencil
ESS-epworth sleepiness scale	Completed training day only	Not collected	Not collected	Not collected	Not collected	Paper and pencil

*Summary of Countermeasures:

FATIGUE COUNTERMEASURES - FLIGHT DECK	**TICK**	**QUANTITY**
Cockpit napping		Time & duration
Activity Breaks		Time & duration
Caffeine intake		No. Cups
Crew communications		
Increased monitoring and cross checking		
Workload sharing/offload		
Increased briefing times and time for task actioning		
Automation application/reliance		
Cockpit lighting		
Crew offload/replacement		Sector & no. crew
OTHER (LIST):		

FATIGUE COUNTERMEASURES – CABIN CREW	**AMOUNT**
No countermeasures employed	
Activity Breaks (number)	
Caffeine intake (cups)	
Crew communications (tick)	
Increased monitoring & checking (tick)	
Workload sharing/offload (tick)	
Increased briefing times and time for task actioning (tick)	
Reduced services (tick)	
Crew offload/replacement (tick)	
OTHER (LIST):	

2. Crew Performance

Background

NASA Ames Research Center’s Human Systems Integration Division’s (Code TH) role is to analyze the physiological and neuro-cognitive measurements to track levels of fatigue and performance for each subject across the time-course of the roster. Individual fatigue and performance levels will also be analyzed in conjunction with personal profile variables to identify the most likely proximate and latent causal factors of fatigue. The goals of this research are to

1. Identify scheduling factors, physiological measures, and cognitive variables that have potential for predicting degraded levels of cognitive functioning.
2. Develop recommendations for fatigue mitigation and scheduling adjustments to maximize performance and alertness levels during aviation operations.
3. Identify the fatigue-related measures and data sources of individual-performance effectiveness that are feasible for inclusion in an FRMS so as to enable continuous monitoring and management of fatigue-related safety and performance risks during aviation operations.

Approach to Data Analysis

The data collected during actual flight operations will be analyzed for each participant over the course of each data-collection period to:

1) Determine if there is a causal correlation between decrements in measurements of individual performance and measurements of the subject’s level of fatigue, and

2) Identify potential causal factors of fatigue.

In pursuing this objective, it is important to note that not all degradations in human performance are due to fatigue and not all levels of fatigue produce significant decrements in human performance.

The causal relationships to be explored for each of the participants are diagramed in Figure 2.

Figure 2- Representation of Crew-Performance Data Analyses

Group statistics will be calculated for correlation of each of the measures of potential causal factors with Level of Fatigue to determine which of them show significant variability among members of the sample group. If any do not, it opens the possibility of generalizing that measure to the full population. We will also assess the informational value of each measurement to identify the minimum number of useful measures. After examining the group data, the focus will turn to analyzing individual data to create an individual personal profile of fatigue-related factors.

Individual Performance

The only measurement of individual crew performance in these studies was the **Psychomotor Vigilance Task (PVT).** The PVT, an objective, neuro-cognitive measure, has become the most widely used measure of fatigue effects on vigilance performance due to its convenience and its high sensitivity to sleep deprivation. (Dinges and Powell 1985) The most commonly used PVT performance metrics are (1) lapses – the cumulative number of reaction times exceeding 500ms – a consistent indicator of deficits in sustained attention; 2) median reaction times – to measure central tendency in response times uninfluenced by outliers; 3) optimum response times – or fastest 10% of reciprocal response times for all trials – an indication of the best performance a participant is capable of producing. However, due to the practical limitations on acquiring data during operations, our primary measure will be the Mean PVT Response Times.[2]

Level of Fatigue

The measurements of level of fatigue (drowsiness) were the **Samn-Perelli** scale and the **BioHarness** measurement of physiological variables.

The Samn-Perelli scale is a subjective self-assessment of the individual's level of fatigue based on the following descriptions

Degree of Fatigue	**Scale Rating**
Fully alert, wide awake	1
Very lively, responsive, but not at peak	2
Okay, somewhat fresh	3
A little tired, less than fresh	4
Moderately tired, let down	5
Extremely tired, very difficult to concentrate	6
Completely exhausted, unable to function effectively	7

Specific scores on the checklist have been used as thresholds to trigger a set of fatigue risk controls, such as the following (modified) example taken from Queensland Government's Fatigue Risk Management System Resource Pack 2009:

Samn-Perelli Scale Rating	**Risk level**	**Controls**
1–3	Low	No specific controls necessary. Except in the presence of higher-level indicators of fatigue (i.e., symptoms, errors or incidents).
4–5	Moderate	Initiate moderate fatigue risk mitigation actions
6	High	Initiate high fatigue risk mitigation actions.
7	Very high	Intolerable risk. No individual rostered beyond this threshold.

---

[2] We also examined PVT Lapses, but found that the data were not adequate for sufficiently reliable analysis.

Each participant in this study was given an ambulatory physiological monitor (BioHarness, Zephyr Technology Inc.) and a laptop computer for downloading data and recharging the BioHarness. The BioHarness (Figure 3) is a small data-logging device that connects to a Velcro fabric strap worn around the chest.

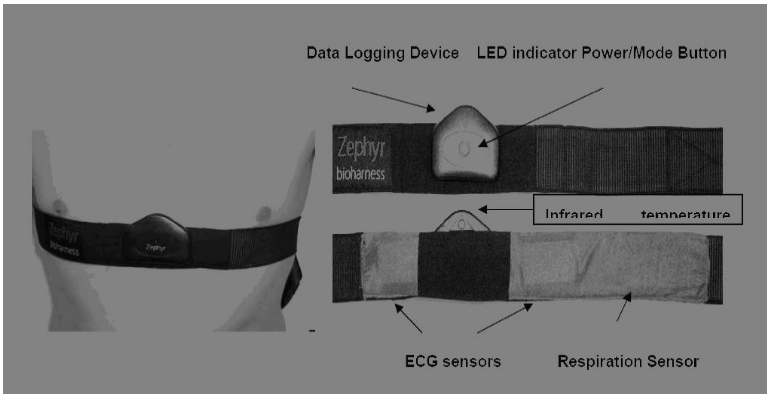


Figure 3- Ambulatory Monitoring System: Zephyr BioHarness

The BioHarness measures (1) Electrocardiography (ECG), (2) Respiration Rate (RR), (3) Skin Temperature (ST), (4) Posture (upright or supine body position), and (5) Physical Activity (three-axis accelerometer). Data can be recorded for up to 16 hours and are downloaded to a laptop computer by using a USB cradle, which also recharges the battery. Participants were instructed to don the BioHarness after awakening (pre-shift) each day and to remove it post-shift, at which time they were instructed to download their data and recharge the BioHarness module.

Laboratory experiments have indicated that there are 4 basic physiological measures needed to provide the information to objectively assess fatigue. The BioHarness provides 2 of these, namely, ECG and RR. A complementary device is the Q-sensor (Figure 4), which, together with the BioHarness can provide information on the other two essential physiological measures, namely, skin conductance level and peripheral blood flow. The Q-Sensor also incorporates a three-dimensional motion sensor so that it can serve the function of an Actiwatch.

Figure 4- The Q-Sensor

While the Q-Sensor could not be acquired in time for the experiment reported on here, it was not considered critical as our primary interest was in evaluating the BioHarness itself in this first operational aviation study in which it was used.

One question to be considered is whether these measures of Level of Fatigue correlate with each other and are causally related to the PVT measures of crew performance. The other question is to what extent are the measured personal profile factors causally correlated with the measures of Level of Fatigue.

Factors of Fatigue

The assumption is that information from the measurements of the proximate and the latent factors that could cause or contribute to fatigue (drowsiness) can be combined to provide reliable correlation with each individual's measure of Level of Fatigue. The personal profile factors measured during this study that could be factors of fatigue are:

Latent

Demographic questionnaires
Morningness/Eveningness Scale (MES)
Checklist Individual Strength (CIS)
Epworth Sleepiness Scale (ESS)
Mood/Alertness Scale (M/AS)

Proximate

Sleep diary
Actigraphy
Work pattern (roster)
NASA Task Load Index (TLX)
Hassle Factors
Fatigue Countermeasures

Following are detailed descriptions of the measurements made of each of these factors.

Latent Factors

Information on the latent factors was obtained through several questionnaires that each participant completed before the start of the experiment and data collection.

In the questionnaire on the subject's **"Demographics",** each participant provided the following information:

Date of Birth
Age
Gender
Nationality
Marital Status
Children at home under 12 yrs
Infants under 2 yrs
Is duty day commute from permanent or temp address
Permanent address commute
Temporary address commute
Lives in a share house
Regular smoker
Weight (kg)
Height (cm)
Total commercial hours
Total hours on type
easyJet hours
Sleep per 24 hours to be fully alert
Most alert times
Most tired times
Duty day minimum sleep (in hrs)
Duty day maximum sleep (in hrs)
What time do you go to bed on non-duty days
What time do you awake on non-duty days
Non-duty days minimum sleep duration
Non-duty days maximum sleep duration

A basic latent factor that can account for differences in the effects of sleep deprivation is the individual behavioral and physiological circadian rhythm, which is referred to as chronotype, circadian type, or diurnal preference. Chronotype is an animal attribute that reflects at what time of the day their physical functions (hormone level, body temperature, cognitive faculties, eating and sleeping) are active, change, or reach a certain level. Human activity-rest patterns are endogenously controlled by circadian rhythms. Morning people (referred to as "larks") are those who wake up early and are most alert in the first part of the day. Evening people (referred to as "owls") are those who are most alert in the late evening hours and prefer to go to bed late. Morning types typically arise and retire about 1-2 hours before evening types. Evening type is associated with a greater need for sleep and less time in bed. These individual differences in circadian rhythm characteristics suggest that there will be significant differences in the capacity of a given crew member to maintain adequate sleep and physiologically adapt to work-rest schedule shifts and shortened sleep-wake cycle schedules. The subject's chronotype was identified at the beginning of the study with the **Morningness/Eveningness Scale (MES)**. This questionnaire has 19 questions (See Appendix A), each with a number of points and the total scores are rated as follows

16-30 definite evening type
31-41 moderate evening type
42-58 intermediate
59-69 moderate morning type
70-86 definite morning type

Each participant also completed the **Checklist Individual Strength (CIS)** questionnaire at the beginning of the study. Respondents rate the extent to which each of the following 20 statements is true for them in the past two weeks on a seven-point Likert scale ranging from "Yes, that is true" to "No, that is not true." For the items: 2, 5, 6, 7, 8, 11, 12, 15, 20, the scoring ranges from 1 for "Yes, that is true" to 7 for "No, that is not true". For the items: 1, 3, 4, 9, 10, 13, 14, 16, 17, 18, 19 the scoring is from 1 for "No, that is not true" to 7 for "Yes, that is true"

1 I feel tired
2 I feel very active
3 Thinking requires effort
4 Physically I feel exhausted
5 I feel like doing all kinds of nice things
6 I feel fit
7 I do quite a lot within a day
8 When I am doing something, I can concentrate quite well
9 I feel weak
10 I don't do much during the day
11 I can concentrate well
12 I feel rested
13 I have trouble concentrating
14 Physically I feel I am in bad condition
15 I am full of plans
16 I get tired very quickly
17 I have low output
18 I feel no desire to do anything
19 My thoughts easily wander
20 Physically I feel in good shape

The responses to these CIS questions can be summarized with respect to the four personal characteristics of fatigue that they capture and the scores for the relevant items added.

Subjective Feeling of Fatigue: Statements 1, 4, 6, 9, 12, 14, 16, and 20
Highest feeling of fatigue = 56
Lowest feeling of fatigue = 8

Reduction in Concentration: Statements 3, 8, 11, 13, and 19
Highest level of reduction in concentration = 35
Lowest level of reduction in concentration = 5

Reduction in Motivation: Statements 2, 5, 15, and 18
Highest level of reduction in motivation = 28
Lowest level of reduction in motivation = 4

Reduction in Physical Activity: Statements 7, 10, and 17
Highest level of reduction in physical activity = 21
Lowest level of reduction in physical activity = 3

Also, a total CIS score can be calculated, by adding the scores in the four dimensions. The total CIS score may range from 20 to 140. Higher total scores indicate a higher degree of fatigue, more concentration problems, reduced motivation, and less activity. CIS was developed as a measure of Chronic Fatigue Syndrome (CFS). CFS is persistent or relapsing fatigue that is not alleviated by rest, is not explainable by medical or psychiatric conditions, and causes significant reduction of activities. (Vercoulen et al 1994; Beurskens et al 2000).

The **Epworth Sleepiness Scale (ESS)** explores the likelihood of falling asleep in various everyday situations in contrast to simply feeling tired. (Johns 1991). Each participant is asked to choose the most appropriate number from the following scale:

0 = Would **never** doze
1 = **Slight** chance of dozing
2 = **Moderate** chance of dozing
3 = **High** chance of dozing

in each of the following 8 situations:

Sitting and Reading
Watching TV
Sitting inactive in a public place
As a passenger in a car for an hour without a break
Lying down to rest in the afternoon when circumstances permit
Sitting and talking to someone
Sitting quietly after lunch without alcohol
In a car, while stopped for a few minutes in traffic

The total ESS score provides an estimate of the individual's level of sleepiness in daily life. The normal range of the total ESS score is considered to be 0-10. A score of 10 or more is considered sleepy and a score of 18 or more is very sleepy. An individual who scores 10 or more on this test should consider the need to improve sleep hygiene.

The data for the **Mood/Alertness Scale (M/AS)** are complementary to the MES, CIS, and ESS data that are used to aid in the analysis of the BioHarness data. The M/AS provides information about the individual's 'Readiness to Perform' based on physical sensations and on his or her subjective perception of readiness. A program for recording the Mood/Alertness data was provided in the laptop computer to which the BioHarness data were downloaded. On Duty Days, participants were asked to enter their Mood/Alertness data before the first sector is flown each day and again at the same time that the BioHarness data were downloaded each evening.

The M/AS consists of eight questions regarding mood states and two questions about the previous night's sleep. The scores for responses to 9 of these questions may range from the unfavorable end of the scale (0) to the favorable end of the scale (10). The 10th question asks for the number of times the previous night's sleep had been interrupted.

The Mood/Alertness scale is further described in Appendix B. The analysts of the BioHarness data use an 'Activation scale' (interpreted as 'Readiness to Perform'), which is a composite score made up of the average of sub-scales:

Motivation
Arousal
Fatigue
Ease of concentration

and an 'Affective scale' (interpreted as 'Perceived Readiness to Perform'), which is a composite score made up of an average of sub-scales:

Tension level
Feels
Physical discomfort
Contentedness

Proximate Factors

Among the most important proximate factors are the timing, quantity, and quality of sleep across the roster period. These were measured with an **Actiwatch** and the recordings in a **Sleep Diary**.

The **Actiwatch** records gross motor activity that can be used to visualize rest activity patterns and quantify physical activity or sleep. It is useful for determining sleep patterns and circadian rhythms and may be worn for several weeks at a time. Actigraphy offers reliable results with an accuracy that is close to those of polysomnography (PSG), which monitors many body functions including brain (EEG), eye movements (EOG), muscle activity (EMG) and heart rhythm (ECG) during sleep.


The Actiwatch

For every night of the study, each participant filled out a **Sleep Diary** to provide subjective measures of sleep and wake behavior during the period of participation. The sleep diary (Shown in Appendix C) consisted of several A4 pages (included in the beginning of the Workbook) on which various details about each day were recorded. All times reported in the diary were based on Universal Time Coordinated (UTC) +1. Participants were requested to record information in their sleep diaries immediately before "lights-out" and 20 minutes after waking, thereby facilitating more accurate fatigue and sleep quality ratings. The sleep diaries asked participants to provide the following information:

1. The sleep date and time of bedtime (i.e. "lights out");
2. Date and time of wake-up;

3. The perceived quality of sleep periods and naps compared to a "normal" sleep period on the following scale:

    1. Very good
    2. Good
    3. Average
    4. Poor
    5. Very poor
    6. Did not sleep

4. The level of fatigue immediately before and after each main sleep period on the following scale:

    1. Fully alert, wide-awake
    2. Very lively, responsive, but not at peak
    3. Okay, somewhat fresh
    4. A little tired, less than fresh
    5. Moderately tired, let down
    6. Extremely tired
    7. Completely exhausted, unable to function effectively

5. Comments on causes of interrupted sleep, nap times, and times of activities, such as shower, when the Actiwatch is not worn.

The other proximate factor is the individual **Work Pattern** in accordance with the FRV diagram of Figure 1. Daily flight schedules included early work shifts beginning at approximately 0600 hours and late work shifts beginning at approximately 1200 hours. Typically, the early shifts entailed two flights, while late shifts had four flights with longer workdays. There were individual differences in number of sectors flown in a day and number of early departures and late arrivals in a block of Duty Days.

An individual's perception of his/her workload and hassle factors may influence (possibly subconsciously) the participant's self-assessment of level of fatigue. The measures of these factors are the **NASA Task Load Index (TLX)** and a rating scale of **Hassle Factors.** They will be examined for evidence of consistent effects. In using the **NASA TLX** Scale (Table 2), participants were asked to assess 6 aspects of workload for each flight on a scale of from very low or perfect (1) to very high or failure (100) in increments of 5:

Table 2 - TLX Rating Scale

Title	Descriptions	Rating
MENTAL DEMAND	How mentally demanding was the task?	Very Low … Very High
PHYSICAL DEMAND	How physically demanding was the task?	Very Low … Very High
TEMPORAL DEMAND	How hurried or rushed was the pace of the task?	Very Low … Very High
EFFORT	How hard did you have to work (mentally and physically) to accomplish	

	your level of performance?	Very Low	Very High
PERFORMANCE	How satisfied were you in accomplishing what you were expected to do?	Perfect	Failure
FRUSTRATION	How insecure, discouraged, irritated, stressed and annoyed were you?	Very Low	Very High

Participants were also asked to indicate the presence of any of the **Hassle Factors** defined in Table 3.

Table 3 - Hassle Factors

Airport Facilities
Cabin Activity
ATC
External Environment
Aircraft Environment
Training
Airline Disruption i.e. crewing, a/c change
Procedures & Documentation
Human Factors
Team Factors
Experience
OTHER (LIST):

The final proximate factor to be considered is the influence of the use of countermeasures on the individual's level of fatigue. Professionals are known to use a variety of strategies to overcome and compensate for the effects of sleepiness on their ability to perform. Flight crews were asked to annotate which of the **Fatigue Countermeasure** strategies identified in Table 4 had been employed at the time they complete the TLX and Hassle scales during sector turnarounds throughout the rostered duties and recovery days. At this time, it is not clear how this information should be incorporated and weighted in considering the causal and contributing factors of drowsiness.

Table 4 - Fatigue Countermeasures

Cockpit napping
Activity Breaks
Caffeine intake
Crew communications
Increased monitoring and cross checking
Workload sharing/offload
Increased briefing times and time for task actioning
Automation application/reliance
Cockpit lighting
Crew offload/replacement
Other (List):

An objective is to determine how the measurements of any or all of these proximate and latent factors of fatigue should be weighted and combined in a reliable predictive model of the measured 'Level of Fatigue' for each of the participants.

Studies of the Data Collected on Crew Performance and Fatigue

This report focuses on the data described above that were collected by flight crews during the period from 1 July through 25 August 2011. With only a few exceptions, the pilots participating in this study were conscientious about recording the data on schedule. However, the quality and quantity of the data collected varied across the participants and with the data sets. The nature of these differences and the suggestions for improvements in the next study are discussed below with respect to each of the data sets.

The focus of the work reported in this section has been on evaluation of data quality and quantity for achieving the objectives of this research and on exploratory analyses. Although primary interest will be on analyses of individual characteristics, some group analyses were performed. The results and possible implications of the group data analysis to the objectives of this study are discussed. After examining the group data for understanding and guidance, the focus turned to individual data so as to be able to create an individual profile based on the measures collected. Our initial analysis addressed correlation between measurements. Then we must show causation of these factors to be able to structure a predictive model.

Individual Performance

Psychomotor Vigilance Task (PVT)

The quantity and quality of the PVT data collected during this study were far better than those collected in the previous two studies. Table 5 indicates missing data in the recordings of PVT. The days highlighted in color indicate the extent of incomplete data on each day as follows:

🟥 - no measurements were taken

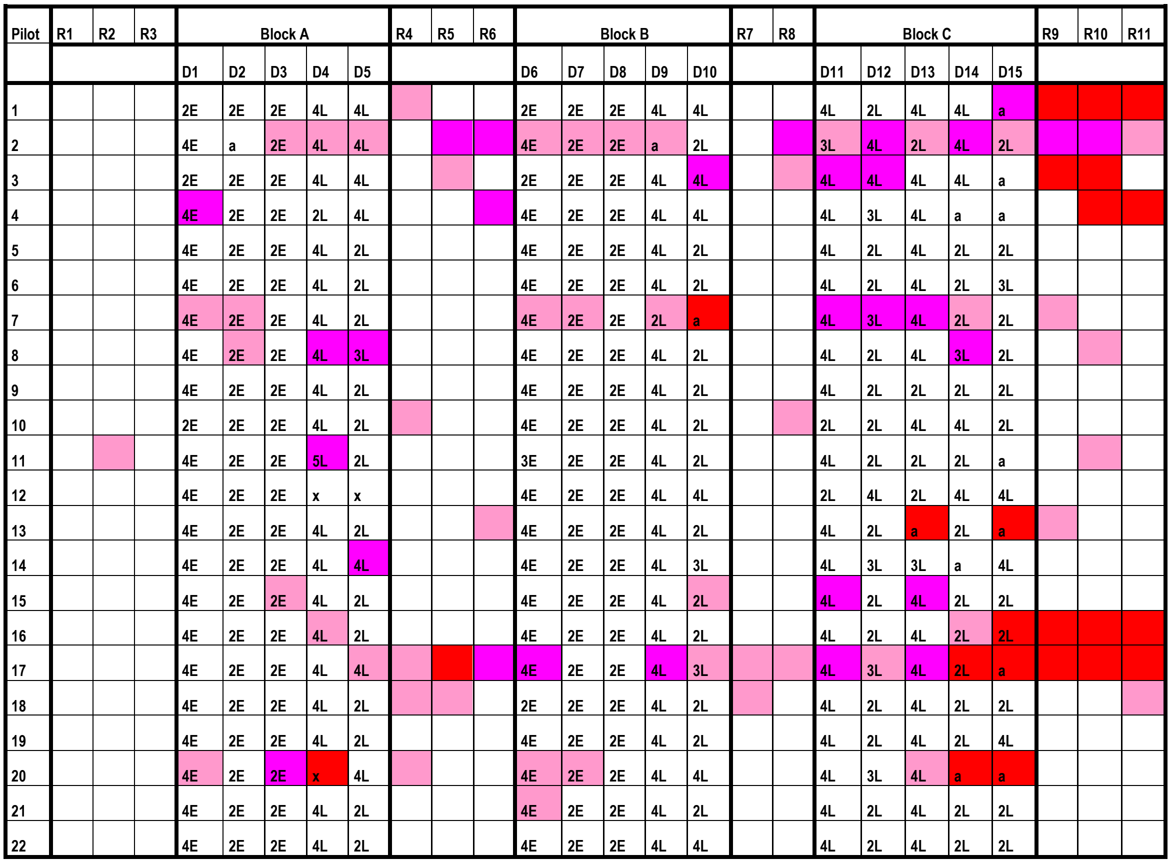
🟪 - 2 or more scheduled measurements are missing

◻ - 1 scheduled measurement is missing

‘nE’ indicates a duty day with an early start with ‘n’ sectors; ‘nL’ indicates a duty day with a late finish with ‘n’ sectors; ‘x’ indicates a sick day, and; ‘a’ indicates absent for other reasons.

For the most part, this group of pilots took the PVT tests as requested; 3 times during each Rest Day, during each flight at top of descent, and following the final flight of each day. Most of the missing data in Table 5 indicate that the pilot failed to take one measurement on that day. However, frequently, a pilot was unable to take the test during flight due to the short time of a sector. Some pilots had PVT-equipment problems or battery failures. Almost all of the Mean PVT Response Times and Lapses for Pilot #2 were so large that we suspect he either had a problem with his PDA equipment or he misunderstood the necessity to minimize lapses while taking the test. We deleted the PVT data for Pilot #2 from any analysis entailing PVT data. The quantity and quality of the PVT data collected by Pilot #17 also appeared questionable, but we chose not to delete these data from analyses. The quality and quantity of the PVT data collected should be sufficient to enable us to conduct useful and valid analyses.

Table 5 - PVT data recorded for each pilot during each day of study



Pilot	R1	R2	R3	Block A					R4	R5	R6	Block B					R7	R8	Block C					R9	R10	R11
				D1	D2	D3	D4	D5				D6	D7	D8	D9	D10			D11	D12	D13	D14	D15			
1				2E	2E	2E	4L	4L				2E	2E	2E	4L	4L			4L	2L	4L	4L	a			
2				4E	a	2E	4L	4L				4E	2E	2E	a	2L			3L	4L	2L	4L	2L			
3				2E	2E	2E	4L	4L				2E	2E	2E	4L	4L			4L	4L	4L	4L	a			
4				4E	2E	2E	2L	4L				4E	2E	2E	4L	4L			4L	3L	4L	a	a			
5				4E	2E	2E	4L	2L				4E	2E	2E	4L	2L			4L	2L	4L	2L	2L			
6				4E	2E	2E	4L	2L				4E	2E	2E	4L	2L			4L	2L	4L	2L	3L			
7				4E	2E	2E	4L	2L				4E	2E	2E	2L	a			4L	3L	4L	2L	2L			
8				4E	2E	2E	4L	3L				4E	2E	2E	4L	2L			4L	2L	4L	3L	2L			
9				4E	2E	2E	4L	2L				4E	2E	2E	4L	2L			4L	2L	2L	2L	2L			
10				2E	2E	2E	4L	2L				4E	2E	2E	4L	2L			2L	2L	4L	4L	2L			
11				4E	2E	2E	5L	2L				3E	2E	2E	4L	2L			4L	2L	2L	2L	a			
12				4E	2E	2E	x	x				4E	2E	2E	4L	4L			2L	4L	2L	4L	4L			
13				4E	2E	2E	4L	2L				4E	2E	2E	4L	2L			4L	2L	a	2L	a			
14				4E	2E	2E	4L	4L				4E	2E	2E	4L	3L			4L	3L	3L	a	4L			
15				4E	2E	2E	4L	2L				4E	2E	2E	4L	2L			4L	2L	4L	2L	2L			
16				4E	2E	2E	4L	2L				4E	2E	2E	4L	2L			4L	2L	4L	2L	2L			
17				4E	2E	2E	4L	4L				4E	2E	2E	4L	3L			4L	3L	4L	2L	a			
18				4E	2E	2E	4L	2L				2E	2E	2E	4L	2L			4L	2L	4L	2L	2L			
19				4E	2E	2E	4L	2L				4E	2E	2E	4L	2L			4L	2L	4L	2L	4L			
20				4E	2E	2E	x	4L				4E	2E	2E	4L	4L			4L	3L	4L	a	a			
21				4E	2E	2E	4L	2L				4E	2E	2E	4L	2L			4L	2L	4L	2L	2L			
22				4E	2E	2E	4L	2L				4E	2E	2E	4L	4L			4L	2L	4L	2L	2L			

The Mean PVT Response Times recorded during tests taken on Rest Days, shown in Figure 5, are the average values for the group of 21 pilots (excluding Pilot #2) in the morning, afternoon, and evening for each of the Rest Days. The first block of 3 days prior to first flights are indicated as R1, R2, and R3, the next block is R4, R5, and R6, next is a 2-day block, R7 and R8, and the final 3 recovery days after last flight are the block denoted as R9, R10, and R11. Figure 5 shows the expected rise from morning through afternoon to evening each day and a small, but statistically significant increase in Mean PVT Response Times for the group from the first Rest Day through the final block of rest days in the FRV schedule. The large standard deviations are indicators of the differences among the pilots due to differences in life styles and sleep patterns.

Figure 6 shows the average Mean PVT Response Times recorded during In-flight and Post-flight tests taken on Duty Days for the group of 21 of the pilots. The Mean PVT Response Times for the group did not rise to disturbing levels (i.e., > 350 msec.) at any time during the Rest or Duty Days. The variation in the Mean PVT Response Times across all the Duty Days as well as the standard deviations on the measurements are not nearly as large as they were in the measurements made during Rest Days. Nevertheless, the individual differences indicated by the standard deviations in the Mean PVT Response Times during Duty Days are significant.

A tentative explanation of the indication that there were no extreme variations or values in the Mean PVT Response Times recorded by the group of 21 pilots across the entire roster is that professionals have coping strategies for overcoming low levels of fatigue to deal with occasionally challenging schedules. (See Hart & Wickens 1990) Also, Loh, et al 2004 argue that motivation frequently masks decrement in reaction times on the shorter duration PVT such as in the 5-minute PVT used in this study.

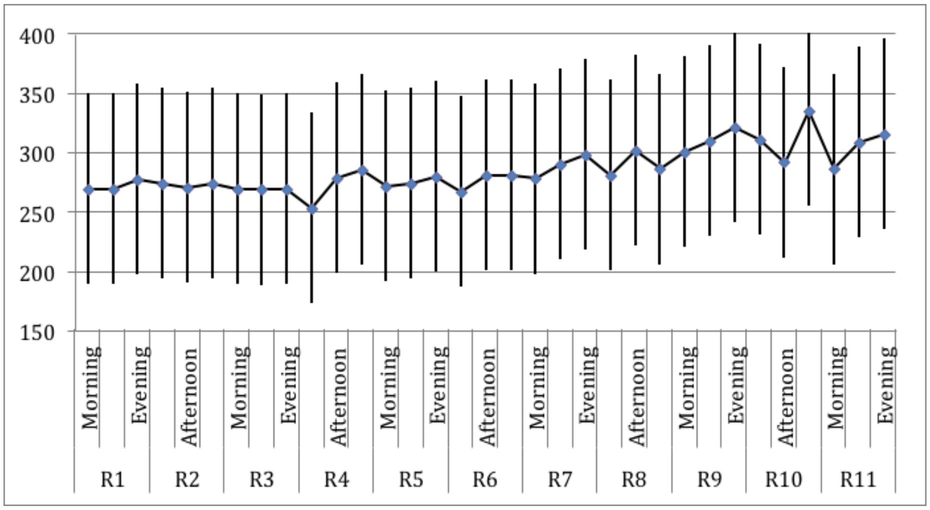


Figure 5- Mean PVT Response Time (in msec) during Rest Days for the Group

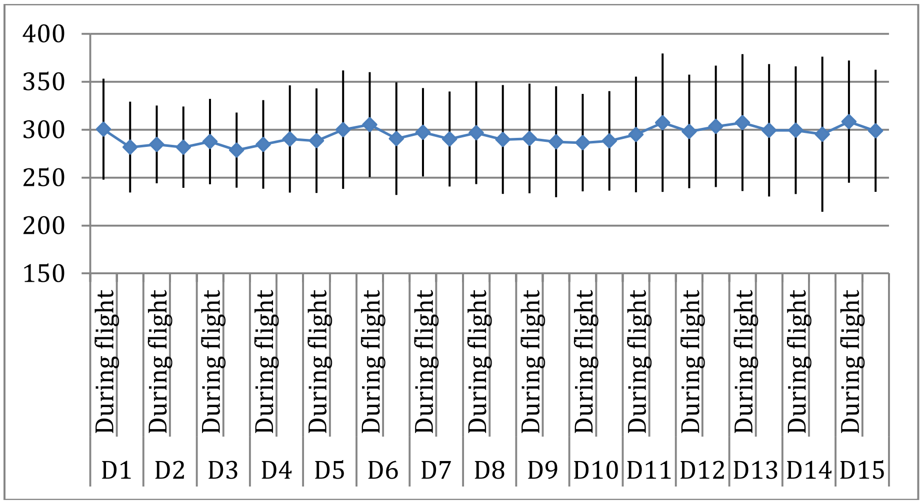


Figure 6- Mean PVT Response Time (in msec) during Duty Days for the Group

PVT Lapses

It has been argued that the best measures of degraded vigilance performance are the lapses during the second 5 minutes of a 10-minute PVT. Tucker et al 2009 compares the first 5 minutes with the second 5 minutes of PVT and shows a difference on the sleep-deprived subjects. Dorian et al 2005 suggest that performance is influenced by time-on-task. Effects of partial sleep deprivation seem to be sensitive to the duration of the task; therefore it is possible that the effects do not show on short tasks but rather on longer tasks and Loh et al 2004 suggests that, while the 5-min PVT may provide a viable alternative to the 10-min PVT for some performance metrics, subjects need to complete the full 10-minute task in order to effectively test for performance degradation.

It was not practical for pilots to take 10-minute PVT's during the short block times of most of the flights in this study. Nevertheless, we examined the lapses recorded in the PVT data that were collected to see whether they might provide any useful information to supplement the mean response times. Descriptive analyses of PVT Lapses for the group of 21 pilots showed a mean M = 2.88 lapses, SD = 1.83. Descriptive statistics for each pilot (excluding Pilot #2) are presented in Table 6 and illustrated in Figure 7. The red line in Figure 7 indicates the mean number of lapses for the group.

Table 6 - Descriptive statistics for PVT lapses for each pilot

Pilot	M	SD	Total Number of Lapses
1	1.18	0.24	27
3	2.21	0.9	53
4	1.28	0.37	31
5	1.33	0.41	34
6	3.21	1.46	84
7	5.63	1.38	141
8	4.7	1.73	122
9	1.19	0.3	31
10	4.24	1.7	110
11	5.12	1.68	128
12	1.11	0.19	29
13	5.11	1.86	123
14	3.36	1.41	87
15	2.92	1.08	76
16	2.27	0.66	50
17	2.43	0.97	49
18	1.7	0.6	44
19	2.69	0.94	70
20	4.84	1.14	111
21	1.25	0.3	33
22	2.76	0.73	72

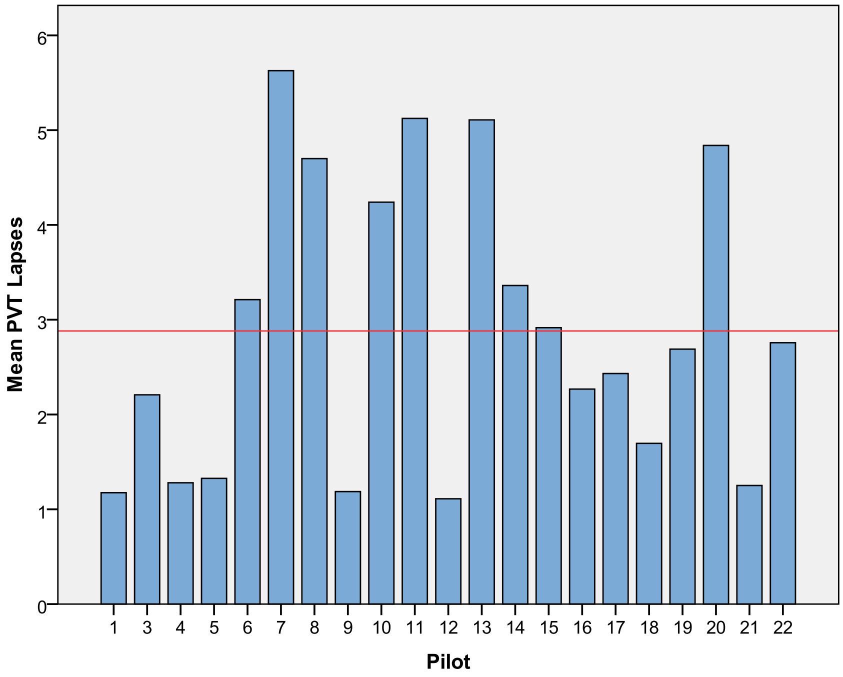


Figure 7- Mean PVT lapses for each pilot

We wanted to determine whether there were significant differences in PVT lapses on rest days, early start days, and late finish days for the pilots as a group and for each pilot. The normality test showed that the lapses were not normally distributed for the group or for most of the individual pilots. To correct the problem with normality, the PVT data were logarithmically transformed. Univariate ANOVA conducted on the normalized PVT lapses indicated that there was a main effect of Rest Days, early-start Duty Days, and late-finish Duty Days, F(2, 521) = 6.98, p= .001 in PVT lapses. Post-hoc comparisons revealed that there were significant differences in PVT lapses between early start Duty Days (M=.56, SD=.18) and Rest Days (M=.51, DS=.19) and between late finish days (M=.58, SD=.19) and rest days, but not between early start and late finish.

The high number of lapses shown for some of the pilots in Table 6 can be due to lack of conformity with the requirements of the test (i.e., push the button as fast as possible) and interruptions during the test, especially when taken in flight (as mentioned by pilots several times in their workbooks).

Univariate ANOVA on PVT lapses comparing Rest Days with Blocks A, B, and C of Duty Days found a significant main effect of Block, F(3, 521)=4.59, p< .001. Post-hoc comparisons showed differences between Blocks B (M=.56, SD=.18) and rest days (M=.51, SD=.19), between Block C (M=.59, SD=.22) and rest days. There was no significant difference between the blocks of Duty Days.

In general pilots made more lapses during Duty Days than during rest days. There was no difference in the number of lapses when comparing early-start with late-finish Duty Days. Although there was an increased level of fatigue during the day it was not reflected in the number of PVT lapses.

The study of the PVT lapses added little to our understanding of the degradation of pilot performance. We relied primarily on analyses of the Mean PVT Response Times for this purpose.

The evidence in our analysis of the Mean PVT Response Times indicated that the pilots experienced degradation of performance very infrequently in flight.

Mean PVT Response Times ≥ 500 msec. are termed performance lapses, or lapses in attention. We chose to use a conservative value of 400 msec. as the boundary between acceptable performance and impaired performance. (See Basner and Dinges 2011 for justification of this decision.)

Of the total number of Mean PVT Response Times recorded on Duty Days (1103 times including both In-flight and Post-flight tests), only 4.53% (50 out of 1103) were greater than 400 msec. and about 1/3rd of these were recorded at Post-flight (end of the day). In-flight, the pilots recorded Mean PVT Response Times greater than 400 msec. for only 4.16% of the flights (34 out of 818). This group of pilots experienced degradation of performance in flight infrequently during the Duty Days of the roster used in this study and even then it was mostly at only moderate levels of degradation.

Only 5.61% (16 out of 285) of Mean PVT Response Times recorded during Post-flight (end of the day) were greater than 400 msec. even though over half (53.03%) of the pilots felt (based on their Post-flight Samn-Perelli scores) at least moderately tired by the end of their Duty Day. Most often, the Mean PVT Response Times recorded at Post-flight was better than the Samn-Perelli score recorded at the same time would imply.

As noted above, 4.53% of the Mean PVT Response Times collected during Duty Days were greater than 400 msec. This is almost the same as the percentage, 4.14%, (26 out of 628), of the Mean PVT Response Times collected during Rest Days that were greater than 400 msec. Based on their Mean PVT response Times, this group of pilots performed as capably during their Duty Days as they did during their Rest days.

On Rest Days, 2.86% (6 out of 210) of the Mean PVT Response Times recorded in the Morning were greater than 400 msec., 4.26% (9 out of 211) were greater than 400 msec. in the Afternoon, and 5.31% (11 out of 207) of the Mean PVT Response Times recorded in the Evening (1 to 2 hours before retiring) were greater than 400 msec. (When compared with the same analysis of Samn-Perelli scores reported later, this is additional evidence that the typical diurnal effect on the self-assessment of sleepiness does not seem to have nearly the same level of effect on performance of tasks as measured by the PVT.)

An important conclusion of our analyses of the Mean PVT Response Times is that this group of pilots experienced degradation of performance very infrequently in flight during the Duty Days as flown during this study and even then it was mostly at only moderate levels of degradation.

Level of Fatigue

Samn-Perelli Scores of sleepiness

We performed analyses of the Samn-Perelli Scores for the group and for each pilot. Again, as in the case of the PVT tests, the quantity and quality of the Samn-Perelli data collected during this study were far better than those collected in the previous two studies. Table 7, which indicates missing data in the recordings of Samn-Perelli scores, is similar to Table 5 for PVT data. As in Table 5, the days highlighted in color indicate the extent of incomplete data on each day as follows:

- 🟥 - no measurements were taken
- 🟪 - 2 or more scheduled measurements are missing
- ▨ - 1 scheduled measurement is missing

'nE' is a duty day of 'n' sectors with early start; 'nL' is a duty day of 'n' sectors with late finish; 'x' indicates a sick day, and; 'a' indicates absent for other reasons. In general, these pilots were very conscientious about recording their Samn-Perelli Scores. Pilot #17 once again had the worst record of number of measurements taken.

Table 7 – Samn-Perelli Scores recorded for each pilot during each day of study

Pilot	R1	R2	R3	Block A					R4	R5	R6	Block B					R7	R8	Block C					R9	R10	R11
				D1	D2	D3	D4	D5				D6	D7	D8	D9	D10			D11	D12	D13	D14	D15			
1				2E	2E	2E	4L	4L	▨			2E	2E	2E	4L	4L			4L	2L	4L	4L	a			
2	▨			4E	a	2E	4L	4L		🟥	🟥	▨ 4E	2E	▨ 2E	a	2L		🟪	▨ 3L	4L	2L	🟪 4L	2L	🟥	🟥	
3				2E	2E	2E	4L	4L				2E	2E	2E	4L	4L			4L	▨ 4L	4L	4L	🟥 a	🟥		
4				4E	2E	2E	2L	4L				4E	2E	2E	4L	4L			4L	3L	4L	a	a		🟥	🟥
5				4E	2E	2E	4L	2L	▨	▨		4E	2E	2E	4L	2L			4L	2L	4L	2L	2L			
6				4E	2E	2E	4L	2L		🟥	🟥	4E	2E	2E	4L	2L			4L	2L	4L	2L	3L			
7				▨ 4E	2E	2E	4L	2L				4E	2E	2E	2L	🟥 a			4L	3L	4L	2L	2L			
8				4E	2E	2E	4L	3L				4E	2E	2E	4L	2L			4L	2L	4L	▨ 3L	2L			
9				4E	2E	2E	4L	2L				4E	2E	2E	4L	2L			4L	2L	2L	2L	2L			
10				2E	2E	2E	4L	2L				4E	2E	2E	4L	2L		▨	2L	2L	4L	4L	2L			
11		▨		4E	2E	2E	🟪 5L	2L				3E	2E	2E	4L	2L	▨		4L	2L	▨ 2L	2L	🟥 a		🟪	
12				4E	2E	2E	x	🟥 x				4E	2E	2E	4L	4L			2L	4L	2L	4L	4L			
13				4E	2E	2E	4L	2L		▨	▨	4E	2E	2E	4L	2L	🟪	▨	4L	2L	🟥 a	2L	🟥 a			🟪
14				4E	2E	2E	4L	4L				4E	2E	2E	4L	3L			4L	3L	3L	a	4L			
15				4E	2E	2E	4L	2L				4E	2E	2E	4L	2L			4L	2L	4L	2L	2L	▨		
16	🟪			4E	2E	2E	4L	2L				4E	2E	2E	4L	2L			4L	2L	4L	2L	2L		🟪	
17				4E	2E	2E	4L	4L		🟥	🟥	4E	2E	2E	🟪 4L	3L	🟥	🟥	🟪 4L	3L	🟪 4L	🟥 2L	🟥 a	🟥	🟥	🟥
18				4E	2E	2E	4L	2L				2E	2E	2E	4L	2L			4L	2L	4L	2L	2L			▨
19				4E	2E	2E	4L	2L				4E	2E	2E	4L	2L			4L	2L	4L	2L	4L			
20				4E	2E	2E	🟥 x	4L				4E	2E	2E	4L	4L			4L	▨ 3L	4L	🟥 a	🟥 a			
21				4E	2E	2E	4L	2L				4E	2E	▨ 2E	4L	2L			4L	2L	4L	2L	2L			
22				4E	2E	2E	4L	2L				4E	2E	2E	4L	4L			4L	2L	4L	2L	2L			

Figures 8 and 9 show the average Samn-Perelli Scores for the group of 22 pilots recorded during tests taken on Rest Days and In-flight and Post-flight on Duty Days, respectively.

The mean values of Samn-Perelli Scores for the group do not rise to disturbing levels (i.e., > 5) at any time and were not significantly different across the Rest or Duty Days. The expected increase in the subjective fatigue ratings across the time of day (from morning to evening on Rest Days and from pre-flight to post-flight on Duty Days) is apparent in Figures 8 and 9 of average Samn-Perelli Scores for the group. This manifestation of the effect of circadian rhythm was not so evident in the Mean PVT

Response Times for the group that were shown in Figures 5 and 6. The reason for the difference in the effects of the circadian cycle on self-assessment of fatigue and on performance will be discussed later when we present the results of our studies of correlation between Mean PVT Response Times and Samn-Perelli Scores.

Again, as in the case of the Mean PVT Response Times, the large individual variability in the Samn-Perelli Scores among the 22 pilots was indicated by the large standard deviations in the group data plotted in Figures 8 and 9. The differences among pilots are more significant than might be expected by chance.

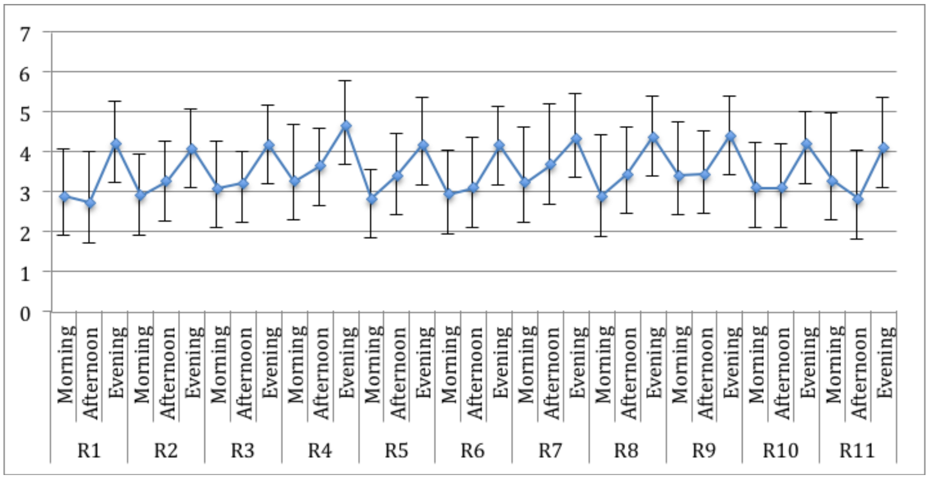


Figure 8 – Average Samn-Perelli scores during Rest Days for the Group

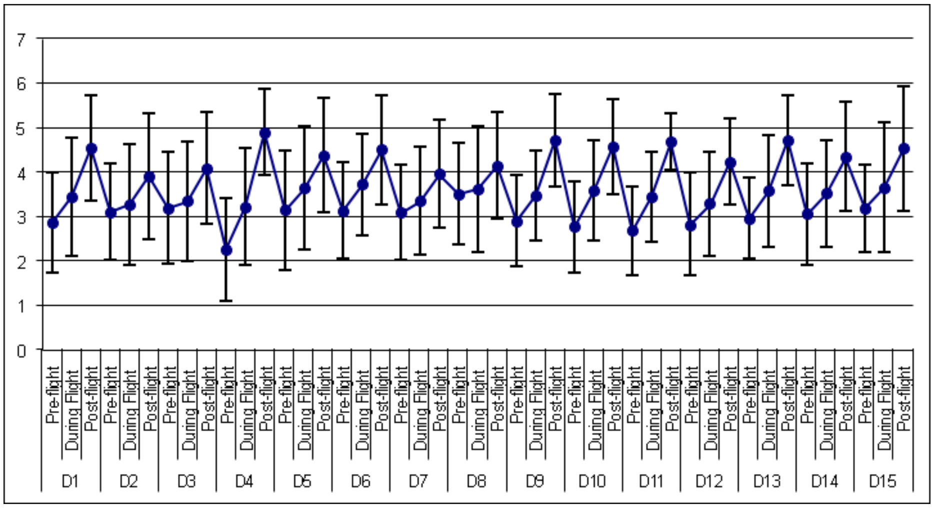


Figure 9 – Average Samn-Perelli scores during Duty Days for the Group

In this study, the participants used the following, currently used, brief descriptions of the Samn-Perelli scale for scoring their levels of fatigue:

Degree of Fatigue	**Scale Rating**
Fully alert, wide awake	1
Very lively, responsive, but not at peak	2
Okay, somewhat fresh	3
A little tired, less than fresh	4
Moderately tired, let down	5

Extremely tired, very difficult to concentrate	6
Completely exhausted, unable to function effectively	7

However, the original definitions taken from Samn and Perelli 1982 are more precise and useful to our discussion of the data.

1 - Unusually wide-awake. Possible performance enhancement.
2 - Very alert, wide-awake. No performance impairment due to fatigue.
3 - Normal level of alertness, typically well rested. No performance Impairment due to fatigue,
4 - Mild fatigue perceived. Performance impairment possible but not a significant factor.
5 - Moderate fatigue. Performance impairment possible. Flying duty permissible but not recommended unless urgent.
6 - Severe fatigue. Performance impairment probable. Flying duty not recommended.
7 - Severe fatigue. Performance definitely impaired. Flying duty not recommended. Safety of flight in jeopardy.

According to Samn and Perelli 1982, a score of 4 *("mild fatigue"*) or less can be interpreted to mean that there is no significant impairment of performance due to fatigue. However, even a score of 5 (*"moderate fatigue"*) represents a level at which performance impairment is only just possible and flight duty is permissible. Scores above 5 represent levels that have high probability of producing significant impairment in performance (i.e., high value of Mean PVT Response Time). A Samn-Perelli score of 5 is often used as the borderline value between acceptable and unacceptable levels of fatigue. As noted in CAA 2005, the concern that arises when average levels of fatigue rise to level 5 on the Samn-Perelli scale is not that the mean level itself is unacceptably high, but that the probability of an extremely high level in a subgroup of individuals is much increased. Moreover, the use of 5 as a "critical" Samn-Perelli score is based on the average for the general population, not even just on the average for the population of commercial airline pilots. In much of the following discussion that deals with group statistics, we have chosen to use a conservative score of 4 as the boundary between acceptable and unacceptable levels of sleepiness. However, when analyzing the data on each individual, a level of 5 is considered acceptable.

Of the total number of Samn-Perelli Scores recorded on Duty Days (1526 values including Pre-flight, In-flight, and Post-flight scores), about ¼ (382/1526) were greater than 4.

Only 7.0% (22 out of 314) of the Samn-Perelli Scores recorded during Pre-flight were greater than 4, indicating that most of the pilots felt alert at the start of their Duty Days.

For 21.5% of the flights (194 out of 899), the pilots recorded Samn-Perelli Scores greater than 4. The majority of these scores, 19.1%, were 5; only 2.4 % of these were 6; and none were at 7. These results indicate that this group of pilots experienced levels of sleepiness greater than a moderate level (at which *"flying duty is permissible")* very infrequently (during 2.4% of the flights) while in flight during this study.

The <u>Post</u>-flight (i.e., end of the day) Samn-Perelli Scores account for almost half (166 out of 382) of all the scores greater than 4 recorded during the Duty Days. Also, of all of the Samn-Perelli Scores recorded during Post-flight, 53.03% (166 out of 313) were greater than 4, indicating that about half of the pilots felt at least moderately tired by the end of their Duty Day. At least some of this sense of tiredness on Duty Days of late arrivals can be attributed to the normal diurnal effect of circadian rhythm, which we see on Rest Days as well.

As noted above, 25.0% of the Samn-Perelli Scores collected during Duty Days were over a value of 4. However, in comparison, a substantially higher percentage, 29.5%, (150/508), of the Samn-Perelli Scores collected during Rest Days were greater than 4. While it appears that the pilots are not using their Rest Days adequately, this may also be interpreted to mean that the flying schedules of the Duty Days in the rosters flown in this study were not inducing levels of fatigue that were worse than those experienced during Rest Days.

On Rest Days, 13.8% (31 out of 224) of the Samn-Perelli Scores recorded in the Morning were greater than 4; 11.06% (24 out of 217) were greater than 4 in the Afternoon; and 39.07% (84 out of 215) of the Samn-Perelli Scores recorded in the Evening were greater than 4. This effect on Rest Days is representative of the typical diurnal effect of the circadian cycle on the self-assessment of sleepiness.

These statistical analyses of the group data indicate that the participating pilots experienced more than just mild levels of sleepiness on fewer than 3% of the flights they flew during the study. A question that we will investigate further in this report is whether any of these 3% of the flights entailed levels of sleepiness that caused degraded performance.

In general, very good days in which the Samn-Perelli Scores tended to be very low were the 2nd or 3rd Duty Days after a block of Rest Days that entailed departure times between 6:00 and 10:00 AM and 4 to 6 duty hours. Very bad days in which Samn-Perelli Post-flight Scores were high occurred when the 3rd, 4th, or 5th sectors ended late in the day with departure times as late as 21:00 to 23:00 hours and as long as 13-hour duty days.

The evidence from our analyses of the data on Samn-Perelli Scores is that very few of the 22 pilots experienced significant levels of fatigue during any of their Duty Days in the roster they flew during this study.

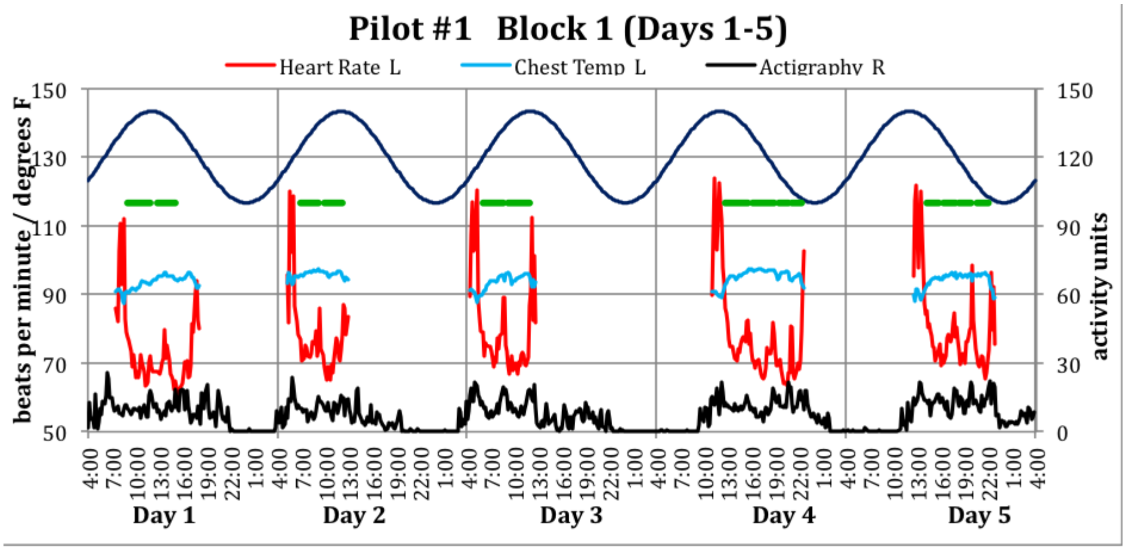
BioHarness Data

The BioHarness was used to collect physiological measures on 20 pilots during the 15 Duty Days of the FRV roster shown in Figure 1. Data were not collected during their Rest Days. The duration of the daily physiological data recordings ranged from 6 to 12 hours where longer recordings occurred during the late shifts. Data files were retrieved from the laptops and electronically sent to NASA researchers for processing and analyses.

Data files were imported to signal processing software (Dadisp, Inc.) consisting of customized routines for removal of noise artifacts (filtering and other smoothing methods) from physiological waveforms, peak detection algorithms for extracting waveform features of interest (e.g., r-peaks from ECG), and routines for calculating metrics (e.g., means and standard deviations over selected time periods). In addition, the software includes manual editing features for those cases where artifacts are missed by automated cleaning routines. Other routines include power spectral analyses for examining periodicity in heart rate variability. Processed data are then exported to Excel for data visualization that include graphing means of physiological metrics and examining relationship to specific study events (e.g., flight times and other study metrics such as PVT, Samn-Perelli, CIS, ESS, M/AS, TLX, sleep diaries, etc.).

Examples of Pilot Data:

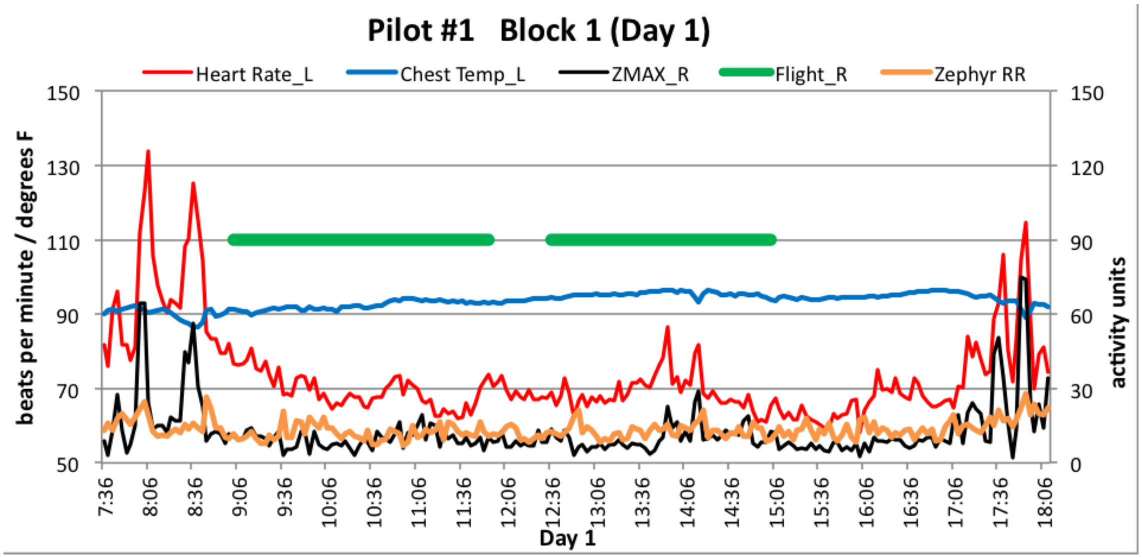
Figure 10 below shows 10-minute means of physiological measures of Pilot #1 during his first 5-Duty-Day cycle. The sine wave represents a repeated 24-hour circadian cycle beginning at 0400 hours. Actigraphy data (black) shows relative activity while awake (more active), and asleep (little or no activity). These data indicate longer sleep periods with later wake times on days 4 and 5 that correspond to the start of late shift duty periods (day 4: ~1030; day 5: ~1200). The green horizontal bars (segmented to show flights and turnarounds) indicate times when the flights occurred (2 on days 1-3; 4 on days 4-5). Increases in activity (possibly exercise) prior to the start of Pilot #1's work shift each day are consistent with observed changes in physiological measures (accelerated heart rate and lower chest temperature). Notable increases in heart rate can be seen with the activity during turnaround periods.



Note: Actigraphy is referenced to right axis; heart rate is referenced to the left axis.

Figure 10 – Activity, heart rate and chest temperature of Pilot #1 on Duty Days 1-5

Figure 11 shows greater detail of Pilot #1's physiological data plotted as 3-minute means over his first day on duty. Large increases in activity (ZMAX_R) were observed at (0800) and (1730) which were consistent with increases in heart rate and respiration rate and decreases in chest temperature. These physiological changes can normally occur in individuals who exercise, go jogging, or bicycle riding. There was a notable increase in heart rate and activity occurring at ~1400 during his second flight; however, there were no indications from his workbook that might account for these changes.



Note: Activity (ZMAX-R) data is referenced to right axis; Heart Rate_L is referenced to left axis.

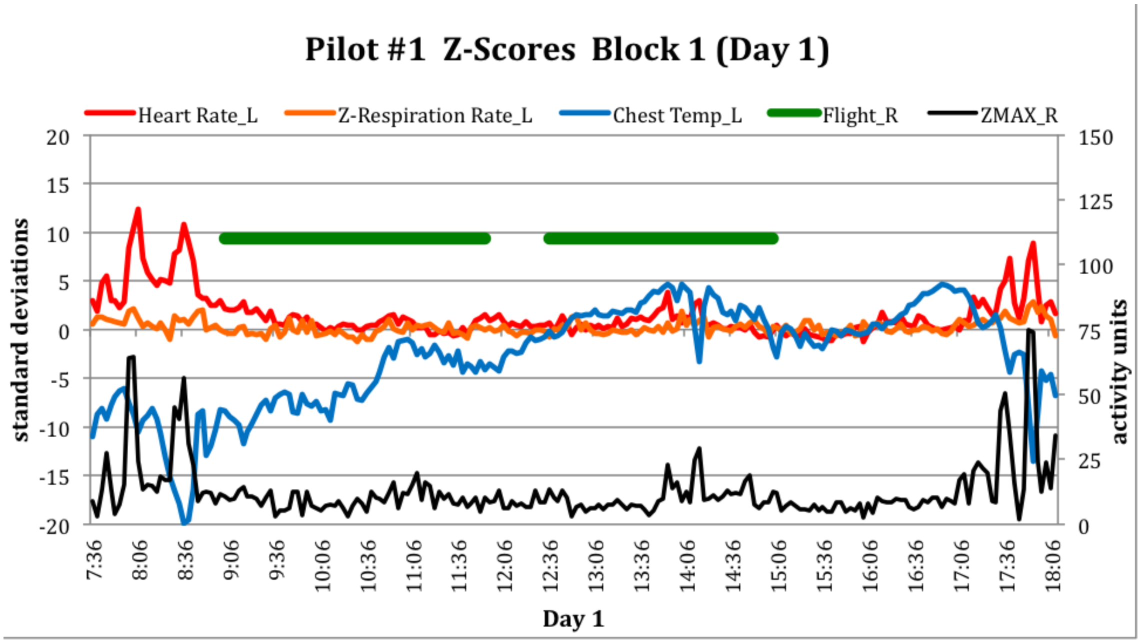
Figure 11 – Activity, heart rate and chest temperature of Pilot #1 on his first flight day

Z-score transformation is a useful statistical method for representing several different measures on the same scale. A Z-score quantifies the original score in terms of the number of standard deviations that that score is from the mean of the distribution. The formula for converting from an original or "raw" score to a Z-score is:

$$Z = (\text{score} - \text{mean}) / \text{standard deviation}$$

Means and standard deviations were calculated from baseline measures taken for each physiological variable. Baseline was defined as a time period when an individual is awake and relaxed but not exercising. The baseline for Pilot #1 was selected as 60-minute period between flights on Day 1. For Pilot #1, Z-scores were computed as the difference of measures (3-minute means) taken during the entire work shift and the baseline period (overall mean and standard deviation).

Figure 12 shows the '**physiological response profile**' of Pilot #1 where data are expressed as Z-scores. A physiological response profile shows the magnitude of each physiological response from its optimal resting level (baseline) to a known study event (e.g., PVT performance), showing covariance between responses, and the time course for their return to baseline levels. These profiles are highly idiosyncratic but are normally quite stable within individuals. Figure 12 indicates that the magnitude of chest-temperature decreases from baseline during flight 2 was greater than 5 standard deviations, but these may have been due to changes in ambient air temperature in the cockpit.



Note: Activity (ZMAX_R) is data referenced to right axis.

Figure 12 – Z-scores of heart rate, respiration rate, and chest temperature (left axis) of Pilot #1 on his first flight day

Heart rate variability (HRV) spectral analyses will be performed on heart rate data derived from ECG. These metrics will include spectral power in very low frequency (VLF, $\leq$ 0.004 Hz), low frequency (LF, 0.04-0.15 Hz), and high frequency (HF = 0.15-0.40 Hz) bands, and the ratio of LF to HF power, which is an index of autonomic balance between sympathetic and parasympathetic nervous system activity.

Correlation analyses will be performed to determine the strength of the linear relationship between Z-score transformations of BioHarness indicators of fatigue and the self-assessments of fatigue (Samn-Perelli Scores) and PVT data.

Some Early Indications:

Heart rate variability increases in different frequency bands as pilots become more fatigued. These increases are seen as large variations in beat-to-beat heart rates possibly due to irregular beats (missed or extra beats). Unexpected increases in VLF power (sympathetic dominance) and decreases in HF power (parasympathetic or vagal dominance) were seen in Pilot #1 when he reported increases in fatigue on both Samn-Perelli Scores and Mood/Alertness Scales. However, these sympathetic responses were associated with improved PVT scores (median RT) at the end of the long-late shifts. These improved PVT scores may reflect a "practice effect" or simply "last test phenomena" where participants tend to perform well on the very last test of a day. A further examination of data from other pilots is needed. (Further evidence that heart rate is a reliable metric of fatigue is presented in Chua et al 2012.)

Factors of Fatigue

A personal profile of the proximate and the latent factors that could cause or contribute to fatigue (i.e., drowsiness) was developed from information based on the following measurements:

Latent

Demographic questionnaires
Morningness/Eveningness Scale (MES)
Checklist Individual Strength (CIS)
Epworth Sleepiness Scale (ESS)
Mood/Alertness Scale (M/AS)

Proximate

Sleep diary
Actigraphy
Work pattern (roster)
NASA Task Load Index (TLX)
Hassle Factors

Latent Factors

Demographic Characteristics of Participants

Table 8 presents a summary of the demographics of the participants. Twenty-two commercial aviation pilots (18 males and 4 females) volunteered to participate in the study during July and August 2011. The participants were between 22 and 46 years of age (M = 33, SD = 7.70). Twenty-one of these pilots were British and 1 was Dutch. They had commercial flight experience of from 250 to 12,000 hours (M = 3,650, SD = 3,795.89). Most of the data flights for this study were flown in A-319 or A-320 aircraft in which the pilots in this group had from 100 to 6,000 hours (M = 457.5, SD = 2,092.67) of experience on type. Two pilots (#1 and #3) flew only B-737 aircraft during this study.

During this study, all pilots returned to their home base each duty day and the commute times from their permanent residences ranged from 15 minutes to 2 hours (M = 37.5 minutes, SD = 25.66 minutes). Only Pilot #10 indicated in his sleep diary that some nights during duty he slept in a hotel.

Table 8 – Statistics for Demographic Characteristics

	N	**Minimum**	**Maximum**	**Mean**	**Std. Dev.**
Age	22	22	46	33.09	7.70
Permanent Address Commute Time in minutes	22	15	120	43.17	25.66
Temporary Address Commute Time in minutes	3	5	75	30.00	39.05
Total Commercial Hours	22	250	12000	4236.36	3795.89
Total Hours on Type	22	100	6000	1813.41	2092.67
Total easyJet Hours	22	365	8000	2265.68	2233.89

This group of pilots said their minimum sleep duration on non-duty days was between 4 and 9 hours (M = 6.95, SD = 1.24) and their maximum sleep time on non-duty days was between 8 and 14 hours (M = 10.05, SD = 1.56). These pilots said their minimum sleep duration on duty days was between 4 and 8 hours (M = 5.07, SD = 1.30) and maximum sleep time on duty days was from 6 to 12 hours (M = 8.70, SD = 1.34).

This small group of volunteers cannot be considered representative of the full population of easyJet pilots because of significant differences in the statistical characterizations of the two groups. Female pilots constituted 18% of the study group, whereas 5% of the full pilot group at easyJet is female. The median age of the participants in the study was 33, whereas the median age of the full easyJet pilot group is 37. The average commercial flight time experience of the study group was 4,236 hours with a standard deviation of 3,796 hours, whereas the average was 3,205 hours with a standard deviation of 1,877 hours for the full pilot group.

The fact that the study group was not representative of the full population may not be important to this study as we are focused on the individual performances and we are finding very large differences among the 22 subjects of the study in all of the factors related to fatigue and its effects on performance. These large and significant individual differences are likely to be found in the full pilot population as well.

Morningness/Eveningness Scale (MES)

The ratings of the 22 pilots on the MES ranged from 35 to 67 where

16-30 definite evening
31-41 moderate evening
42-58 intermediate
59-69 moderate morning
70-86 definite morning

On the basis of the scores they gave to the MES questions, 5 pilots were rated as Moderate Larks, 2 were rated as Moderate Owls, and 15 were rated as neither Owls nor Larks. However, there were 3 of the 15 (Pilots #4, #8, and #18) whose sleep habits indicated a tendency toward being Owls and another 3 (Pilots #6, #13, and #16) whose sleep habits indicated they have tendency toward being Larks. Furthermore, Pilots #8 and #18 show an obvious pattern of an Owl in their Actigraphy data and, based on their subjective “most alert” and “most tired” times from their demographic questionnaires, Pilots #4 and #8 are clearly Owls. The Actigraphy data shows Lark patterns for Pilots #13 and #16 and their subjective “most alert” and “most tired” times indicate that all three (Pilots #6, #13, and #16) are Larks.

Checklist Individual Strength (CIS)

Although it had been planned that the pilots would complete the CIS questionnaire at the beginning of the study and also during the recovery period at the end of the study, in fact, it was completed only at the beginning of the study, which greatly reduces the value of this self-assessment.

The scores assigned by each of the participants to the 20 questions in the CIS at the start of the study showed that this group experienced a Level of Fatigue between 12 and 52 (M = 27.36, SD = 10.99) out of a possible range of 8 to 56. Their Reduction in

Concentration was from 5 to 25 (M = 14.95, SD = 5.41) in a possible range of 5 to 35. Their Reduction in Motivation was rated between 7 and 20 (M = 11.77, SD = 3.64) in a possible range of 4 to 28 and their Reduction in Physical Activity was 3 to 18 (M = 8.82, SD = 3.97) in a possible range of 3 to 21. Total CIS scores ranged from 32 to 105 in a possible range of 20 to 140. Two pilots had total CIS scores of 102 (#5) and 105 (#16), which are disconcertingly high for people in this profession. The next highest score in this group was 90. The lowest CIS score of 32 was for a 38-year old female pilot (#2).

Epworth Sleepiness Scale (ESS)

The data on the ESS were only collected at the beginning of the study. On a scale from 0 (Would **never** doze) to 3 (**High** chance of dozing), these pilots showed the following responses to the 8 situations posed:

Sitting and Reading: 0 to 3 (M = 1.23, SD = 1.07)
Watching TV: 0 to 3 (M = 1.41, SD = 0.85)
Sitting inactive in a public place: 0 to 2 (M = 0.5; SD = 0.60)
As a passenger in a car for an hour without a break: 0 to 3 (M = 0.91, SD = 0.75)
Lying down to rest in the afternoon: 1 to 3 (M = 2.23, SD = 0.81)
Sitting and talking to someone: 0 to 1 (M = 0.18, SD = 0.39)
Sitting quietly after lunch without alcohol: 0 to 3 (M = 0.82, SD = 0.85)
In a car, while stopped for a few minutes in traffic: 0 to 2 (M = 0.27, SD = 0.55)

The total ESS scores ranged from 1 to 19 (M = 7.55, SD = 4.06). 17 of the 22 pilots had total ESS ratings less than 10 (i.e., in the Normal range), 3 were in the range of 10 to 12 (i.e., Borderline sleepy), and 2 pilots had ratings greater than 12 (i.e., considered Abnormally sleepy). A 40-year old pilot (#16) showed the highest total ESS rating of 19 and a 39-year old pilot (#13) had the other rating above 12 (13), but this does not appear to be age related because two 46-year old pilots (#6 and #10) had total ESS scores of 2 and 8. The lowest total ESS of 1 was for a 38-year old female pilot (#2) who also had the lowest total CIS score. The pilot (#16) with the highest total ESS score of 19 also had the highest CIS score of 105 reported above.

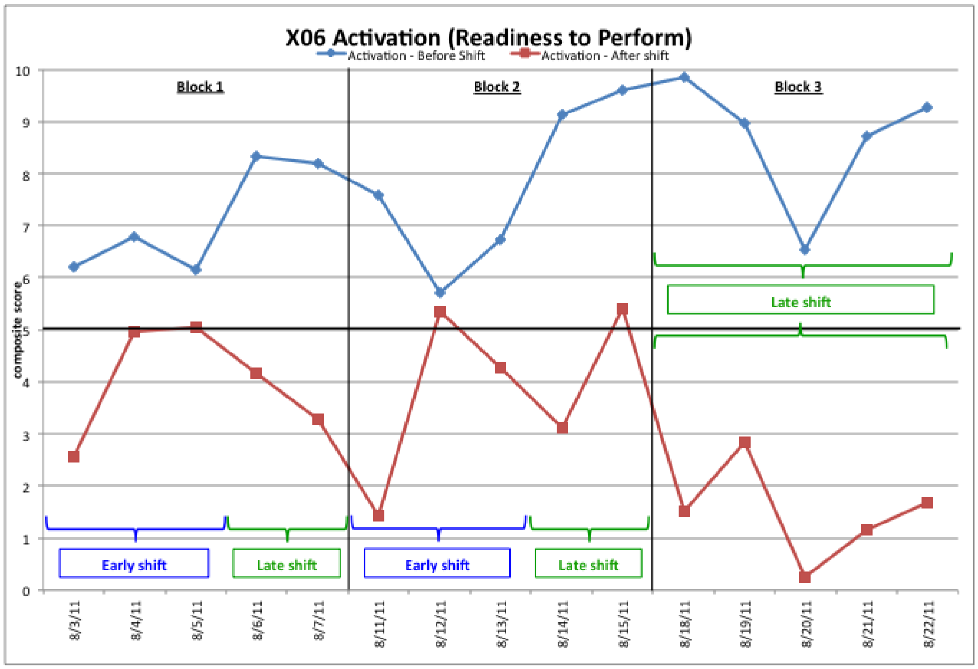
Mood/Alertness Scale (M/AS)

The Mood/Alertness Scale questionnaire was in the laptop computer used with the BioHarness. As the BioHarness was not worn on Rest Days, the M/AS data were also not collected on Rest days.

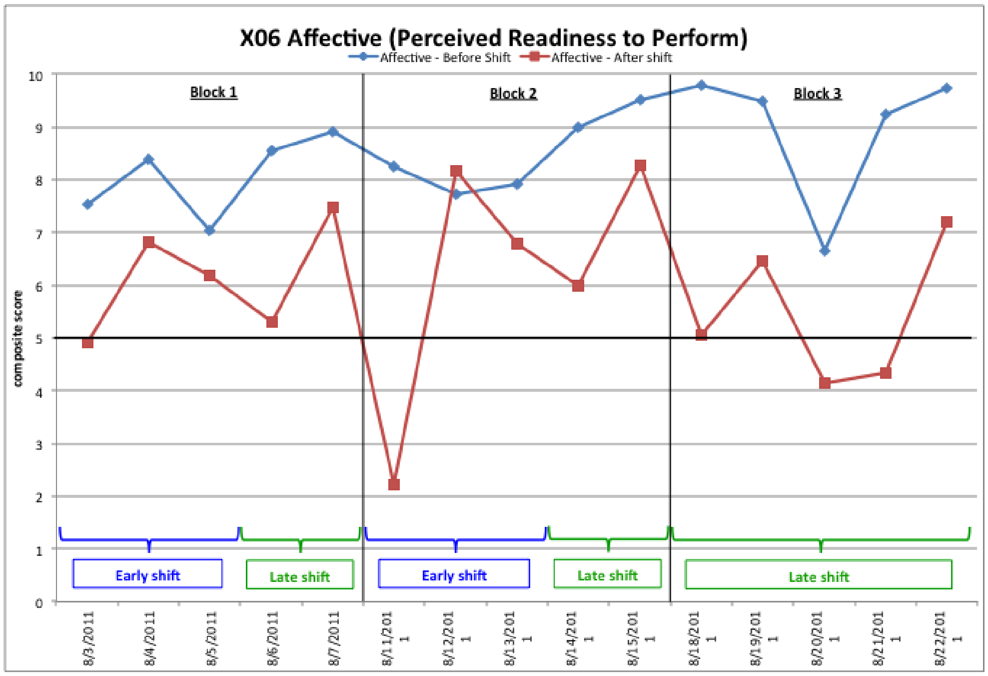
On the basis of the M/AS data, every one of the 20 pilots who completed this data set was significantly more ready to perform in the early hours (M=6.01, SD=2.10) than he or she was later in the day (M=3.78, SD=1.91). However, each pilot perceived himself or herself to be equally ready to perform in the PM hours (M=5.83, SD=2.04) as in the AM hours (M=6.31, SD=1.81). Among all the pilots, the self-perception of being ready to perform was higher than his or her readiness to perform and this was especially true in the evening hours.

These characteristics are exemplified in the data from one pilot (#6) shown in Figures 13.a and 13.b below. The figures show the high level of readiness (M=7.85, SD=1.40) and Pilot #6's perception of a high level of readiness (M=8.51, SD=0.99) in the early hours (Before Shift in blue). It shows the typical lower level of readiness (M=3.13, SD=1.69) and the perception of a high level of readiness (M=5.95, SD=1.65) in the evening hours (After Shift in red). These features are independent of whether the Duty Days entailed early or late shift flights.

The pilot (#16) who had the highest ESS and CIS scores also had M/AS data that differed significantly from those of the rest of the group. Consistent with his high ESS and CIS scores, Pilot #16 presented the lowest value in the group of readiness to perform (M=2.66, SD=0.90 in morning and M=1.64, SD=0.89 in evening) with mean values for both early and late hours that were more than 2 standard deviations lower than the mean values for the group. The data for this pilot (Figures 14.a, b) indicate that Pilot #16 perceived himself equally ready to perform in the early hours (M=5.88, SD=0.65) and evening hours (M=5.50, SD=0.66) despite his actual low readiness levels at either time.

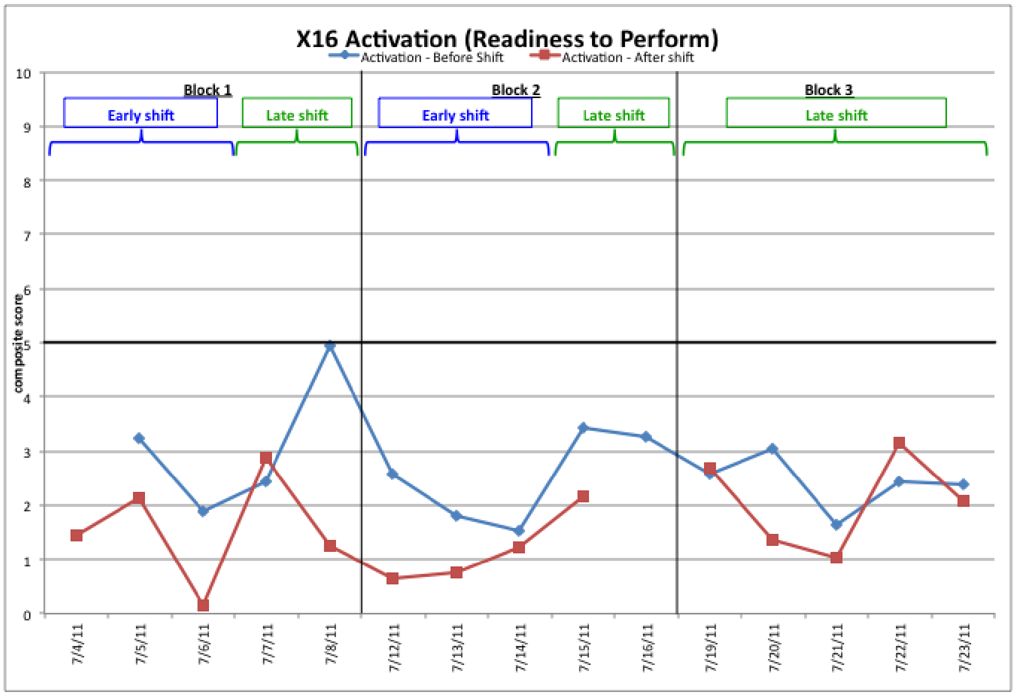


a. – Pilot #6's Readiness to Perform

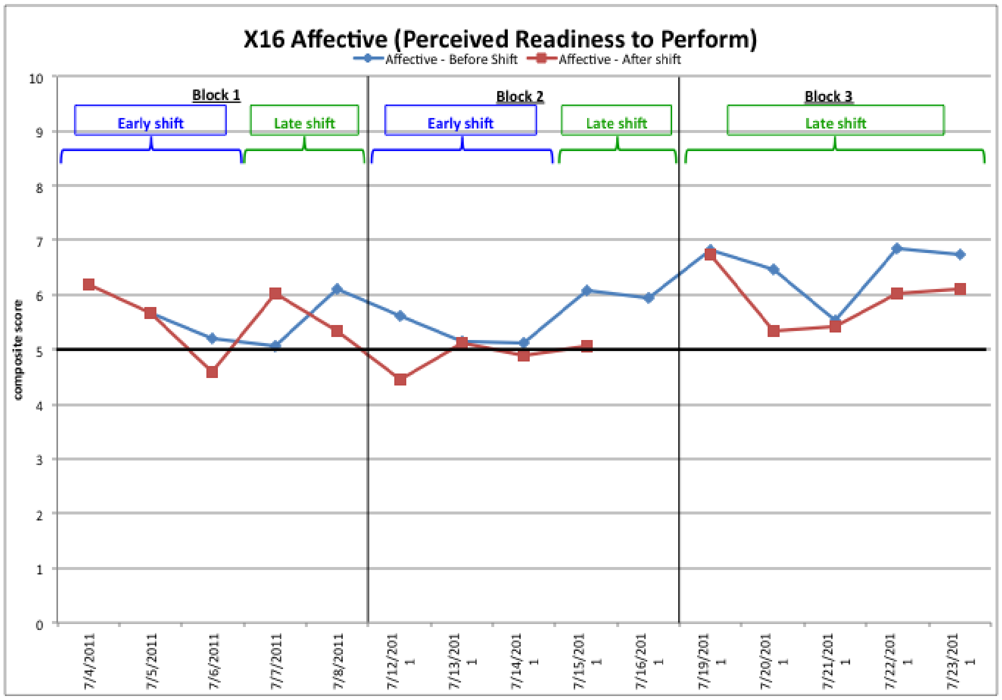


b. – Pilot #6's Perceived Readiness to Perform

Figure 13 – Readiness and Perceived Readiness to perform based on M/AS for Pilot #6



a. – Pilot #16's Readiness to Perform



b. – Pilot #16's Perceived Readiness to Perform

Figure 14 – Readiness and Perceived Readiness to perform based on M/AS for Pilot #16

Proximate Factors

Actigraphy

The activity monitor[3] provided objective data from which to infer sleep and wake activity. The activity monitors were worn during the period of participation (except when showering or swimming). For the purposes of the current study, the activity information was stored in 2-minute epochs as activity counts. The activity movements recorded by the activity monitor were downloaded from the monitor using Actiwatch communications and Sleep Analysis Software (Respironics, Inc.). A range of algorithms was then used to score wake and sleep. From the sleep-scoring functions of the Actiware software and from the bedtime and get-up time data recorded in the sleep diary, a range of sleep measures was determined. The following parameters are the most important in Actigraphy:

1. Time in Bed (TIB) - the time elapsed between the Start Time and the End Time of the given interval.
2. Total Sleep Time (TST) – the total number of epochs between the Start Time and the End Time of the given interval scored as SLEEP multiplied by the Epoch Length in minutes. For all analyses, the minutes were transformed into hours of sleep.
3. Sleep Efficiency – the percentage of Scored Total Sleep Time to Interval Duration minus Total Invalid Time (Sleep/Wake), for the given Rest Interval.

---

[3] Actiwatch 64, Mini Mitter Co., Inc., OR, USA

4. Wake After Sleep Onset (WASO) – the total number of epochs between the Start Time and the End Time of the given Sleep Interval scored as WAKE by the Actiware software multiplied by the Epoch Length in minutes.
5. Sleep Onset Latency – the time elapsed between the Start Time of a given rest interval and the following Sleep Start Time, in minutes.

The Actograms were examined to determine the missing data for each individual. Table 9 indicates the number of days when the pilots wore the Actiwatch. The days highlighted in purple indicate incomplete data. ‘nE’ indicates a duty day with an early start with ‘n’ sectors; ‘nL’ indicates a duty day with a late finish with ‘n’ sectors; ‘x’ indicates a sick day, and; ‘a’ indicates absent for other reasons.

Table 9 - Actigraphy data for each pilot during each day of study

Pilot	R1	R2	R3	Block A					R4	R5	R6	Block B					R7	R8	Block C					R9	R10	R11
				D1	D2	D3	D4	D5				D6	D7	D8	D9	D10			D11	D12	D13	D14	D15			
1				2E	2E	2E	4L	4L				2E	2E	2E	4L	4L			4L	2L	4L	4L	a			
2				4E	a	2E	4L	4L				4E	2E	2E	a	2L			3L	4L	2L	4L	2L			
3				2E	2E	2E	4L	4L				2E	2E	2E	4L	4L			4L	4L	4L	4L	a			▨
4				4E	2E	2E	2L	4L				4E	2E	2E	4L	4L			4L	3L	4L	a	a		▨	▨
5				4E	2E	2E	4L	2L				4E	2E	2E	4L	2L			4L	2L	4L	2L	2L			
6				4E	2E	2E	4L	2L				4E	2E	2E	4L	2L			4L	2L	4L	2L	3L			▨
7				4E	2E	2E	4L	2L				4E	2E	2E	2L	a			4L	3L	4L	2L	2L			
8				4E	2E	2E	4L	3L				4E	2E	2E	4L	2L			4L	2L	4L	3L	2L			▨
9				4E	2E	2E	4L	2L				4E	2E	2E	4L	2L			4L	2L	2L	2L	2L			
10				2E	2E	2E	4L	2L				4E	2E	2E	4L	2L			2L	2L	4L	4L	2L			
11	▨	▨	▨	4E	2E	2E	5L	2L	▨	▨	▨	3E	2E	2E	4L	2L	▨	▨	4L	2L	2L	2L	a	▨	▨	▨
12				4E	2E	2E	x	x				4E	2E	2E	4L	4L			2L	4L	2L	4L	4L			
13				4E	2E	2E	4L	2L				4E	2E	2E	4L	2L			4L	2L	a	2L	a			
14				4E	2E	2E	4L	4L				4E	2E	2E	4L	3L	▨		4L	3L	3L	a	4L			
15				4E	2E	2E	4L	2L				4E	2E	2E	4L	2L			4L	2L	4L	2L	2L			
16				4E	2E	2E	4L	2L				4E	2E	2E	4L	2L			4L	2L	4L	2L	2L			
17				4E	2E	2E	4L	4L				4E	2E	2E	4L	3L			4L	3L	4L	2L	a	▨	▨	▨
18				4E	2E	2E	4L	2L				2E	2E	2E	4L	2L			4L	2L	4L	2L	2L			
19				4E	2E	2E	4L	2L				4E	2E	2E	4L	2L			4L	2L	4L	2L	4L			
20				4E	2E	2E	x	4L				4E	2E	2E	4L	4L			4L	3L	4L	a	a			
21				4E	2E	2E	4L	2L				4E	2E	2E	4L	2L			4L	2L	4L	2L	2L			
22				4E	2E	2E	4L	2L				4E	2E	2E	4L	4L			4L	2L	4L	2L	2L			

Analyses were conducted to examine the first three parameters (TIB, TST, and Sleep Efficiency). The WASO and Sleep Onset latency will be analyzed in forthcoming analyses of the data as needed. The following results are based on analysis of the combined data for the group of 21 pilots. The data for Pilot #11 were excluded from these analyses because of a defective Actiwatch, which recorded multiple sets of data for him.

Overall TST and TIB obtained from Actiwatch and TST obtained from the Sleep Diary were strongly correlated for the whole group as shown in Table 10 (with further details in Appendix D) and for each individual (Table 11), except for Pilot #12. The TST data from sleep diary were not correlated with TIB and TST from Actiwatch for Pilot #12. After reviewing the comments section from the sleep diary for this pilot, we concluded that his records of times were questionable due to several stressful events he experienced throughout the study period and because of an illness of which he spoke.

Table 10 - Correlations between Actiwatch TST, TIB, and TST from sleep diary for the group of 21 pilots

	Time in Bed from Actiwatch	Total Sleep Time from Diary
Total Sleep Time from Actiwatch	.831**	.769**
Time in Bed from Actiwatch	-	.886**

**.Indicates that correlation is significant at the 0.01 level (2-tailed).

Table 11 - Correlations between Actiwatch TST, TIB, and TST from sleep diary for each pilot

Pilot ID		Time In Bed from Actiwatch	Total Sleep Time from Diary
1	Total Sleep Time from Actiwatch	.959**	.956**
	Time In Bed from Actiwatch	-	1.000**
2	Total Sleep Time from Actiwatch	.874**	.838**
	Time In Bed from Actiwatch	-	.847**
3	Total Sleep Time from Actiwatch	.795**	.898**
	Time In Bed from Actiwatch	-	.970**
4	Total Sleep Time from Actiwatch	.914**	.924**
	Time In Bed from Actiwatch	-	.916**
5	Total Sleep Time from Actiwatch	.921**	.884**
	Time In Bed from Actiwatch	-	.933**
6	Total Sleep Time from Actiwatch	.966**	.963**
	Time In Bed from Actiwatch	-	.993**
7	Total Sleep Time from Actiwatch	.856**	.875**
	Time In Bed from Actiwatch	-	.966**
8	Total Sleep Time from Actiwatch	.835**	.884**
	Time In Bed from Actiwatch	-	.969**
9	Total Sleep Time from Actiwatch	.913**	.867**
	Time In Bed from Actiwatch	-	.967**
10	Total Sleep Time from Actiwatch	.970**	.963**
	Time In Bed from Actiwatch	-	.953**
12	Total Sleep Time from Actiwatch	.777**	.068
	Time In Bed from Actiwatch	-	-.072
13	Total Sleep Time from Actiwatch	.588**	.850**
	Time In Bed from Actiwatch	-	.623**
14	Total Sleep Time from Actiwatch	.896**	.939**
	Time In Bed from Actiwatch	-	.856**
15	Total Sleep Time from Actiwatch	.896**	.808**
	Time In Bed from Actiwatch	-	.961**
16	Total Sleep Time from Actiwatch	.958**	.952**
	Time In Bed from Actiwatch	-	.972**
17	Total Sleep Time from Actiwatch	.955**	.794**
	Time In Bed from Actiwatch	-	.862**
18	Total Sleep Time from Actiwatch	.977**	.944**
	Time In Bed from Actiwatch	-	.968**
19	Total Sleep Time from Actiwatch	.896**	.878**
	Time In Bed from Actiwatch	-	.979**
20	Total Sleep Time from Actiwatch	.958**	.788**
	Time In Bed from Actiwatch	-	.770**
21	Total Sleep Time from Actiwatch	.898**	.858**
	Time In Bed from Actiwatch	-	.980**
22	Total Sleep Time from Actiwatch	.944**	.913**
	Time In Bed from Actiwatch	-	.969**

** Indicates that correlation is significant at the 0.01 level (2-tailed).

Time in Bed (TIB)

The data for TIB were normally distributed when considered by Duty Day Blocks A, B, and C, by Late-finish days, and by Number of Sectors, but were non-normally distributed when considered by Rest Days or Duty Days with an Early Start.

The TIB for all pilots across the participation period is illustrated in Figure 15. The mean hours of TIB for all 21 pilots (indicated by the red line in Figure 15) was M = 8.14, SD = .48. The lowest number of hours spent in bed corresponds to the last nights of Rest Days before Duty Days with an early start.

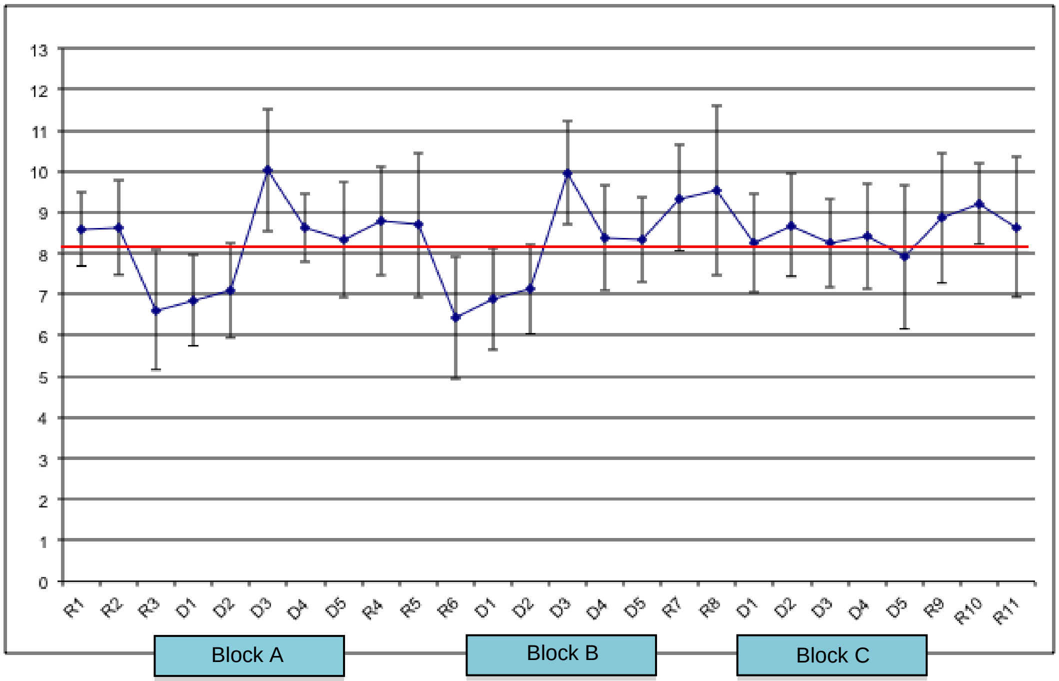


Figure 15 – Average Time in Bed for the group of 21 pilots

Total Sleep Time (TST)

The Kolmogorov-Smirnov test shows that data for TST were normally distributed when considered by Duty Day Blocks, by Early start or Late finish, and by Number of Sectors.

Univariate ANOVA indicated a significant difference between TST during Rest Days and Duty Days, F(16, 526) = 7.83, p< .001. Multiple pair-wise comparisons using Bonferroni test indicated that the mean score for TST (M = 5.04, SD = 1.33) for the night before first day of duty in Block A (i.e., the night of R3) was significantly lower than the TST (M = 6.66, SD = 1.32) during the previous night of sleep (i.e., on R2) and that the mean score for TST (M = 5.04, SD = 1.230) for the night before first day of duty in Block B (i.e., the night of R6) was significantly lower than TST (M = 6.50, SD = 1.37) during the previous night of sleep (i.e., on R5).

There were significant differences in the mean score of TST between the Duty Days D3 (M = 5.39, SD = .85) and D4 (M = 7.48, SD = 1.21) for Block A and between Duty Days D3 (M = 5.58, SD = .94) and D4 (M = 7.72, SD = 1.15) for Block B. These differences are highlighted in red as significant in Figure 16. The mean hours of TST for the group of 21 pilots was M = 6.43, SD = .64 denoted by the red line in Figure 16.

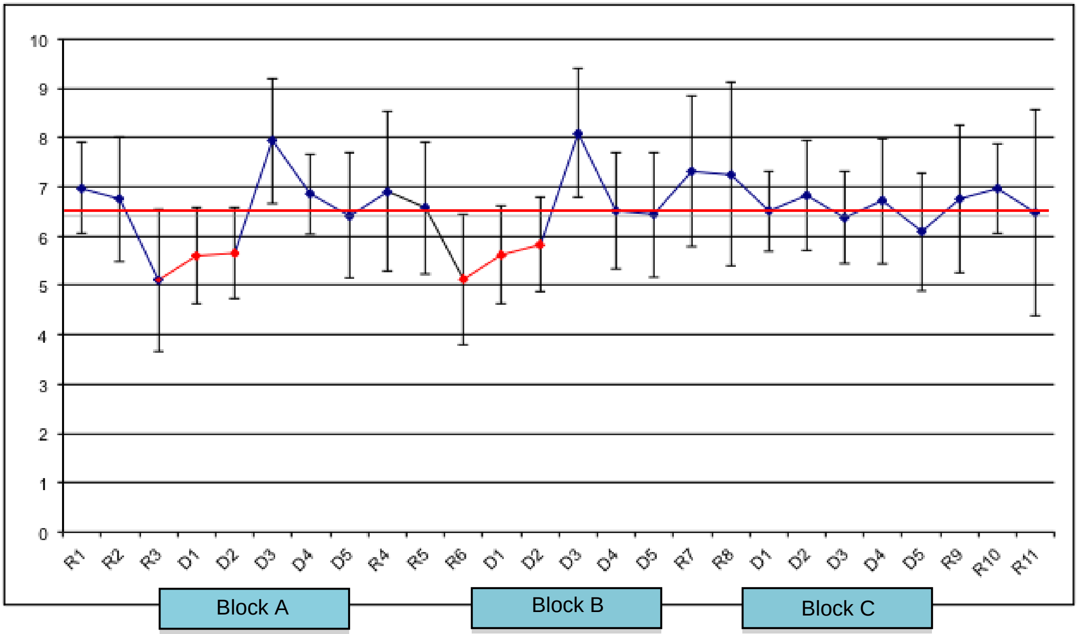


Figure 16 – Average Total Sleep Time for the group of 21 pilots

These results suggest that duty periods with early start times limit the amount of sleep obtained by the pilots on the preceding night while the duty periods with late finish followed by a late start the next day allow for more sleep. Discounting absent days and only 2 exceptions, the rosters for all 22 pilots had early departures and 2 sectors on Days 2 and 3 followed by late arrival and 4 sectors on Day 4 of both Blocks A and B.

Graphs representing TST for each pilot provided in Appendix E indicate a tendency among all pilots for less sleep on a night prior to a day of early departure and more sleep on a night following late finish. This is likely explained by the fact that Duty Days that finish late have start times mostly in early afternoon or, occasionally, late morning. Humans are able to extend their sleep time in the morning more easily than they are able to start their sleep time early in the evening. Sleep cannot be scheduled at anytime of day and expected to be of equal quality. However, this lower sleep time on nights of R3 and on the nights of early-departure Duty Days D1, D2 of Block A and on the nights of R6 and on the nights of early-departure Duty Days D1, D2 of Block C are not reflected in exceptionally high values of Samn-Perelli Scores for any of the following days.

Appropriately, in the rosters of all 22 pilots, there was no case of a late finish Duty Day with an early departure on the following day.

Sleep efficiency

The average Sleep Efficiency for all pilots in this study group is shown in Figure 17. The mean score for Sleep Efficiency for this group of pilots (denoted by the red line in Figure 17) was 79.18% with a standard deviation of 6.77%.

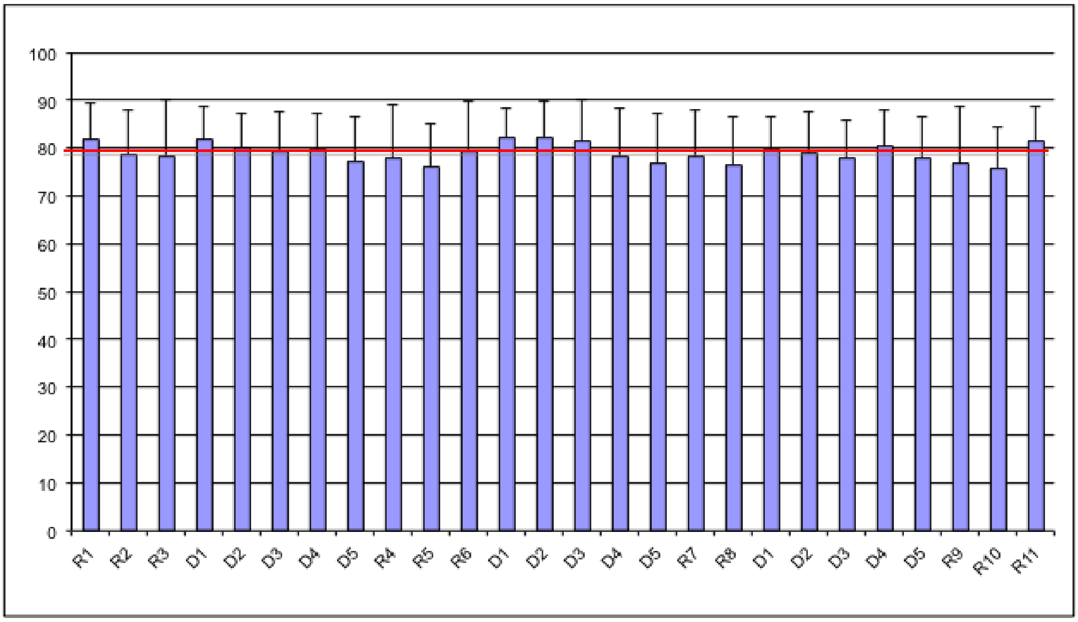


Figure 17 – Average Sleep Efficiency for the group of pilots

Examination of Sleep Efficiency data for each pilot showed large differences among individuals as seen in Figure 18.

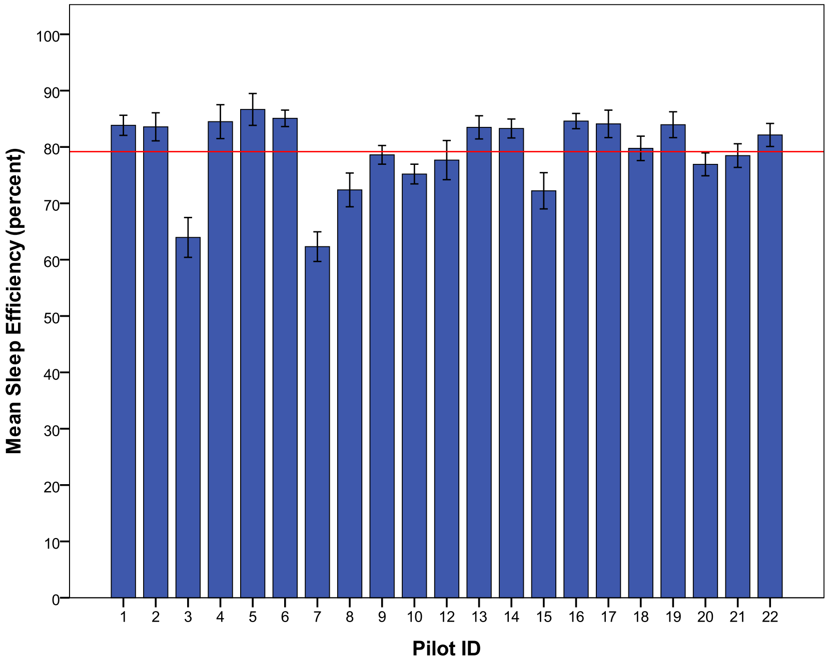


Figure 18 – Sleep Efficiency for each pilot

Normal sleep efficiency is considered to be at least 85% (i.e., asleep 85% of the night). Although most of the pilots had sleep efficiencies of about 80% or higher, only a couple of them achieved 85%. Pilots #3 and #7 show very low sleep efficiency compared to other pilots. Exploratory analyses for these two pilots are presented below.

Note on Pilot #3.

The recorded Sleep Efficiency for Pilot #3 for each night during the participation period is presented in Figure 19. Overall, this pilot presented very low sleep efficiency.

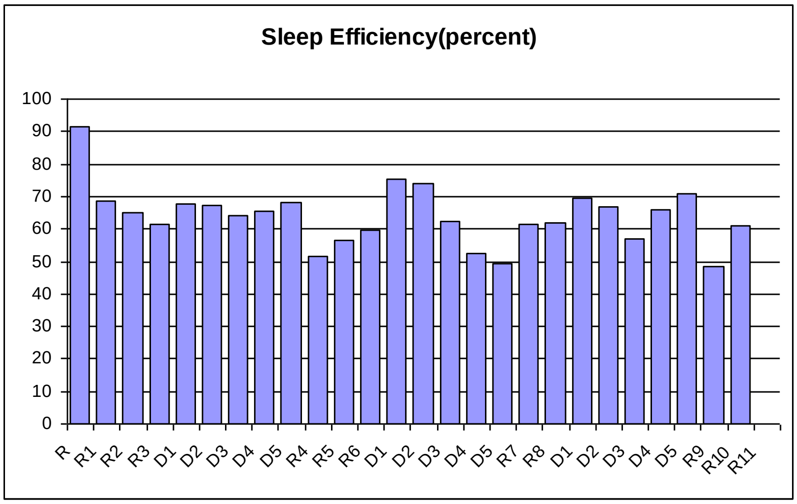


Figure 19 – Sleep Efficiency for Pilot #3

The Actiwatch data for Pilot #3 showed large difference between TIB and TST as shown in Figure 20, even though the two were highly correlated and were also correlated with sleep time from the diary as shown in Table 12.

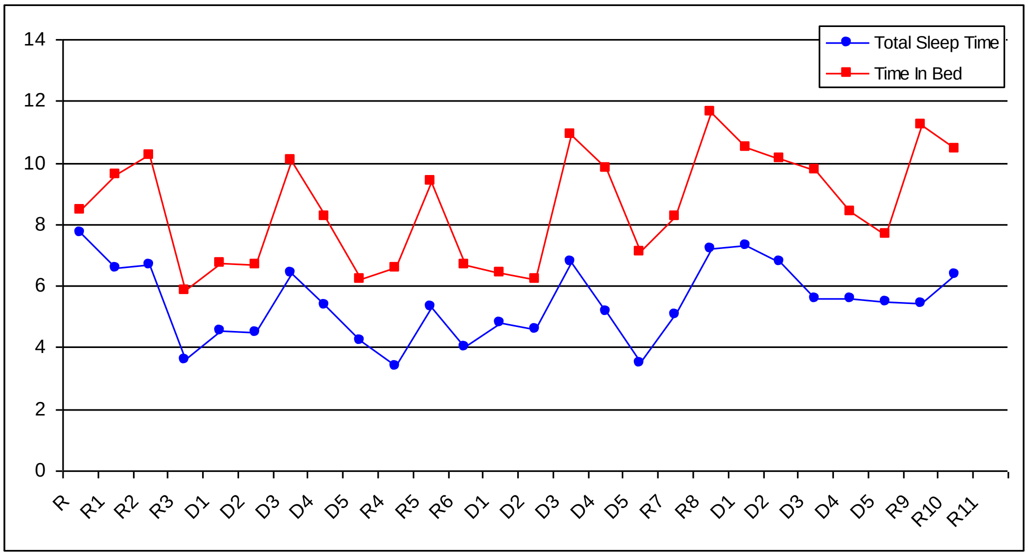


Figure 20 – TST and TIB for Pilot #3

Table 12 – Correlations between TIB, TST, and TST from diary for Pilot #3.

	Time in Bed from Actiwatch	**Total Sleep Time from Diary**
Total Sleep Time from Actiwatch	.795**	.898**
Time in Bed from Actiwatch		.970**

**. Indicates that correlation is significant at the 0.01 level (2-tailed).

The Actogram provides a graphic view of the distribution of rest and activity throughout the day. Examples of the Actograms for Pilot #3 are shown in Figure 21. Activity data are indicated in black, rest periods are in the aqua areas, and the sleep intervals are marked in the blue areas. An examination of the sleep periods in the blue areas shows that this pilot recorded numerous activity counts during his sleep periods every night. This accounts for his low sleep efficiency.

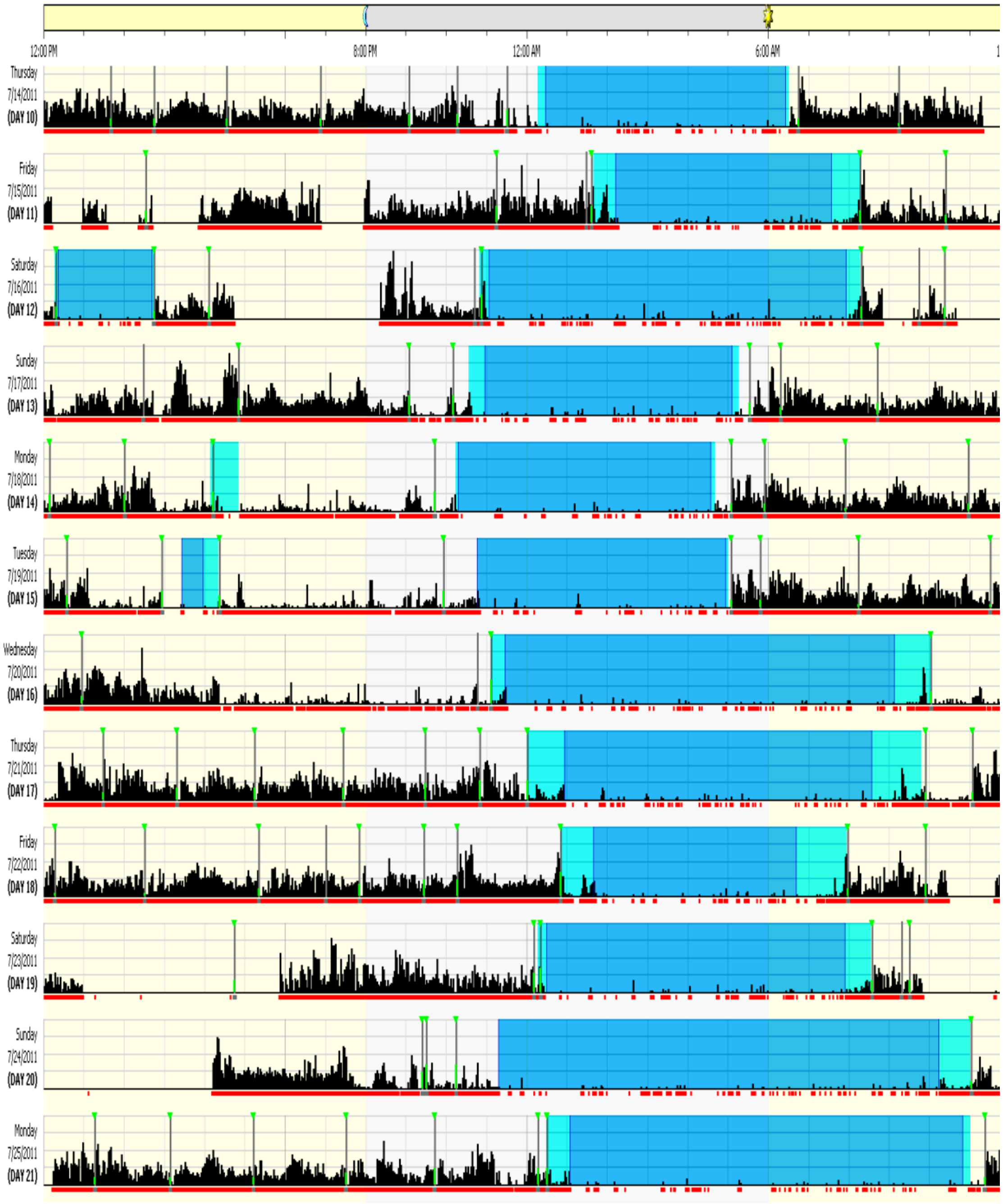


Figure 21 – Actogram representing a sample of night of sleep for Pilot #3

Although Pilot #3 consistently reported in his sleep diary that his sleep quality was between "very good and good", the objective measures show that this pilot had low sleep efficiency. Future studies may reveal relationships of Pilot #3's self-assessment of sleepiness and his performance that are influenced by his low sleep efficiency.

Note on pilot 7

The recorded Sleep Efficiency for Pilot #7 for each night during the study period is presented in Figure 22. Overall, this pilot also presented very low sleep efficiency.

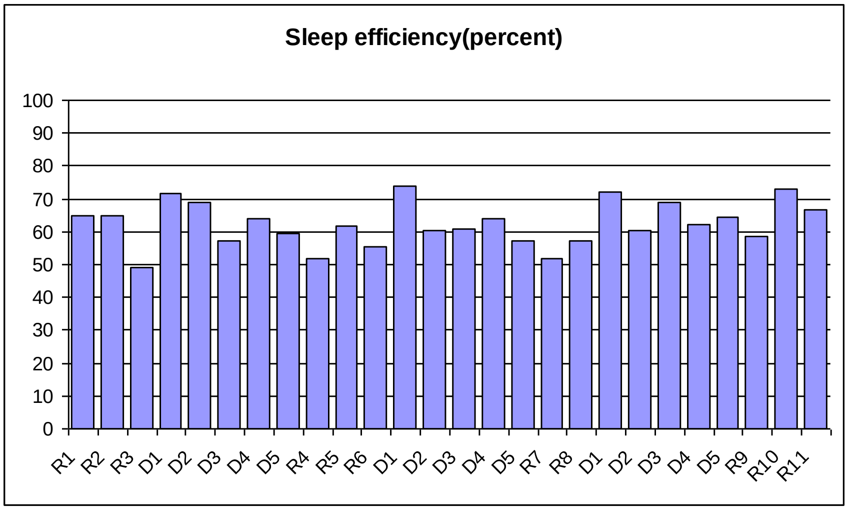


Figure 22 – Sleep Efficiency for Pilot #7

Like Pilot #3, the Actiwatch data for Pilot #7 showed large difference between TIB and TST as seen in Figure 23, even though they were highly correlated as shown in Table 13.

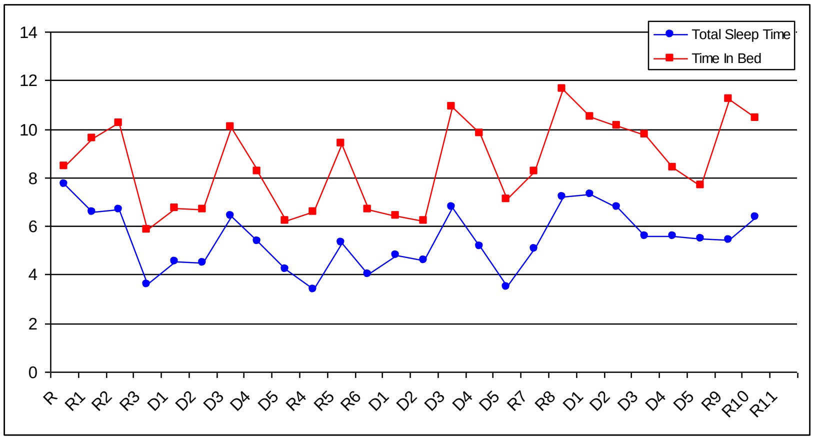


Figure 23 – TST and TIB for Pilot #7

Table 13 - Correlations between TIB, TST, and TST from diary for Pilot #7

	Time in Bed from Actiwatch	**Total Sleep Time from Diary**
Total Sleep Time from Actiwatch	.856**	.875**
Time in Bed from Actiwatch		.966**

** Indicates that correlation is significant at the 0.01 level (2-tailed).

As in the case of pilot #3, Pilot #7's Actogram also showed many activity counts during sleep periods. Examples of the Actograms for Pilot #7 are shown in Figure 24. Activity data are in black, rest periods are indicated in aqua, and sleep intervals are in blue. Examinations of the comments section from Pilot #7's sleep diary did not indicate any explanation for his low sleep efficiency.

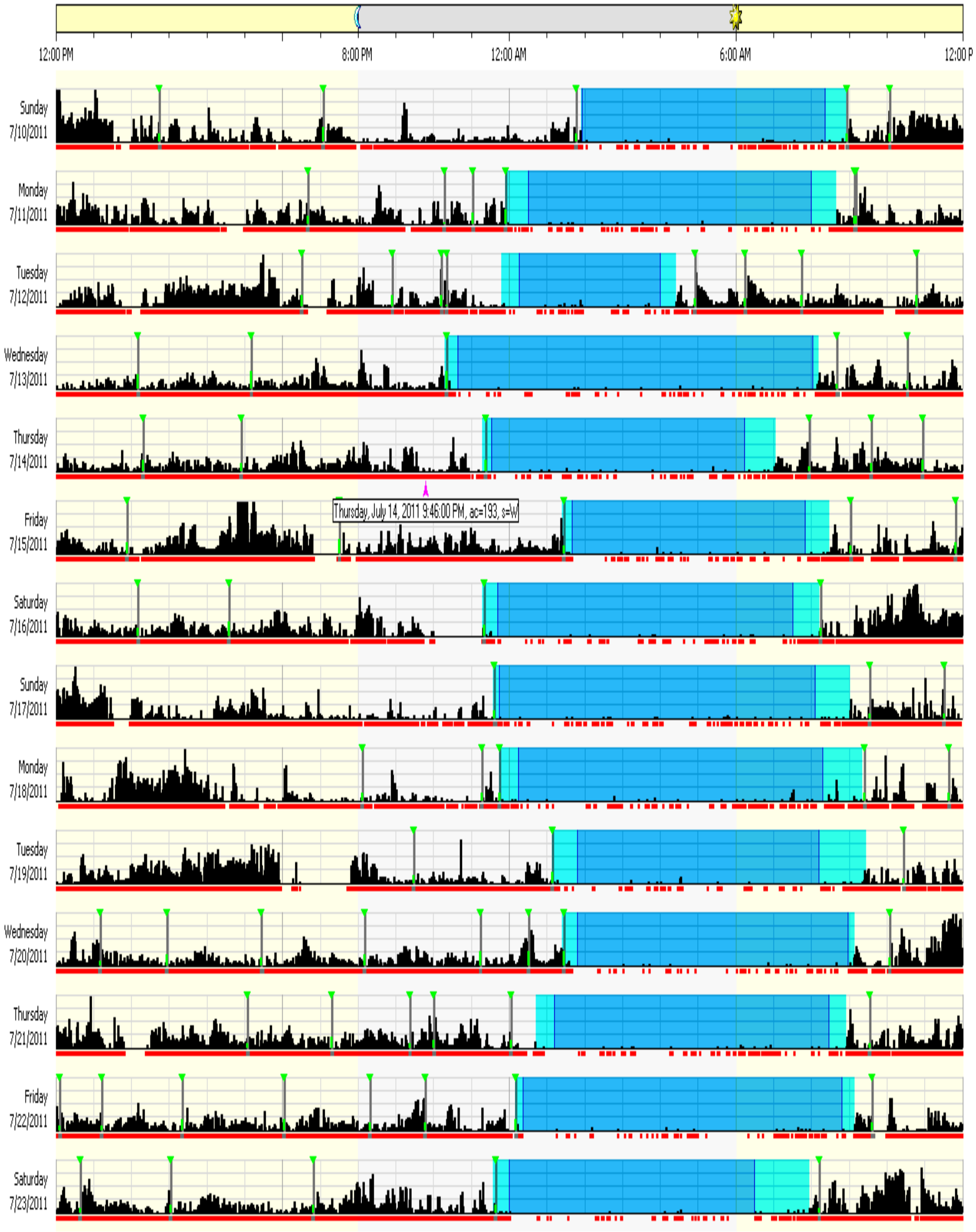


Figure 24 – Actogram representing a sample of nights of sleep for Pilot #7

Sleep Diary

As shown in Table 14, the participants in this study were very conscientious about filling out the Sleep Diary. The days highlighted in purple indicate incomplete data. E indicates a duty day with an early start with ‘n’ sectors; L indicates a duty day with a late finish with ‘n’ sectors; “x” indicates a sick day, and; “a” indicates absent for other reasons.

Table 14 - Sleep diary data completed by each pilot during each day of study

Pilot	R1	R2	R3	Block A					R4	R5	R6	Block B					R7	R8	Block C					R9	R10	R11
				D1	D2	D3	D4	D5				D6	D7	D8	D9	D10			D11	D12	D13	D14	D15			
1				2E	2E	2E	4L	4L				2E	2E	2E	4L	4L			4L	2L	4L	4L	a			
2				4E	a	2E	4L	4L				4E	2E	2E	a	2L			3L	4L	2L	4L	2L	▨		
3				2E	2E	2E	4L	4L				2E	2E	2E	4L	4L			4L	4L	4L	4L	a	▨		
4				4E	2E	2E	2L	4L				4E	2E	2E	4L	4L			4L	3L	4L	a	a		▨	▨
5				4E	2E	2E	4L	2L				4E	2E	2E	4L	2L			4L	2L	4L	2L	2L			▨
6				4E	2E	2E	4L	2L				4E	2E	2E	4L	2L			4L	2L	4L	2L	3L			
7				4E	2E	2E	4L	2L				4E	2E	2E	2L	a			4L	3L	4L	2L	2L			
8				4E	2E	2E	4L	3L				4E	2E	2E	4L	2L			4L	2L	4L	3L	2L			
9				4E	2E	2E	4L	2L				4E	2E	2E	4L	2L			4L	2L	2L	2L	2L			
10				2E	2E	2E	4L	2L				4E	2E	2E	4L	2L			2L	2L	4L	4L	2L			
11		▨		4E	2E	2E	5L	2L				3E	2E	2E	4L	2L			4L	2L	2L	2L	a			
12				4E	2E	2E	x	x				4E	2E	2E	4L	4L			2L	4L	2L	4L	4L			
13				4E	2E	2E	4L	2L				4E	2E	2E	4L	2L			4L	2L	a	2L	a			
14				4E	2E	2E	4L	4L				4E	2E	2E	4L	3L			4L	3L	3L	a	4L		▨	▨
15				4E	2E	2E	4L	2L				4E	2E	2E	4L	2L			4L	2L	4L	2L	2L			
16				4E	2E	2E	4L	2L				4E	2E	2E	4L	2L			4L	2L	4L	2L	2L			
17				4E	2E	2E	4L	4L				4E	2E	2E	4L	3L			4L	3L	4L	2L	a	▨	▨	▨
18				4E	2E	2E	4L	2L				2E	2E	2E	4L	2L			4L	2L	4L	2L	2L			▨
19				4E	2E	2E	4L	2L				4E	2E	2E	4L	2L			4L	2L	4L	2L	4L			
20				4E	2E	2E	x	4L				4E	2E	2E	4L	4L			4L	3L	4L	a	a			▨
21				4E	2E	2E	4L	2L				4E	2E	2E	4L	2L			4L	2L	4L	2L	2L			
22				4E	2E	2E	4L	2L				4E	2E	2E	4L	4L			4L	2L	4L	2L	2L			

The sleep diaries required participants to provide information each day on bedtime and wake-up time, the quality of the sleep, level of fatigue before and after sleep, interruptions, and times of naps and showers. Total sleep times were calculated for each individual based on information from his or her sleep diary and were correlated with his or her actigraphy data.

Group analyses were conducted on the sleep diary parameters. As discussed above, the overall TST and TIB obtained from Actiwatch and TST obtained from Sleep Diary were strongly correlated for the 21 pilots taken as a group. Overall, pilots reported being more tired during duty days than rest days as shown in Figure 25. Before sleep, the majority of the pilots reported feeling “*moderately tired, let down*”. At the end of the sleep period, the majority of the pilots reported feeling “*okay, somewhat fresh*”. Naps were omitted from these calculations but will be accounted for in forthcoming fatigue analyses of the data.

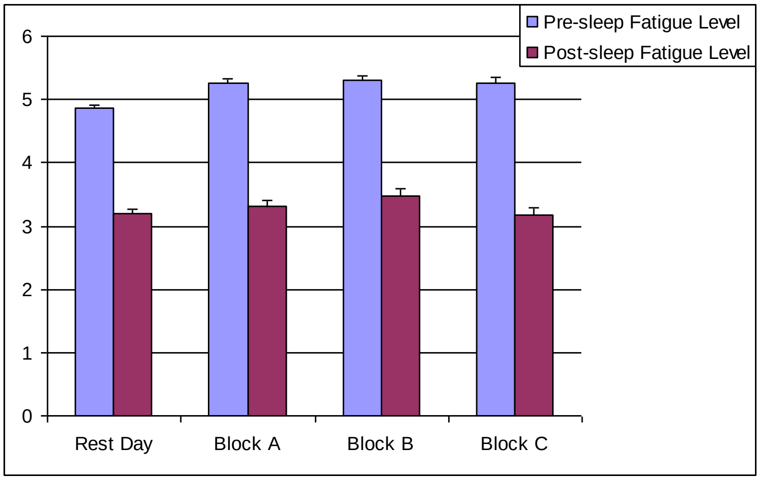


Figure 25 – Subjective pre-sleep and post-sleep fatigue levels by Duty Block (A, B, C) and Rest Days for all pilots

Sleep quality was reported as being "*good*" to "*average*" as illustrated in Figure 26.

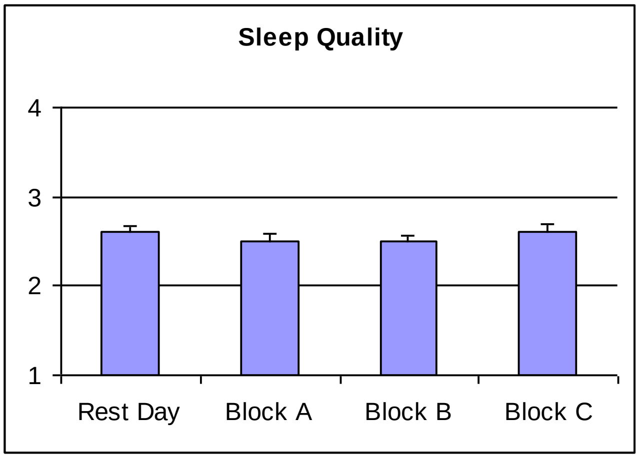


Figure 26 – Subjective sleep quality by Duty Block (A, B, C) and Rest Days for all pilots

Sleep time was reported as being shorter during the first three nights of Duty Blocks A and B, which generally correspond to duty days with early departure (see Figure 27). This suggests that during Duty Days with early departure, pilots sleep less than during Duty Days with late finish or during Rest Days. We saw precisely the same pattern in the Actiwatch data shown in Figure 16, which is reproduced below for comparison with the subjective sleep time from the diary. The patterns are identical with the subjective sleep time being about 1 hour longer than the actual sleep time across the entire sequence.

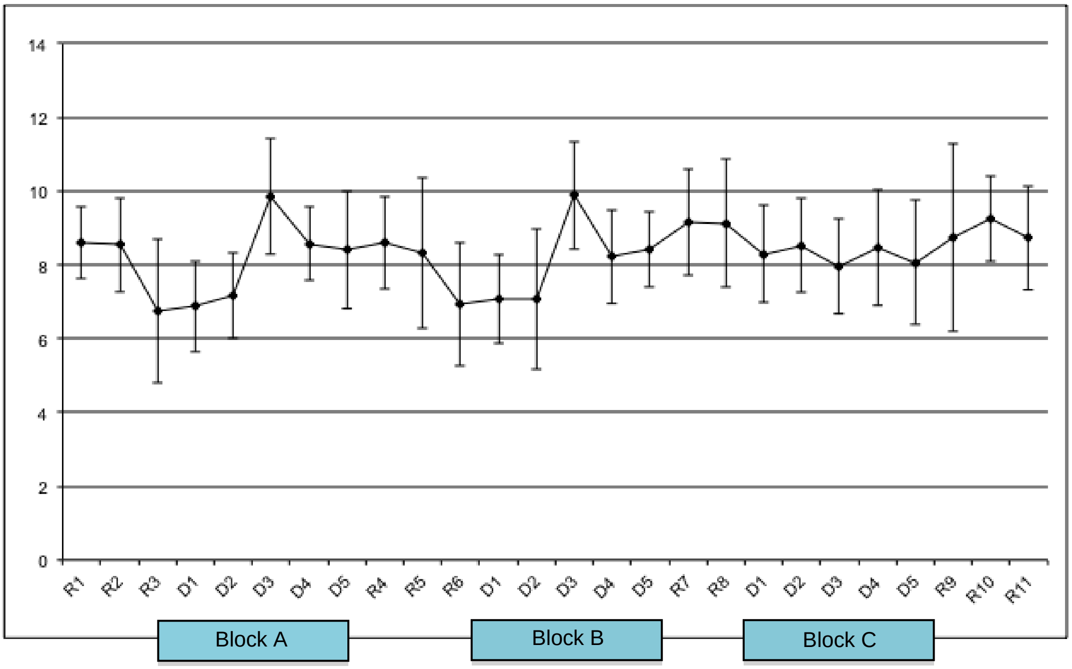
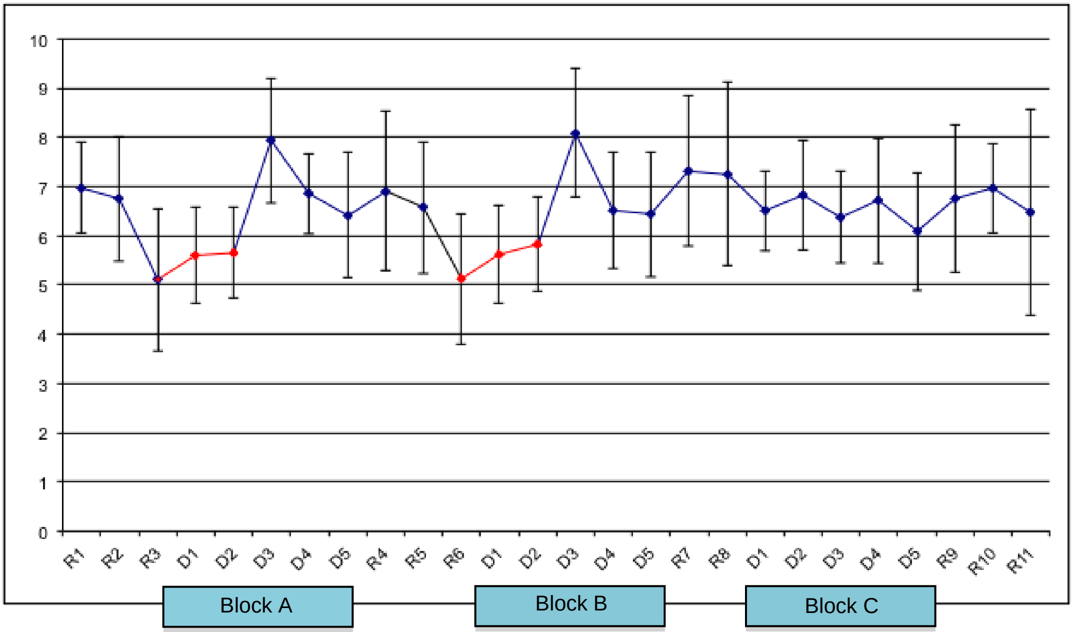
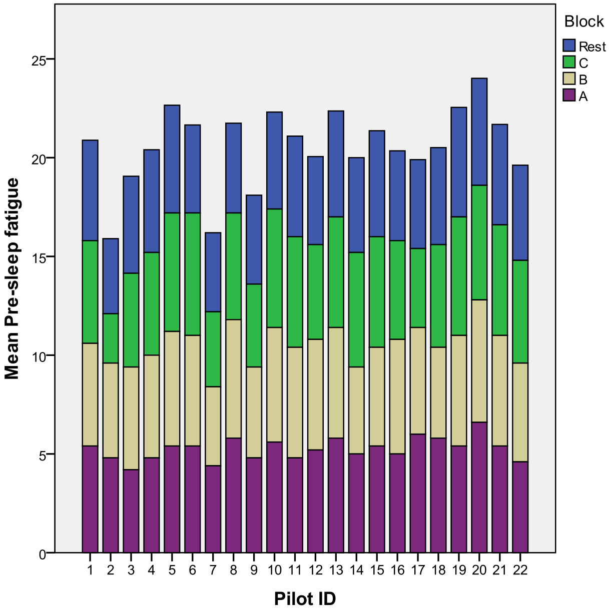


Figure 27 – Average Subjective Sleep Time from Sleep Diaries for the group

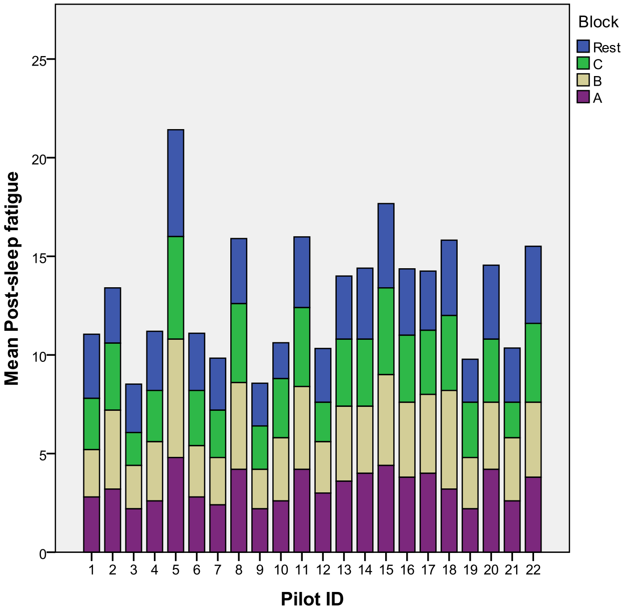


Copy of Figure 16 – Objective sleep time by Rest and Duty Days for each Block for all pilots

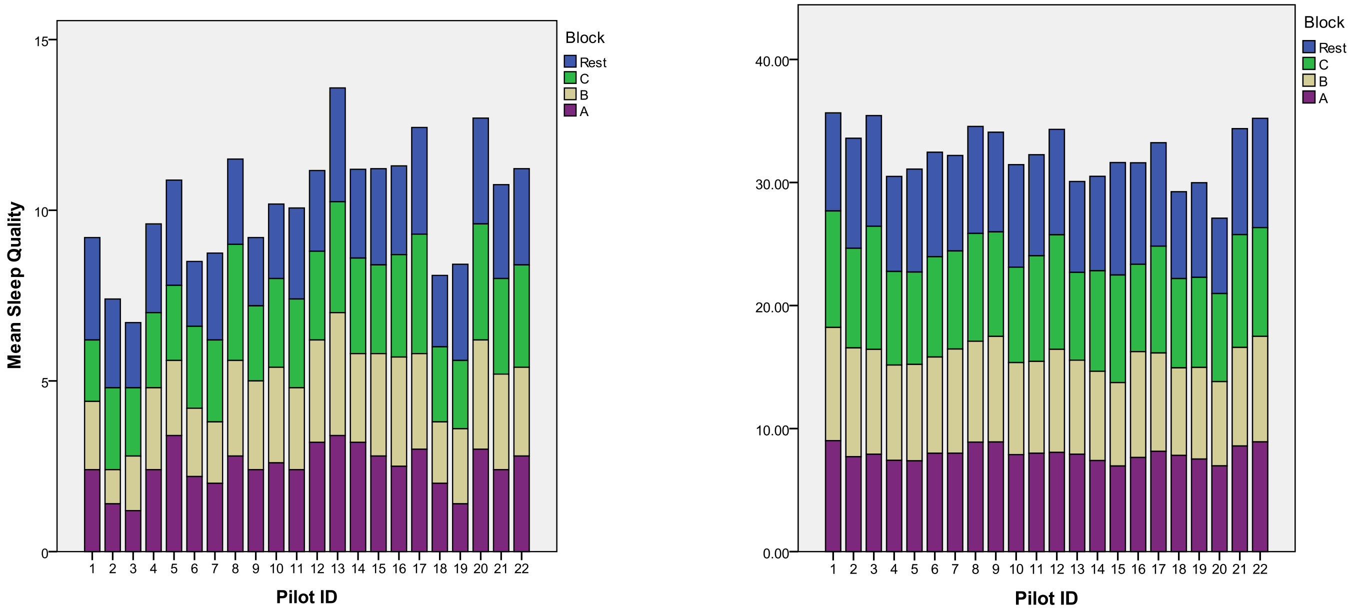
Statistical descriptive analyses were conducted for each individual's self-assessments of pre-sleep fatigue level, post-sleep fatigue level, sleep quality, and sleep time. A summary of these analyses is illustrated in Figures 28.a, b, c, and d and they are described further in the table in Appendix F.



a - Subjective pre-sleep fatigue



b - Subjective post-sleep fatigue



c - Subjective sleep quality

d - Subjective sleep time

Figure 28 – Subjective pre-sleep fatigue, post-sleep fatigue, sleep quality, and sleep time by Rest and Duty Days for each block for all pilots

Whereas pre-sleep fatigue appeared fairly constant across all pilots on Rest Days, pre-sleep fatigue differed significantly among pilots for Duty Days as shown in Figure 28.a. As indicated in Figure 28.b, Pilot #5 reported the highest levels of post-sleep fatigue for all blocks and rest periods, while Pilots #3 and #9 reported the lowest levels of post-sleep fatigue. Figure 28.c shows that sleep quality was the worst for Pilot #13 and best for Pilot #3. Although Pilot #3's self assessment of sleep quality recorded in his sleep diary was invariably between "very good and good", the objective measures from the Actiwatch discussed previously show that this pilot had low Sleep Efficiency. There were significant differences in the mean sleep times among pilots as indicated in Figure 28.d.

The results of these analyses of the data on sleep from the Actiwatch and the Sleep Diary give us high confidence in the quality and reliability of our evaluations of sleep for each of the pilots. This is an important finding because, of course, sleep is a fundamental factor in this study.

NASA TLX and Hassle Factors

The pilots provided evaluations of workload by scoring the 6 subscales of NASA TLX (i.e., Mental Demand, Performance, Physical Demand, Effort, Temporal Demand, and Frustration) (Figure 29) and their explanations of those workloads by completing the checklist of Hassle Factors on their Duty Days (Figure 30).

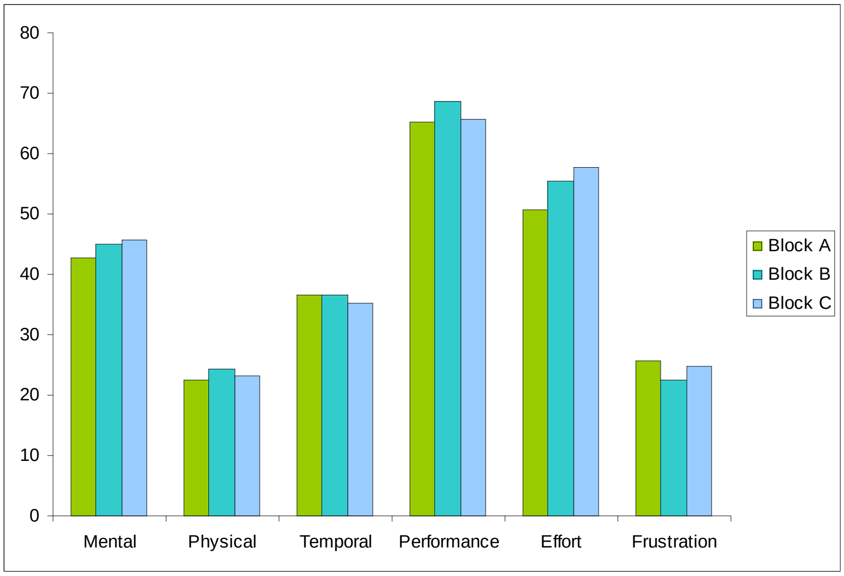


Figure 29 – Averaged NASA TLX Workload Ratings for each Duty Block for all Pilots

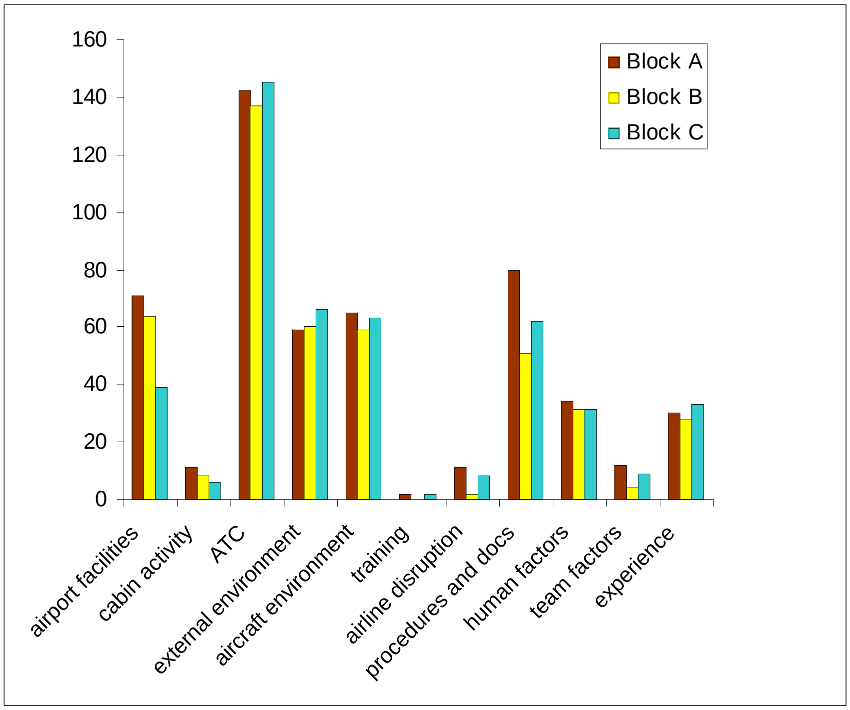


Figure 30 – Summary of Hassle Factors Encountered during each Duty Block for all pilots

At this time, we do not yet know how we will incorporate this information into the analysis or modeling of fatigue or the effects of perceived workload on crew performance.

As might be expected, ATC was frequently identified as a primary hassle factor and was often associated with high levels of Mental Demand, Physical Demand, and Effort, especially at foreign airports.

External Environment is also mentioned often as a hassle factor and it too seems to be associated primarily with increased Mental Demand, Physical Demand, and Effort.

Summary of results and conclusions from our review of the data collected

1. While the data collected by several pilots, particularly of the key measurements of PVT and Samn-Perelli Scores, are highly questionable, the quality and quantity of the data collected in this study is adequate for reliable analyses to achieve the objectives.
2. The average Mean PVT Response Times for the group does not rise to levels indicative of significantly degraded vigilance performance at any time during the Rest or Duty Days of this roster. However, the standard deviations in the average Mean PVT Response Times indicate that individual differences are significant.
3. An important conclusion of our analyses of the Mean PVT Response Times is that, individually, these pilots experienced degradation of performance very infrequently in flight during the Duty Days as scheduled for this study.
4. Like the PVT results, the average values of Samn-Perelli Scores for the group do not rise to disturbing levels at any time during the Rest or Duty Days. The subjective fatigue ratings were not significantly different across the Rest Days or Duty Days.

5. The large standard deviations in the group data of the average Samn-Perelli Scores are indicative of the large individual variability that we also saw in the PVT data.
6. An important conclusion of our analyses of the Samn-Perelli Scores is that very few of the 22 pilots experienced significant levels of fatigue during any of their Duty Days in the roster they flew during this study. These pilots said they experienced more than just mild levels of sleepiness during fewer than 3% of the flights they flew during the study. The rosters flown in this study were not inducing levels of fatigue that were worse than those experienced during Rest Days.
7. In contrast with the results for the Mean PVT Response Times, the Samn-Perelli Scores are strongly correlated with the time of day at which the test was recorded over the entire roster, over the Duty Days, and, especially, over the Rest Days. This is evidence of the strong influence of the circadian rhythm on the feeling of sleepiness that does not necessarily impact the performance of an assigned task by a professional.
8. It is important to recognize that, based on Mood/Alertness Scale data, the self-perception of being ready to perform was invariably higher than the readiness to perform and especially in the evening hours.
9. There seems to be a tendency among all pilots to sleep less on a night prior to a day of early departure and more on a night following late finish. This is likely explained by the fact that Duty Days that finish late have start times mostly in early afternoon or, occasionally, late morning. Humans are able to extend their sleep time in the morning more easily than they are able to start their sleep time early in the evening. Sleep cannot be scheduled at anytime of day and expected to be of equal quality.
10. Pilots need to be taught the circumstances where the likelihood of fatigue is elevated and the importance of using their Rest Days properly. They should be trained to optimize sleep opportunities and, in particular, to maximize the amount of sleep obtained prior to early-morning starts.

Analysis of Data Correlation

In this section, we present the results of correlation analyses of the data collected on crew performance, level of fatigue, and factors of fatigue as the initial step to understanding the relevance of measured factors to the questions we posed for this study. While not all of the results are yet clearly understood, some aspects have become evident. For example, a feature that we have already mentioned and is further revealed and discussed in this section is the profound individual variability in the data. Our analyses confirm the finding by easyJet in their very first study of the problem of fatigue that inter-individual differences (traits and lifestyles) were important determinants of performance, and ultimately fatigue risk. (Stewart 2009) This was evident in the large standard deviations of the means in our analysis of the group data on almost every factor and in the variations in the results of the analyses of the data for the individual pilots.

The analyses discussed in this section focused on correlations to identify

1. statistically significant relationships between performance as measured by PVT and the measurements of fatigue and
2. relationships between self-assessment of sleepiness as indicated by the Samn-Perelli Scores and the latent and proximate factors of fatigue measured during this study.

Tests for normality of distribution revealed that most of the data were non-normally distributed. Therefore, we used Kendall's tau rank-correlation coefficient rather than Pearson's coefficient as a better estimate of correlations in a population with non-normal distribution. It conveys the extent to which pairs of values are in the same rank order. The numerical values indicate the order of the correlation with ±1.0 being perfect correlation. In all of the tables showing Kendall's tau in the following sections, * means p< .05 (significantly correlated), ** means p< .01, and ** means p< .001 (very strongly correlated).

Correlations between Samn-Perelli and PVT

An important question was whether the data revealed correlation between the objective Mean PVT Response Times and the subjective Samn-Perelli Scores. First, we examined the correlation between the group mean of the Samn-Perelli Scores and the group mean of the Mean PVT Response Times. The relationship between Samn-Perelli Scores and the Mean PVT Response Times recorded by the pilots at the same times during the study is shown in Figure 31.

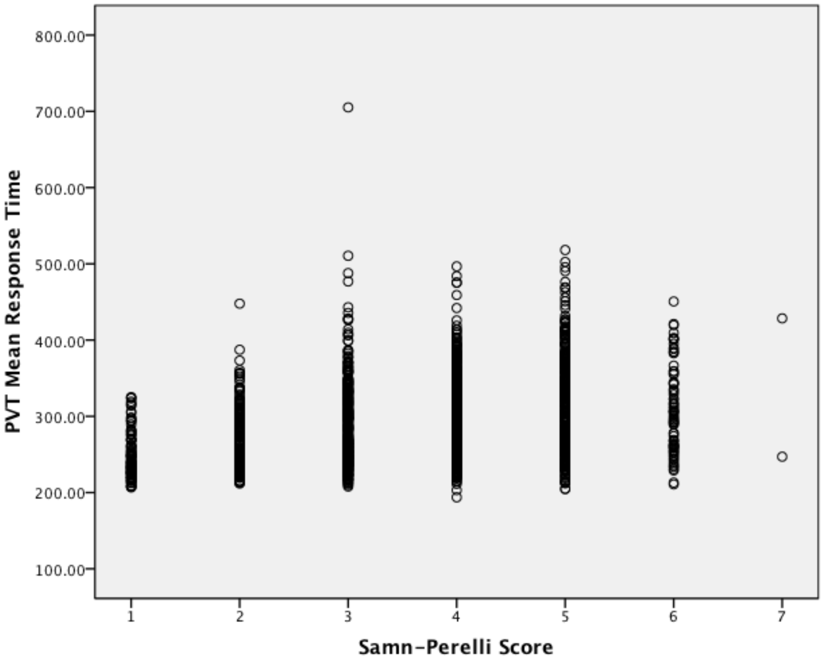


Figure 31 – Mean PVT Response Time (in msec) compared with Samn-Perelli Scores for the Group

Kendall's tau (τ) rank-correlation indicated that the correlation of group data on Mean PVT Response Times and Samn-Perelli Scores was statistically significant, τ = .305, p< .01.

On Figure 32 (a repeat of Figure 31), a solid red line is drawn at the value of Mean PVT Response Time of 400 msec., which we chose as the conservative boundary between acceptable performance and possible degraded vigilance performance. The solid blue line is at a value of Samn-Perelli Score of 4. A score of 4 or less can be interpreted to mean that there is little possibility of significantly impaired performance due to fatigue.

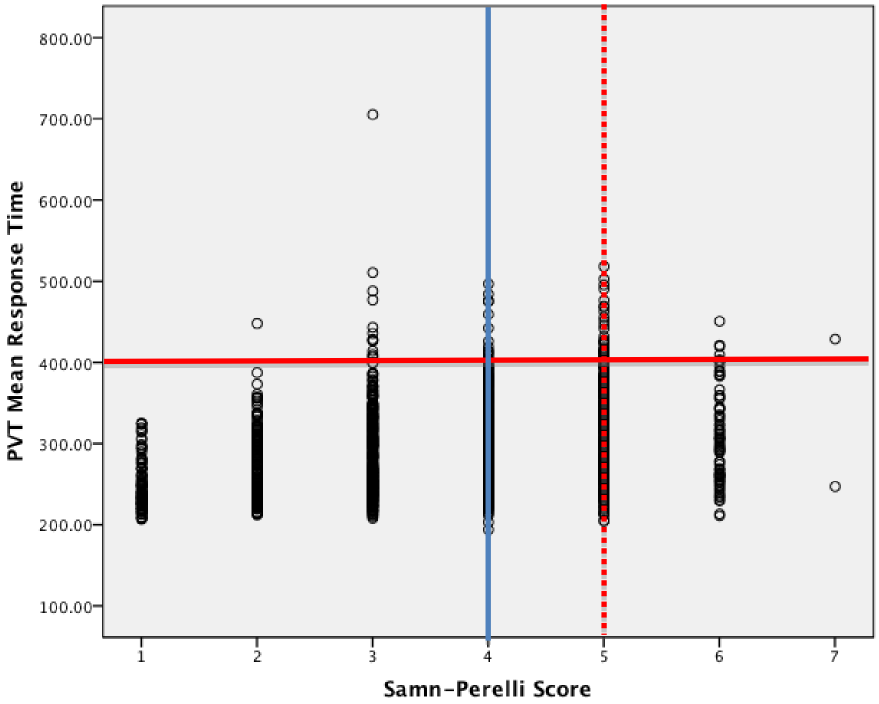


Figure 32 – Comparison of Mean PVT Response Times and Samn-Perelli Scores for the Group

Figure 32 is useful to identify the individual flights to be examined more carefully. There are only 40 data points (out of a total of 1731 data points, 2.3%) in the upper right quadrant of this figure where the Mean PVT Response Time is longer than 400 msec. and the Samn-Perelli Score is higher than 4. Only 19 of these points were recorded In-flight (2.3% of 818 flights), 12 were Post-flight, and 9 were on Rest Days. If we consider the critical value of the Samn-Perelli score to be 5 rather than 4, there are only 7 data points in the uppermost right hand quadrant of Figure 32; 1 of which was recorded In-flight (0.1% of 818 flights); 5 Post-flight, and 1 at the end of a Rest Day. The single data point at a Samn-Perelli score of 7 in that quadrant was recorded during Post-flight. The evidence is that there was statistically insignificant potential for degraded vigilance performance due to fatigue during this study.

We studied in detail the 19 data points in the upper right quadrant of Figure 32 that were recorded In-flight to see if they might indicate a direction for further exploration. Of these, 13 were recorded during the last (C) Block of Duty Days in the Flexible Roster Variation (FRV) schedule. 12 of the 19 data points were recorded during the 3rd or 4th flight leg (sector) flown on a day of late arrival (L). (See Figure 1 for explanation of these references to the schedule Flexible Roster Variation (FRV) used in this experiment.) During only one of the 19 flights was the recorded Samn-Perelli Score a 6. The 19 cases were associated with just 7 of the 21 pilots. One pilot (#14) accounted for 6 of the 19 cases and one other pilot (#11) accounted for 4.

It is significant to note that there were many cases of Samn-Perelli Scores above the moderate sleepiness level of 5 in flight that did not also have high Mean PVT Response Times. This supports the evidence in the literature that there is low probability of a Samn-Perelli Score of 5 causing a decrement in performance.

The question of whether these 19 cases of degraded vigilance performance recorded in flight were associated with degraded aircraft performance is addressed in Section 4.0.

We examined the relationships between the Samn-Perelli Scores and the Mean PVT Response Times for each individual pilot. Table 15 that presents the results of Kendall's tau (τ) rank-correlations applied to the data for each pilot shows the relation between Samn-Perelli Scores and Mean PVT Response Times was significant for 12 out of 21 pilots. (Pilot #2 was not included in this analysis because of invalid PVT data.) This result is an important finding for the plan of this study as represented by the process of analyses diagrammed in Figure 2. We have found that a subjective measure of fatigue correlated with an objective measure of cognitive performance for the group and for most of the pilots. While these correlations are not strong, they are statistically significant. There is nothing that would force such a correlation.

Table 15 - Kendall's tau rank-correlations between Samn-Perelli and PVT for each pilot over the full roster

Pilot	N	Non-normalized		Normalized	
		Kendall's tau		Kendall's tau	
		Correlation Coefficient	Sig. (1-tailed)	Correlation Coefficient	Sig. (1-tailed)
1	79	.034	.341	.034	.341
3	76	.001	.496	.001	.496
4	81	.399**	.000	.399**	.000
5	87	.077	.168	.077	.168
6	83	.378**	.000	.378**	.000
7	73	.080	.186	.080	.186
8	84	.373**	.000	.373**	.000
9	88	.470**	.000	.470**	.000
10	86	.203**	.006	.203**	.006
11	78	.359**	.000	.359**	.000
12	86	.142*	.043	.142*	.043
13	73	.093	.152	.093	.152
14	93	.182*	.012	.182*	.012
15	82	.060	.236	.060	.236
16	74	.266**	.001	.266**	.001
17	55	-.212*	.020	-.212*	.020
18	84	-.028	.361	-.028	.361
19	92	.347**	.000	.347**	.000
20	75	.263**	.001	.263**	.001
21	88	.042	.297	.042	.297
22	92	.250**	.001	.250**	.001
No. of significant		12/21		12/21	
Chance		2/21		2/21	

There are large individual variations in the Samn-Perelli Scores and smaller, but statistically significant, individual variations in the Mean PVT Response Times among these pilots during similar rest and work patterns. We also considered normalized values of Samn-Perelli Scores and Mean PVT Response Times for each pilot. As seen in Table 15, the correlation coefficients and the significance levels of the non-normalized and the normalized data were identical.

We calculated the Kendall's tau rank-correlations between Samn-Perelli Scores and Mean PVT Response Times for each pilot over just their Rest days (Table 16) and their Duty Days (Table 17).

Table 16 - Kendall’s tau rank-correlations between Samn-Perelli and PVT for each pilot over Rest Days

Pilot	N	Kendall’s tau	
		Correlation Coefficient	Sig. (1-tailed)
1	22	.282*	.043
3	25	.263*	.047
4	25	.494**	.001
5	30	.268*	.027
6	27	.106	.244
7	32	-.053	.354
8	32	.273*	.026
9	33	.541**	.000
10	31	.198	.075
11	29	.487**	.001
12	33	.356**	.005
13	24	.240	.070
14	33	.093	.254
15	31	.291*	.018
16	21	.436**	.006
17	11	-.166	.254
18	29	-.280*	.025
19	33	.629**	.000
20	32	.227*	.046
21	33	.301*	.011
22	33	.162	.123
Correlation		13/21	

Table 17 - Kendall’s tau rank-correlations between Samn-Perelli and PVT for each pilot over Duty Days

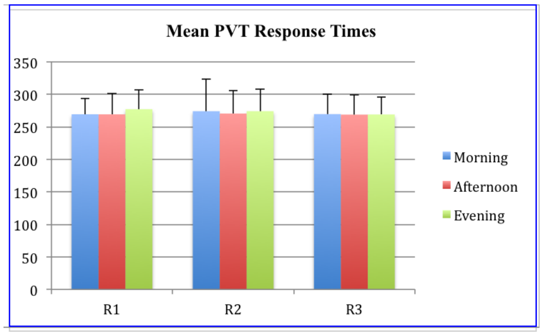
Pilot	N	Kendall’s tau	
		Correlation Coefficient	Sig. (1-tailed)
1	57	-.102	.154
3	51	-.014	.449
4	56	.351**	.000
5	57	-.013	.449
6	56	.398**	.000
7	41	.196	.050
8	52	.394**	.000
9	55	.425**	.000
10	55	.207*	.022
11	49	.287**	.005
12	53	.016	.442
13	49	.007	.474
14	60	.159	.056
15	51	-.083	.224
16	52	.145	.090
17	44	-.373**	.001
18	55	.086	.195
19	59	.181*	.033
20	43	.188	.054
21	55	-.104	.151
22	59	.302**	.002
Correlation		8/21	

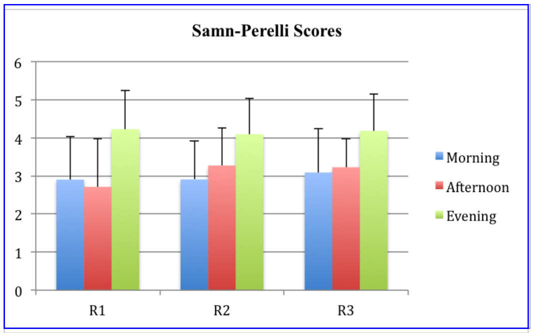
The data for about the same number of pilots showed correlation between Samn-Perelli Scores and Mean PVT Response Times during Rest Days as during the full roster, but they were not the same. Fewer pilots showed correlation over their Duty Days, but all of them were among those for whom there was correlation over the full roster. Although the Samn-Perelli Scores and the Mean PVT Response Times are considered significantly correlated over the full roster and over the Rest Days, there are meaningful differences between them on Duty Days that explain the lesser correlation. Some of these differences can be seen in the graphical presentations of the averages shown over Rest Days and over Duty Days in Figures 33 and 34, respectively.

For low levels of Samn-Perelli scores, sleepiness seems to have little effect on the performance of these pilots. This might be explained by the ability of well–trained professionals to adapt to and overcome low levels of tiredness to perform his or her job acceptably.

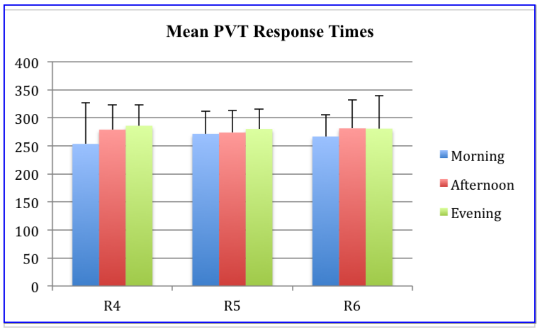
Rest Days

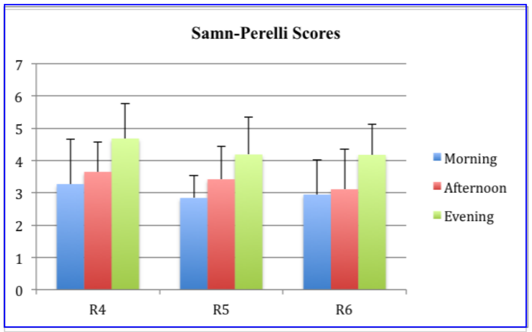
R1-R2-R3



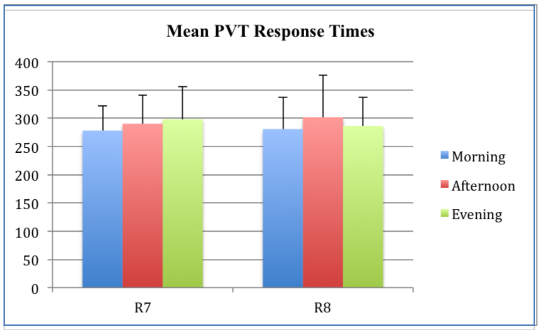


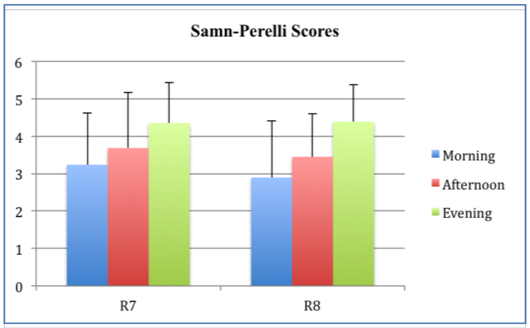
R4-R5-R6





R7-R8





R9-R10-R11

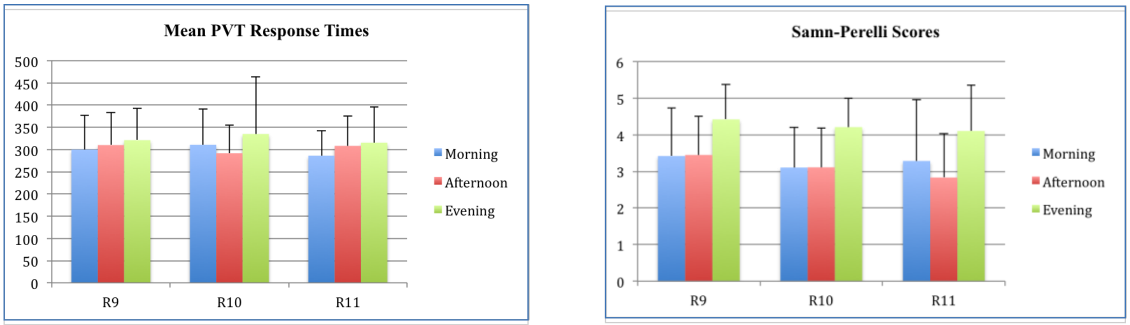
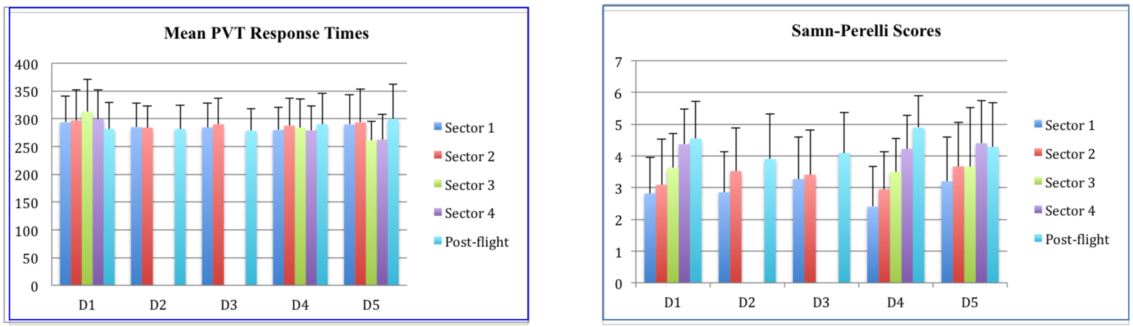


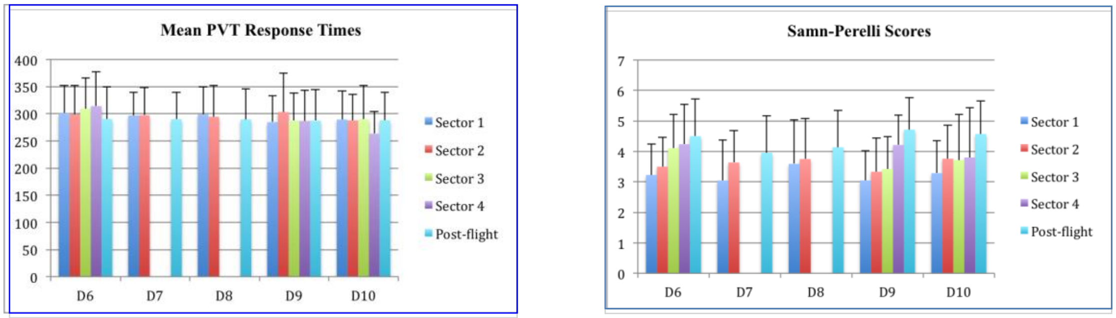
Figure 33 –Averages of PVT Response Times and Samn-Perelli Scores for Rest Days

Duty Days

BLOCK A



BLOCK B



BLOCK C

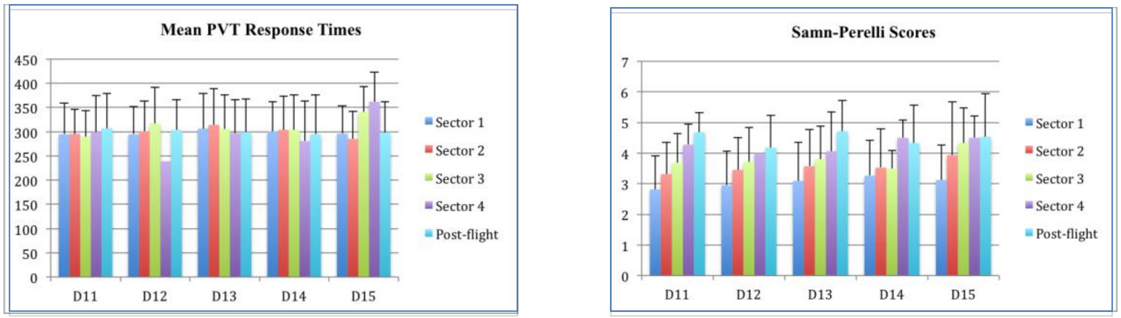


Figure 34 –Averages of PVT Response Times and Samn-Perelli Scores for Duty Days

Correlation between Mean PVT Response Times and Samn-Perelli Scores is so fundamental to our approach to this study that we invested time to try to understand why the there was not some level of correlation between performance and self-assessment of sleepiness for all of the pilots. We examined plots of the Samn-Perelli Scores and Mean PVT Response Times for each pilot to explore the correlations, but there are no visually obvious differences between the ones identified as being statistically correlated and those that are not correlated.

We compared several of the characterizing features of the group of 12 pilots for whom their Samn-Perelli Scores correlated with their Mean PVT Response Times with those of the group of 9 pilots for whom there was no such correlation. Table 18, for example, is a comparison of the descriptive statistics (i.e., Mean and Standard Deviation) for the pilots in the two groups. Group A includes all the pilots for whom the Samn-Perelli Scores and the Mean PVT Response Times were correlated. Group B includes those pilots for whom they were not correlated.

Table 18 – Comparison of Descriptive Statistics for Group A for which Samn-Perelli and PVT were correlated and Group B for which they were non-correlated

	N	**Minimum**	**Maximum**	**Mean**	**Std. Deviation**
Group A: Mean PVT Response Times	12	208.00	496.62	301.06	52.69
Group B: Mean PVT Response Times	9	193.78	705.14	275.89	57.26
Group A: Samn-Perelli Scores	12	1	7	3.67	1.18
Group B: Samn-Perelli Scores	9	1	7	3.57	1.39
Group A: Sleep prior to Rest Day	12	5.61	7.53	6.6636	.57949
Group B: Sleep prior to Rest Day	9	4.68	7.78	6.4740	.92750
Group A: Sleep prior to Duty Day	12	4.56	6.95	5.7582	.61415
Group B: Sleep prior to Duty Day	9	4.74	6.52	5.5920	.66001
Group A: CIS	12	38	105	64.50	21.55
Group B: CIS	9	38	102	64.22	18.25
Group A: ESS	12	2	19	7.67	4.60
Group B: ESS	9	4	13	8.11	2.89

With the exception of the highlighted value of Mean PVT Response Time of 705.14 for Group B (which may well be due to bad data), there are no significant differences in the Means or the Standard Deviations of the Mean PVT Response Times or the Samn-Perelli Scores between Group A and Group B.

The descriptive statistics for CIS show that there are no differences between the two groups of pilots while the ones for ESS show just a small difference between the two groups with the pilots in Group B being slightly more tired overall, but the variability is larger for Group A.

We examined the correlation between Samn-Perelli Scores and Mean PVT Response Times for Group A and for Group B. Of course, the first group showed correlation with Kendall's tau rank-correlation = .372 and $p< .01$. The surprising result was that there was correlation (Kendall's tau = .233 and $p< .01$) for the second group as well when all pilots from the group were included even though none of the pilots in this group presented correlation of their individual Samn-Perelli Scores and their Mean PVT Response Times.

We examined correlation between the Total Sleep Time (TST) during the previous 24 hours and Samn-Perelli Scores and Mean PVT Response Times for the two groups and found no significant differences. The Kendall's tau rank-correlation for Mean PVT Response Times and TST was $\tau = -.148$, $p< .01$ for Group A and $\tau = -.197$, $p< .01$ for Group B. Similarly, the correlation for Samn-Perelli Scores and TST was $\tau = -.152$, $p<.01$ for Group A compared with $\tau = -.142$, $p<.01$ for Group B.

We considered the work patterns and found that 1 pilot in the group for whom the Samn-Perelli Scores and Mean PVT Response Times correlated and 2 pilots in the group for whom they did not flew different schedules during Blocks A and B than the rest of the pilots. However, this cannot explain the difference between the two groups in the SP-PVT correlations. We also found no significant correlations for either group between the number of sectors flown and the Samn-Perelli Scores or the Mean PVT Response Times.

In summary, our analyses of the data for the two groups of pilots did not reveal any differences that help us understand why the Samn-Perelli Scores correlate with the Mean PVT Response Times for some pilots and not for others. There are some questionable Samn-Perelli and PVT data for several of the pilots in Group B that seem to indicate they either misunderstood the instructions or failed to follow them precisely. However, even if these were deleted from consideration, there would still remain a large number of pilots in the group for which their Samn-Perelli Scores did not correlate with their Mean PVT Response Times.

To complete our study of correlation between PVT data and Samn-Perelli Scores, we examined the correlation of Samn-Perelli Scores with lapses in the PVT. We found that PVT Lapses correlated with Samn-Perelli Scores for only 7 of 21 pilots and even these were of low magnitude. We failed to find any distinctive differences in the patterns of the scatter plots of Samn-Perelli Scores and PVT Lapses between those pilots with a high number of lapses and those with fewer lapses. We examined the Samn-Perelli Scores and the PVT Lapses for 2 sector and 4-sector Duty Days. Typically, there was an increase in the Samn-Perelli Score during the day for both 2-sector days and 4-sector days, but it was not reflected in increased PVT Lapses.

Our conclusion on the use of PVT Lapses was that the data from this study were not sufficiently reliable to use as a measure of fatigue. We recommend, in future experiments, that consideration be given to using the 10-minute PVT on Rest Days for a baseline and that subjects be cautioned about interruptions during tests.

Correlations of Mean PVT Response Times and Samn-Perelli Scores with time of test

We examined the effect of the time of the day at which the test was taken on Mean PVT Response Times and on Samn-Perelli Scores. Tables 19, 20, and 21 present Kendall's tau rank-correlations for Mean PVT Response Times with the time of day at which the test was taken over the entire roster (Table 19), during Rest days (Table 20), and during Duty Days (Table 21). In the all the tables below, * means *p* < .05, ** means *p* < .01, and *** means *p* < .001.

Table 19 – Kendall's tau correlation of Mean PVT Response Time and Time of Test over the entire roster

Pilot #	N	Kendall's tau	
		Correlation Coefficient	Sig. (1-tailed)
1	80	-.147*	.027
3	77	.052	.253
4	81	.344**	.000
5	90	-.303**	.000
6	89	.099	.085
7	73	.012	.441
8	84	.165*	.013
9	88	.107	.069
10	86	.103	.081
1	80	.293**	.000
12	90	-.161*	.012
13	80	.040	.302
14	93	.179**	.006
15	83	-.047	.263
16	76	.024	.380
17	60	-.099	.132
18	84	-.249**	.000
19	92	.109	.062
20	76	.001	.493
21	90	.018	.401
22	92	.121*	.044
Correlation		5/21	

Table 20 – Kendall's tau correlation of Mean PVT Response Times and Time of Test over Rest Days

Pilot #	N	Kendall's tau	
		Correlation Coefficient	Sig. (1-tailed)
1	23	.071	.317
3	25	.173	.112
4	25	.193	.088
5	33	-.119	.164
6	33	.032	.396
7	32	.254*	.021
8	32	.131	.146
9	33	.229*	.030
10	31	.110	.193
11	31	.286*	.012

12	33	-.036	.384
13	31	.058	.323
14	33	.104	.197
15	32	.282*	.012
16	23	.012	.468
17	16	-.183	.161
18	29	.052	.347
19	33	.309**	.006
20	32	.254*	.021
21	33	.127	.150
22	33	-.017	.445
Correlation		6/21	

Table 21 – Kendall's tau correlation of Mean PVT Response Time and Time of Test over Duty Days

Pilot #	N	Kendall's tau	
		Correlation Coefficient	Sig. (1-tailed)
1	57	-.283**	.001
3	52	-.040	.338
4	56	.060	.255
5	57	-.428**	.000
6	56	.162*	.039
7	41	-.198*	.034
8	52	.411**	.000
9	55	.022	.405
10	55	.192*	.019
11	49	.323**	.001
12	57	-.254**	.003
13	49	.032	.372
14	60	.232**	.004
15	51	-.076	.215
16	52	.064	.251
17	44	-.123	.120
18	55	-.397**	.000
19	59	.006	.471
20	44	-.019	.428
21	56	-.086	.175
22	59	.164*	.034
Correlation		6/21	

Contrary to expectation, the Mean PVT Response Times show weak and even negative correlation with the time of the day at which the test was taken even on Rest Days. In contrast, Tables 22, 23, and 24 show strong correlations for Samn-Perelli Scores with the time of day over the entire roster (Table 22), during Rest days (Table 23), and during Duty Days (Table 24).

Table 22 – Kendall's tau correlation of Samn-Perelli Scores and Time of Test over the entire roster

Pilot #	N	Kendall's tau	
		Correlation Coefficient	Sig. (1-tailed)
1	79	.109	.097
2	58	-.026	.395
3	76	.503**	.000
4	81	.344**	.000
5	87	.100	.105

6	83	.345**	.000
7	73	.176*	.025
8	84	.297**	.000
9	88	.356**	.000
10	86	.650**	.000
11	79	.561**	.000
12	86	.388**	.000
13	73	.504**	.000
14	93	.618**	.000
15	83	.176*	.017
16	74	.162*	.034
17	55	-.002	.491
18	84	-.027	.366
19	92	.519**	.000
20	75	.318**	.000
21	88	.483**	.000
22	92	.327**	.000
Correlation		17/22	

Table 23 – Kendall's tau correlation of Samn-Perelli Scores and Time of Test over Rest Days

Pilot #	N	Kendall's tau	
		Correlation Coefficient	Sig. (1-tailed)
1	22	.086	.300
2	17	.378*	.026
3	25	.614**	.000
4	25	.316*	.022
5	30	-.250*	.036
6	27	.373**	.007
7	32	.164	.123
8	32	.434**	.001
9	33	.354**	.004
10	31	.721**	.000
11	29	.384**	.005
12	33	.314*	.011
13	24	.625**	.000
14	33	.519**	.000
15	32	.541**	.000
16	21	.306*	.038
17	11	.583*	.010
18	29	.173	.113
19	33	.477**	.000
20	32	.582**	.000
21	33	.617**	.000
22	33	.292*	.019
Correlation		18/22	

Table 24 – Kendall's tau correlation of Samn-Perelli Scores and Time of Test over Duty Days

Pilot #	N	Kendall's tau	
		Correlation Coefficient	Sig. (1-tailed)
1	57	.109	.137
2	41	-.069	.284
3	51	.489**	.000
4	56	.361**	.000

5	57	.320**	.001
6	56	.341**	.000
7	41	.188	.058
8	52	.370**	.000
9	55	.383**	.000
10	55	.634**	.000
11	50	.685**	.000
12	53	.420**	.000
13	49	.451**	.000
14	60	.691**	.000
15	51	-.098	.186
16	52	.122	.130
17	44	-.059	.311
18	55	-.191*	.028
19	59	.541**	.000
20	43	.225*	.026
21	55	.402**	.000
22	59	.362**	.000
Correlation		15/22	

In contrast with the results for the Mean PVT Response Times, the Samn-Perelli Scores are strongly correlated with the time of day at which the test was recorded over the entire roster, over the Duty Days, and, especially, over the Rest Days. This is further evidence of the strong influence of the circadian cycle on the feeling of sleepiness that does not necessarily impact the performance of a professional on an assigned task.

Correlations of Samn-Perelli Scores with sleep

To gain information of the effect of sleep on the self-awareness of sleepiness, we examined the correlations of Samn-Perelli Scores with Total Sleep Time (TST) on the previous day (including naps), with Sleep Loss the previous night, and with Cumulative Sleep Loss since the start of roster.

Analysis of variance (ANOVA) of the group data indicated that the TST during the previous day (including naps) and the Cumulative Sleep Loss significantly affected the averaged Samn-Perelli Scores of the pilots taken as a group over the entire roster, but that Sleep Loss the previous night had no significant effect as indicated in Table 25.

Table 25 - ANOVA on Samn-Perelli scores
for the group with sleep over entire roster

Cumulative sleep loss	TST (Previous day’s sleep)	Sleep Loss
F(1,1566)= 13.999***	F(1,1566)= 72.86***	F(1,1566)= .51

*** Indicates correlation level at p< 0.001

Kendall’s tau rank correlations between TST on the previous day (including naps) and the Samn-Perelli Scores of the pilots taken as a group showed significant correlation (τ = -.110, p< .05) during Rest Days and showed high correlation (τ = -.162, p< .01) during Duty Days. Kendall’s tau rank correlations between TST on the previous day (including naps) and the Samn-Perelli Scores of the pilots taken as a group showed high correlation during both Duty Days with 2 sectors (τ = -.163, p< .01) and Duty Days with 4 sectors (τ = -.210, p< .01).

Table 26 is a summary of the Kendall's tau rank-correlations of Samn-Perelli Scores for each pilot with their Cumulative Sleep Loss, TST, and Sleep Loss and. Pilot #11 was omitted due to missing Actigraphy data.

Table 26 - Kendall's tau rank-correlations of Samn-Perelli scores for each pilot over entire roster

Pilot	Cumulative sleep loss	TST (Previous day's sleep)	Sleep Loss
1	.042	-.093	.101
2	-.063	-.219	.219
3	.113	.094	-.121
4	-.068	-.126	.167
5	.182*	-.063	.080
6	-.023	-.288***	.366*
7	-.150	-.151*	.149
8	.028	-.184*	.271
9	-.083	-.042	.014
10	.110	-.050	.128
12	-.070	-.197*	.267
13	.075	-.223**	.306
14	.153*	.131	-.143
15	-.075	-.327***	.397**
16	.194*	-.226**	.231
17	-.055	-.308**	.485**
18	.026	-.302***	.403**
19	.174*	.060	-.099
20	-.143	-.134	.234
21	.106	-.093	.194
22	.319***	-.182*	.225
# sig	5/21	10/21	4/21

*$p < .05$, **$p < .01$, ***$p < .001$

The row labeled "#sig" indicates the number of significant correlations in that column.

The Samn-Perelli Scores for an individual pilot correlated strongly (negatively) with the TST the previous day. TST on the previous day (including naps) had a significantly larger effect on the feeling of sleepiness during the following day than did either Sleep Loss the previous day or Cumulative Sleep Loss from start of roster.

Correlations of Samn-Perelli Scores with work pattern

We represented the effects of the work pattern over the roster in several ways in order to examine its effect on Samn-Perelli Scores. We considered the following representations:

> Rest Days vs. Duty Days: Used to represent the effects of the differences between the activities and sleep over Rest Days and the activities and sleep over Duty Days over the course of the 26-day roster. This representation is the coarsest category, simply contrasting Samn-Perelli Scores on Rest Days vs. Samn-Perelli Scores on Duty days.
>
> Sequence of Days within Rest/Duty Blocks: Used to represent effects of the succession of days within the blocks of Rest Days or within the blocks of Duty Days over the course of the 26-day roster. Sequence is an intermediate level of categorization. This would be sensitive to any kind of trend that is consistent within Duty Blocks or consistent within Rest Blocks such as better performance toward the beginning of a Duty Block, or better performance toward the end of a Duty Block.

Time: Used to represent the effects of the times of the daily and cumulative activities and sleep over the course of the 26-day roster. It is literally the time of the day in 2-hour increments (using the 24-hour clock) continuously from first measurement in the morning of the first Rest Day, R1, to the last measurement in the evening of the third recovery day, R11. This is specifically designed to pick up on circadian patterns, such as decreased performance very early in the morning and/or very late at night.

We conducted ANOVA on the Samn-Perelli Scores for both the group of pilots and the individual pilots to explore for the significant main effects of the work pattern. Overall, ANOVA on the Samn–Perelli data for the whole group of pilots revealed significant main effects of all three characterizations of work pattern.

Table 27 - ANOVA on Samn-Perelli Scores with work pattern for the group

Rest days vs. Duty days	Sequence of days within Rest/Duty blocks	Time
F(1,1566)= 15.23***	F(6,1566)= 3.99***	F(11,1566)= 22.36***

*** Indicates correlation level at p< 0.001

The results of this analysis of the effects of time using 2-hour increments in Table 27 agrees with the results reported previously of the correlation of the Samn-Perelli Scores with the time of the day at which they were recorded presented in Table 22. As we showed in Figures 8 and 9, the highest Samn-Perelli scores were recorded at the end of each day, i.e., in the evening of each Rest Day and at the Post-flight measurement on a Duty Day and, to some extent, this is a reflection of the diurnal effect of circadian rhythm on sensation of drowsiness. This daily effect was not seen in the Mean PVT Response Times shown in Figures 5 and 6 and in Tables 19, 20, and 21. This is additional evidence, perhaps, that trained professionals overcome low levels of sleepiness to perform an assigned task.

Although the differences in the mean Samn-Perelli Scores on Rest Days compared with Duty Days presented for the group in Figures 8 and 9 do not appear to be large, the ANOVA results in Table 27 show that the differences were statistically significant.

There was a main effect of the sequence of the Rest and Duty Days on the group means of the Samn-Perelli scores, which requires further study for interpretation.

Table 28 is a summary of the correlations between the Samn-Perelli Scores for each pilot and Time, Rest Days vs. Duty Days, and Sequence of Rest Days or Duty Days.

The subjective ratings of sleepiness were influenced by the time of day for the majority of pilots (16 out of 20). More than half (11 out of 20) showed significant differences in subjective sleepiness between Rest Days and Duty Days, while half of them (10 out of 20) showed significant differences in subjective sleepiness as a function of sequence of days.

All of these results are evidence that the Samn-Perelli scores when considered for the group or for the individual pilots were strongly influenced by any and all of our representations of the work pattern. However, the correlation is stronger with time of day than with the work pattern for most pilots. (See Table 28) This might be interpreted as signifying that the self-perception of fatigue is more strongly influenced by circadian rhythm than by work schedule. However, we need to consider other ways to represent work pattern to confirm these results.

Table 28 - ANOVA on Samn-Perelli scores with work pattern for each pilot

Pilot	Rest Days vs. Duty Days	Sequence of days within Rest/Duty blocks	Time
1	F(1, 62)=2.01	F(6, 62)=1.79	F(9, 62)=1.90
3	F(1, 60)=20.02***	F(6, 60)=2.37*	F(8, 60)=11.32***
4	F(1, 63)=0.11	F(6, 63)=1.06	F(10, 63)=5.79***
5	F(1, 70)=1.03	F(6, 70)=2.23	F(9, 70)=2.27*
6	F1, 67)=27.55***	F(6, 67)=3.75**	F(8, 67)=8.23***
7	F(1, 56)=0.21	F(6, 56)=2.56*	F(9, 56)=2.54*
8	F(1, 66)=10.17**	F(6, 66)=2.75*	F(9, 66)=3.09**
9	F(1, 71)=1.18	F(6, 71)=2.04	F(8, 71)=8.32***
10	F(1, 69)=12.88***	F(6, 69)=2.52*	F(9, 69)=16.11***
12	F(1, 70)=0.04	F(6, 70)=3.41**	F(9, 70)=2.27*
13	F(1, 57)=13.13***	F(6, 57)=2.15	F(8, 57)=7.50***
14	F(1, 76)=13.92***	F(6, 76)=2.78*	F(9, 76)=13.29***
15	F(1, 66)=0.04	F(6, 66)=3.67**	F(9, 66)=2.99**
16	F(1, 57)=7.65**	F(6, 57)=3.77**	F(9, 57)=1.06
17	F(1, 38)=21.64***	F(6, 38)=0.93	F(9, 38)=1.52
18	F(1, 65)=5.36*	F(6, 65)=2.50*	F(11, 65)=1.14
19	F91, 75)=2.53	F(6, 75)=1.34	F(9, 75)=8.85***
20	F(1, 57)=14.63***	F(6, 57)=0.23	F(10, 57)=5.47***
21	F(1, 71)=1.48	F(6, 71)=1.33	F(9, 71)=7.89***
22	F(1, 75)=7.27**	F(6, 75)=2.13	F(9, 75)=4.36***
#sig	11/20	10/20	16/20

*p < .05, **p < .01, ***p < .001

Correlations of Total Sleep Time (TST) with work pattern

The analyses of the TST group data, on which we reported earlier, offered evidence of the strong influence of work pattern on sleep time. See, for example, Figure 16 that shows the average TST for the group of 21 pilots across participation period. This strong effect was further substantiated by the ANOVA of the TST data for the individual pilots.

As shown in Table 29, ANOVA on TST individual-pilot data found statistically significant main effects of Rest Days vs. Duty Days for 14 pilots (out of 20) and of the sequence of Rest Days and Duty Days for 17 pilots (out of 20). The high levels of correlations among 17 of the 20 pilots shown in Table 29 is consistent with the small standard deviations on the mean TST for the group shown in Figure 16.

Table 29 - ANOVA of TST with work pattern for each pilot

Pilot	Rest days vs. Duty days	Sequence of days within Rest/Duty blocks
1	F(1, 81)=3.77	F(6, 81)=17.88***
3	F(1, 77)=1.14	F(6, 77)=3.89**
4	F(1, 77)=0.06	F(6, 77)=5.13***
5	F(1, 82)= 5.81*	F(6, 82)=1.87
6	F(1, 78)=17.59***	F(6, 78)=4.36***
7	F(1, 69)=4.70*	F(6, 69)=11.52***
8	F(1, 80)=0.08	F(6, 80)=4.80***
9	F(1, 78)=30.01***	F(6, 78)=27.58***
10	F(1, 77)=8.47 **	F(6, 77)=6.59***
12	F(1, 79)=26.76***	F(6, 79)=13.32***
13	F(1, 71)=24.44***	F(6, 71)=12.16***
14	F(1, 84)=32.05***	F(6, 84)=3.42**
15	F(1, 79)=14.03***	F(6, 79)=19.29***
16	F(1, 76)=4.96*	F(6, 76)=1.95

17	F(1, 70)=27.52***	F(6, 70)=15.83***
18	F(1, 80)=1.33	F(6, 80)=6.36***
19	F(1, 84)=5.40*	F(6, 84)=16.67***
20	F(1, 74)=11.03**	F(6, 74)=11.10***
21	F(1, 83)=0.36	F(6, 83)=1.28
22	F(1, 81)=7.21**	F(6, 81)=18.13***
#sig	14/20	17/20

*$p < .05$, **$p < .01$, ***$p < .001$

We have found indications of strong influence of work pattern on both the Samn-Perelli Scores and on the TST. This is consistent with the results of the analysis that showed strong correlation between Samn-Perelli Scores and TST. Such consistent results, give us confidence in the quality of the data.

In Appendix E, we present plots of TST for each of the pilots across their roster schedules. Examination of these plots together with the small standard deviations shown on Figure 16 of the Total Sleep Time for the group of 21 pilots across the entire participation period support our previous findings that the majority of pilots follow a similar sleep pattern (i.e., less sleep during the nights before a Duty Day with an Early start and more sleep during nights before Duty Days with Late start and finish).

Summary of major results and conclusions from analyses of data correlation

1. The correlation of group data on Mean PVT Response Times and Samn-Perelli Scores was statistically significant. Correlation analysis of the data for each pilot shows the relation between Samn-Perelli Scores and Mean PVT Response Times was significant for 12 out of 21 pilots. We cannot explain why the Samn-Perelli Scores and the Mean PVT Response Times are well correlated for some pilots and not for others.
2. The subjective measure of fatigue correlated significantly with the objective measure of cognitive performance for the group and for most of the pilots.
3. There is much less variation in the Mean PVT Response Times than in the Samn-Perelli Scores over the same periods of time. The Samn-Perelli Scores seldom exceed 5 and the low levels of sleepiness for most pilots during this study had little effect on their performance. This might be explained by the ability of a well–trained professional to adapt to and overcome low levels of tiredness to perform his or her job acceptably.
4. The evidence is that there was statistically insignificant potential for degraded vigilance performance due to fatigue during this study. Mean PVT Response Times > 400 msec. and Samn-Perelli Scores > 4 were recorded In-flight for only 2.3% of the 818 flights during which data were acquired for this study. This percentage drops to 0.1% if we consider a Samn-Perelli Score of 5 as the critical value rather than 4.
5. The evidence is that the Total Sleep Time and the Samn-Perelli Scores, whether considered for the group or for the individual pilots, were strongly influenced by any and all of our representations of the work pattern.

Our Approach to Modelling

Bio-mathematical models are used to predict level of fatigue, alertness, performance, or risk associated with a particular roster, and can therefore be used as part of safety-related decision-making in a fatigue-management context (Dawson et al. 2011). The goal of predictive fatigue models is to increase the safety and efficiency of crew scheduling. We need to develop a reliable model of fatigue to predict the effects and to assess the safety risks of proposed interventions by using it to explore their impacts prior to implementation.

We will focus on developing bio-mathematical models that take into account individual differences because our data have confirmed the evidence (e.g., Van Dongen et al. 2003a,b, Kandelaars et al. 2006, and Roach et al. 2002) that individuals differ significantly not only in their response to sleep loss and sleep patterns, but also in the consistency of their sleep/wake behavior across similar situations. Our research has shown very large individual differences in reactions to circadian disruption, quantity and quality of sleep, and levels of drowsiness.

The work we have completed in correlation analyses will provide a basis for identifying the main factors of fatigue to be considered in building a reliable model for each pilot. Figure 35 shows the concept for building the model of fatigue for each pilot. All of the data collected relevant to fatigue will be examined for their contributions to causation. We expect to develop a generic structure that will be individualized by appropriately weighting each of these factors. The model will be validated by comparing its predictions to the levels of fatigue measured with BioHarness and Samn-Perelli Scores across the entire roster.

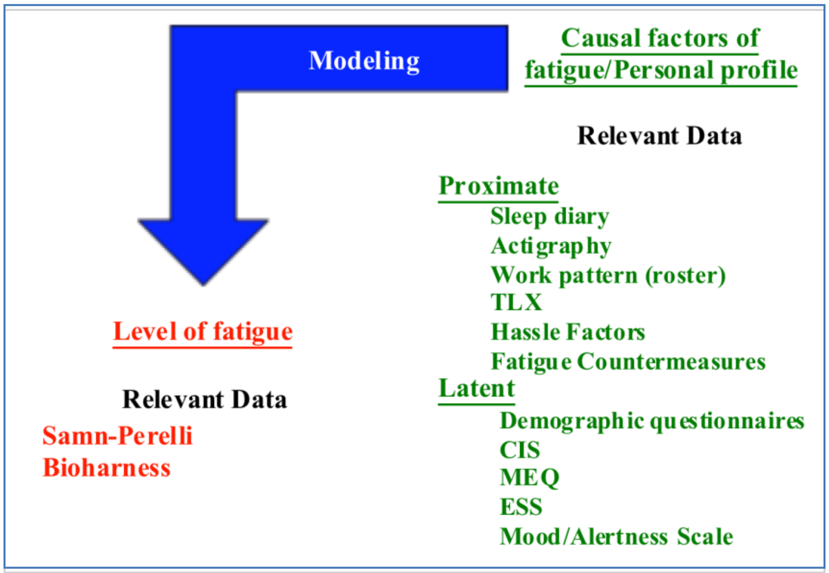


Figure 35 – A Concept for Building a Model of Fatigue for Each Pilot

We have established that circadian rhythm is one of the main individualistic factors of fatigue. Therefore, we have taken a preliminary step toward developing a model for each pilot by determining the circadian cycle for each using his or her Actiwatch, Sleep Diary, and other data averaged over his or her Rest Days. An example of the fit of the data • to the circadian cycle (red curve) is shown in Figure 36 for Pilot #1using his averaged data for the first three Rest Days. The goodness of the fit is 0.019, p< .001. The acrophase (i.e., the upper part of a sine wave fitted to a measurement of a circadian pattern or other biological rhythm), which is a key parameter of such a model, is 15.2 hours in this case. The best-fit curve is only intended to be used to interpolate data for times other than when they were recorded.

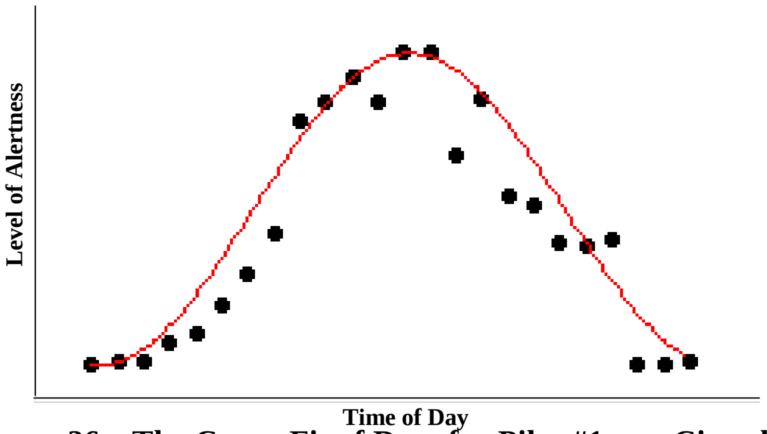


Figure 36 – The Curve Fit of Data for Pilot #1 to a Circadian Cycle

3. Aircraft Performance

Background

NASA Ames Research Center's Intelligent Systems Division's (Code TI) role is to develop and apply algorithms that automatically extract information on anomalous events from the flight-recorded data. The aim is to identify relationships between anomalous events and/or exceedances identified in the flight data and their possible causes identified in the ASR's. Additionally, Code TI will work with Code TH to identify events in flight-recorded data that can serve as surrogates for indications of fatigue in physiological data collected during some flights. The idea is to use these surrogates in the clear majority of flights for which physiological data cannot be collected. The developed algorithms will contribute to NASA and easyJet's joint efforts to develop efficient and reliable methodologies to extract and merge information from large, diverse data sources. The goals of this research are to assist aviation safety analysts to:

1. Identify expected and unexpected operationally significant events or trends in aircraft performance that could compromise the safety of the system.
2. Identify the latent and proximate causal and contributing factors of the events identified to enable data-driven decisions on interventions or mitigations.
3. Identify the events in which flight-crew fatigue was a contributing or causal factor.

Approach to Flight-Data Analysis

As discussed in the previous annual report the purpose of Code TI's work in flight-data analyses is to discover the unexpected events that could compromise the safety of operations to complement and supplement the search for the expected events that is currently typically performed under the Flight Operational Quality Assurance (FOQA) programs. The expected events are usually defined by a single variable that has exceeded a value during a particular phase of flight that is considered outside of the established Standard Operating Procedures (SOP's).

The algorithms developed by Code TI search for sets of continuous parameters and binary switches that contribute to an event that is considered statistically anomalous in a multivariate comparison with normal operations.[4] The automatic identification of the

---

[4] Safe "Normal" operations are not always completely consistent with the SOP's.

contributions of the particular continuous and discrete parameters entailed in the identified anomalous event assist the domain expert in ascertaining its operational significance. The algorithms are designed to process very large data sets (collected at the rate of over 10,000 Kb per flight and about 2.5 Tb per year) in nearly real time. This is achieved by applying the two algorithms discussed in the previous annual report (MKAD and SequenceMiner).

Currently, Airbus and its contractor using its proprietary software called AirFASE are identifying the FOQA-like exceedances for easyJet. Through existing agreements among Airbus, easyJet, and ONERA, ONERA undertook the task of identifying the exceedances in the flight data that were collected for the HFMP study reported here. Through existing agreements among ONERA, easyJet, and NASA, these exceedances were made known to NASA. Code TI personnel, with assistance from subject-matter experts, compared the results of the search for the unexpected events using NASA's anomaly detection algorithms with the exceedances identified by ONERA using AirFASE. As we had expected, there were anomalous events that were identical to those prescribed exceedance events found using AirFASE. The multivariate information obtained using the algorithms for anomaly detection complement the single variable exceedances information and give the safety analyst a better understanding of the event.

There was a second category of events that were identified as exceedances for which no anomalous event was identified. Each of these is likely to have a different explanation. Reasons that we have found in previous studies have been exceedances based on computed parameters within the AirFASE processing that were not recorded and available to the search for anomalies or exceedances that occur so frequently that they are not identified as anomalous events when compared to normal operations. There were also flight data that were analyzed with AirFASE that were not available for analysis with the anomaly-identification algorithms.

The third category was of events that were identified as anomalous but were not found using AirFASE. These events comprise new discoveries and, if deemed operationally significant, could be used to define a pattern to conduct a routine search of past or future flights for similar events.

Subject-matter experts reviewed the events in all of these categories along with the identification of the parameters that caused them to be considered a statistically significant event to identify the ones considered to be operationally significant.

Flight-recorded Numerical Data

NASA has worked to further improve the tools for anomaly discovery and pattern searching techniques. The MKAD algorithm has been extended to be able to scale to very large data sets and has been successfully tested on the equivalent of 10TB of raw CSV data or approximately 940,000 flights. As of March 1st 2012, NASA received approximately 5.86TB of flight data in the amount of 557,907 flights from easyJet over the period of 22 months. During the analysis of these data, two new types of anomalies deemed operationally significant by both the domain experts and easyJet have been identified. These identified anomalous operations are 1) near stall conditions upon takeoff and 2) possible cases of mode confusion during approach. The list of flights found in these two categories can be found in Table 30.

Table 30 - List of anomalies identified by MKAD and SequenceMiner along with similar events.

Flight Name	Anomaly Descriptions

Near Stall	**Flights found to have come within 15 knots of estimated stall speed. (Stall speed is based on gross weight, flap position, and only at takeoff and landing altitudes.)**
FlightData__20100701T092903	6 Kt>stall @500 ft. Flaps retracted early.
FlightData__20110127T223549	7 Kt>stall @2400 ft. Flaps raised at low airspeed.
FlightData__20101014T175030	10 Kt>stall @1950 ft. ~20 sec drop in airspeed
FlightData__20101025T072003	10 Kt>stall @1250 ft. During GA ~15 sec drop in airspeed while starting to climb.
FlightData__20110322T142709	10 Kt>stall @1800 ft. During GA ~12 sec drop in airspeed after starting climb.
FlightData__20100901T102333	11 Kt>stall @1600 ft. Flaps raised early.
FlightData__20100708T170946	12 Kt>stall @3200 ft. ~15 drop in airspeed.
FlightData__20110228T183041	12 Kt>stall @2500 ft. ~60 sec drop in airspeed.
FlightData__20100630T015221	13Kt>stall @2450 ft. ~20 sec drop in airspeed .
FlightData__20101013T005253	13 Kt>stall @3200 ft. ~50 sec drop in airspeed.
FlightData__20101116T073907	13 Kt>stall @3050 ft. ~90 sec drop in airspeed.
FlightData__20110206T171937	13 Kt>stall @3700 ft. ~20 sec drop in airspeed.
FlightData__20110522T083445	13 Kt>stall @3000 ft. ~60 sec drop in airspeed.
FlightData__20101205T232249	14 Kt>stall @1900 ft. ~80 sec drop in airspeed.
FlightData__20101224T125321	14Kt>stall @2500 ft. ~20 sec drop in airspeed.
FlightData__20110221T203916	14 Kt>stall @1250 ft. ~20 sec drop in airspeed.
FlightData__20110621T002935	14 Kt>stall @2800 ft. ~70 sec drop in airspeed.
FlightData__20110721T185744	14 Kt>stall @500 ft. During drop in airspeed.
Mode Confusion	**Flights found to exhibit possible mode confusion with either repeated switching attempts on vertical modes and/or recycling of the flight director on approach to landing.**
FlightData__20100516T121755	Repeated switching attempts on vertical modes and recycling of the flight director.
FlightData__20110131T093429	Repeated switching attempts on vertical modes and recycling of the flight director.
FlightData__20110712T092249	Repeated switching attempts on vertical modes.
FlightData__20110119T180030	Repeated switching attempts on vertical modes.
FlightData__20100711T004807	Repeated switching attempts on vertical modes.
FlightData__20100428T124000	Repeated switching attempts on vertical modes.
FlightData__20110730T152021	Repeated switching attempts on vertical modes.
FlightData__20100527T131333	Repeated switching attempts on vertical modes.
FlightData__20101016T183310	Repeated switching attempts on vertical modes.
FlightData__20110420T150033	Repeated switching attempts on vertical modes.
FlightData__20110508T073513	Recycling of the flight director.
FlightData__20110427T074943	Recycling of the flight director.
FlightData__20111010T150931	Repeated switching attempts on vertical modes.
FlightData__20101221T130819	Recycling of the flight director.

Analysis of Numerical Flight Data during the Fatigue Study

Flight data during the fatigue study experiment (July 4th – August 22nd 2011) were logged and transferred to NASA to be analysed. However, of the 904 flights flown during the study, ONERA was only able to access flight data for 770 flights for application of AirFASE to identify exceedances and NASA was only able to access data for 631 flights for application of MKAD to identify anomalous events. Domain experts validated the events identified as anomalous and, together with exceedance counts from AirFASE, these events were correlated with crew performance data as discussed in Section 4.

ONERA Validation of MKAD tool on FOQA data

In the technical report by Dr. Nicolas Maille titled, *"Improving analysis methodologies for flight data: lessons provided by the study of operational flights"* January 2012 produced by ONERA, the results found by MKAD are compared with the current state of the art algorithm for discovery of anomalous events (The Morning Report). A number of anomalous events are listed in the report including a few described in detail: 1) a holding pattern with an unusual LOC* capture sequence, 2) high rate of descent below 2000 ft., and 3) possible mode confusion with unusual vertical, lateral, and ATS mode changes. Additionally, other groups of examples were found and described as having anomalous contributions from the following sets of parameters:

- The aircraft configuration setting: landing without a full configuration (these flights are identified in AirFASE with a low severity event “Landing with incorrect flap setting”) or non-standard changes (from configuration 2 to full configuration).
- Unusual use of lateral or longitudinal modes on the autopilot: Vertical speed, Altitude capture, and Track modes.
- The use of speed brakes: Continuous parameters and particularly Rudder, Elevator, Glide, Lateral acceleration…

4. Correlation of Crew Performance with Aircraft Performances

A primary goal of assessing potential causal and contributing factors of fatigue using the data that have been collected in this study is to determine if there is reliable evidence demonstrating that crew fatigue is a causal or primary contributing factor in the occurrence of particular exceedances or anomalies in aircraft performance. We examined the association of the aircraft performance during the course of the roster schedule with the individual performance of the crew on that flight as measured by the Mean PVT Response Times. All data had been time stamped to enable such linkage. It should be noted that not all anomalies or exceedances found in the aircraft performance data will be due to, or even associated with, decrement of human performance and that degradation of human performance will not always cause an anomaly or an exceedance in aircraft performance.

Effects on Aircraft Performance

Aircraft performance is based on the occurrence of unwanted events found by analyses of the in-flight-recorded data for the flights flown by 20 of the pilots[5]. Analysis of the data on the crew performance and aircraft performance on a flight-by-flight basis showed absolutely no correlation between the Mean PVT Response Times and the frequencies of occurrences of anomalies during takeoff, anomalies during landings, exceedances of severity level 1, exceedances of severity level 2, exceedances of severity level 3, or the total number of events for each flight. As the relationship between crew performance and aircraft performance was such a fundamental part of this study, we examined the data for each of the flights in detail to corroborate this finding.

For example, we studied the 19 flights during which the Mean PVT Response Times were >400 msec. and the Samn-Perelli Scores were >4 in flight that we discussed previously. As shown in Table 31, we found no instance in which these circumstances were associated with an inordinate number of anomalies during takeoff, anomalies during landings, exceedances of severity level 1, exceedances of severity level 2, exceedances of severity level 3, or the total number of events for each flight. In all of the 19 flights, there was 1 anomalous event identified, 63 exceedances of severity level 1, 9 exceedances of severity level 2, and only 7 exceedances of severity level 3. These are well within the ranges of the events experienced in the other flights of this study during which the recorded Mean PVT Response Times were considerably less than 400 msec. and the Samn-Perelli Scores were substantially lower than 4. The average number of

---

[5] Pilots #1 and #3 were omitted because they flew only B-737 on which the flight data were not collected.

total events (i.e., exceedances plus anomalies) per flight across all the flights flown by the 20 pilots was 4.63 with a standard deviation of 1.76. For the 19 flights in Table 31, the mean was even smaller; 4.37 events per flight with a standard deviation of 1.86. We studied the nature of the anomalous events and exceedances experienced by these flights and found nothing to suggest that they might have been especially related to fatigue of the flight crew.

Table 31 – Comparisons of Crew Performance and Aircraft Performance

Flight					Crew Performance		Aircraft Performance					
Flt. No.	Date	Pilot	Duty Day	Sector	Mean PVT Response Time	Samn-Pirelli Score	Anomalies		Exceedances			Total Events
							During Takeoff	During Landing	Level 1	Level 2	Level 3	
3478	8-20	6	D13	4th	402.18	5	0	0	4	0	1	5
3010	8-22	6	D15	3rd	412.61	5	0	0	0	0	0	0
3032	7-5	7	D1	4th	405.81	5	0	0	4	1	0	5
3104	7-13	7	D6	4th	419.43	6	0	0	2	0	0	2
8580	8-3	10	D13	4th	441.25	5	No Data		2	1	0	3
5258	8-5	10	D15	2nd	416.98	5	0	0	1	1	0	2
5008	7-20	11	D5	2nd	468.05	5	0	0	5	0	0	5
8536	7-27	11	D9	4th	423.41	5	No Data		3	0	0	3
5498	7-31	11	D11	4th	411.67	5	0	0	4	1	2	7
5360	8-3	11	D14	2nd	410.00	5	0	0	5	0	0	5
3322	7-19	13	D5	2nd	410.00	5	0	0	4	0	0	4
2162	7-29	14	D10	3rd	413.90	5	0	0	5	0	1	6
2442	8-1	14	D11	4th	469.28	5	0	0	4	0	1	5
18	8-2	14	D12	3rd	414.25	5	0	0	2	2	0	4
2161	8-3	14	D13	1st	406.54	5	No Data		3	3	2	8
2197	8-3	14	D13	2nd	427.14	5	No Data		3	1	0	4
186	8-5	14	D15	4th	405.00	5	0	0	5	0	0	5
2123	7-28	20	D11	1st	400.89	5	0	1	5	0	0	6
1318	7-30	20	D13	3rd	445.78	5	0	0	4	0	0	4

It is important to note that the absence of statistically significant correlations between the PVT data and anomaly- and exceedance-related events cannot be attributed to an absence of statistical power in the data. As reported previously we found a number of statistically significant effects, such as work pattern effects on Samn-Perelli Scores and on Mean PVT Response Times; sleep effects on Samn-Perelli Scores; and significant correlations between Samn-Perelli Scores and Mean PVT Response Times. These statistical analyses showed that the collected data had enough statistical power to detect fairly small effects. The absence of correlation between Mean PVT Response Times and aircraft performance events is a reliable statistical conclusion.

We noted above that highest levels of sleepiness encountered during this study were not related to high Mean PVT Response Times. Research reported in the literature has shown that highly trained professionals adapt to and overcome moderate levels of sleepiness. This seems to be especially true when they are operating as a team like flight crews. In a report of study on fatigue and decision-making in a commercial airline environment (Foushee et al., 1986), the authors said that, despite nominally fatigued crews reporting less sleep and higher levels of subjective fatigue associated with recent duty history compared to rested crews, somewhat paradoxically, these crews achieved better overall performance than low-fatigued crews. A report of a simulation study by Petrilli et al 2006, states that all pilots employed performance protection strategies during normal flight operations to cope with fatigue. When pilots were asked whether performance protection strategies were used during the simulator experiments of this study, over 80% of non-rested crews stated they had employed these strategies, whereas only about 20% of rested crews stated they had employed these strategies.

The data collected by this group of pilots in this limited study show that a pilot who has assessed his or her sleepiness at a score of 5.0 (or even greater) is as likely (or unlikely) to exhibit degraded performance (i.e., cause unwanted events) in operating his or her aircraft as some other pilot who has assessed his or her level of fatigue at a value less than 5.0. In any case, (according to these data) an average Samn-Perelli score of greater than 5.0 is not necessarily a safety risk. Nor is a score of less than 5.0 assurance of zero safety risk.

We have found no convincing evidence in the data collected in this study that would permit us to state that the unwanted events experienced by the aircraft flown during this study were associated with degraded vigilance performance of a crewmember due to sleepiness. This does not answer the question of whether higher levels of fatigue would show such a relationship, since the fatigue levels of the pilots in this study remained relatively low throughout.

5. Major Findings and Conclusions

Since the issuance of the first report of the collaborative study under the NASA-easyJet Agreement, we have made significant progress. We completed the review of all of the data and found them adequate in quality and quantity to perform reliable analyses. We found that, on the average for the group of pilots, neither the levels of sleepiness nor the degradations of performance reached disturbing levels at any time during the study. However, individual differences are significant. An important conclusion is that this group of pilots experienced degradation of performance very infrequently in flight during the Duty Days as flown during this study.

A significant finding was that the Samn-Perelli Scores (a subjective measure of fatigue) correlated with the Mean PVT Response Times (an objective measure of cognitive performance) for the group and for 12 of 21 pilots.

Since the Samn-Perelli Scores seldom exceed 5, the fairly low levels of sleepiness for most pilots during this study seem to have had little effect on their performance as measured by their Mean PVT Response Times. This might be explained by the ability of a well–trained professional to adapt to and overcome low levels of tiredness to perform his or her job acceptably. The evidence is that there was statistically insignificant potential for degraded vigilance performance due to fatigue during this study.

Finally, we have found no evidence to say that any of the unwanted events experienced by the aircraft flown during this study were related to degraded vigilance performance of a crewmember due to sleepiness.

Although the small group of pilots who volunteered to participate in this study cannot be considered representative of the general population of easyJet pilots and although the individualistic differences are important, we believe that most of the results and conclusions we have reached so far apply to any group of similarly trained pilots flying the same FRV roster and comparable flights as those flown in this study.

We also offered a few recommendations. We suggested that pilots be encouraged to use their Rest Days, including the first 3 days in this roster, for proper and adequate sleep. Based on our review of the data provided by the pilots on their workloads and hassle factors, we suggested that an update in their CRM training seems appropriate and that, in routine flights, at least one member of the flight crew should have had experience at the destination airport especially if it a foreign-language ATC.

6. References

Basner M. and Dinges, D.F. 2011. Maximizing sensitivity of the psychomotor vigilance test (PVT) to sleep loss. *SLEEP* 34(5): 581-591.

Beurskens, Anna J.H.M.; Bültmann, Ute; Kant, IJmert; Vercoulen, Jan H.M.M.; Bleijenberg, Gijs; and Swaen, Gerard M.H. 2000. Fatigue among working people: validity of a questionnaire measure. *Occup Environ* Med; 57: 353–357.

Civil Aviation Authority. 2005. *Aircrew fatigue: a review of research undertaken on behalf of the UK Civil Aviation Authority*. Norwich, UK: The Stationery Office; CAA Paper 2005/04.

Chua, Eric Chern-Pin; Tan, Wen-Qi; Yeo, Sing-Chen; Lau, Pauline; Lee, Ivan; Ho Mien, Ivan; Puvanendran, Kathiravelu; and Gooley, Joshua J. 2012. Heart Rate Variability Can Be Used to Estimate Sleepiness-related Decrements in Psychomotor Vigilance during Total Sleep Deprivation. *SLEEP*, Vol. 35, No. 3.

Dawson, D.; Noy, I.Y.; Härmä, M.; Åkerstedt, T.; and Belenky, G. 2011. Modeling fatigue and the use of fatigue models in work settings. *Accident Analysis and Prevention* 43, 549–564.

Dorian, J.; Rogers, N.L.; and Dinges, D. F. 2005. Psychomotor Vigilance Performance: Neurocognitive Assay Sensitive to Sleep Loss. In *Sleep deprivation: clinical issues, pharmacology, and sleep loss effects.* Ed. Clete Kushida, New York: Marcel Deckker.

Dinges, David F. and Powell, John W. 1985. Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. *Behavior Research Methods, Instruments, & Computers*, 17 (6), 652-655.

Foushee, C. H.; Lauber, J. K.; Baetge, M. M.; and Acomb, D. 1986. *Crew factors in flight operations: III. The operational significance of exposure to short-haul air transport operations.* Moffett Field, CA NASA Ames Technical Memorandum No. 88322.

Hart, S.G. and Wickens, C.D. 1990. Workload Assessment and Prediction. In HR Booher (Ed). *Manprint: An integrated approach to systems integration* (pp. 257-296). New York: Van Nostrand.

Johns, M.W. 1991. A new method for measuring daytime sleepiness: the Epworth Sleepiness Scale. *Sleep,* p 540-545

Kandelaars, K.J.; Fletcher, A.; Dorrian, J.; Baulk, S.D.; and Dawson, D. 2006. Predicting the timing and duration of sleep in an operational setting using social factors. *Chronobiology International* 23, 1265–1276.

Loh, S.; Lamond, N.; Dorrian, J.; Roach, G.; and Dawson, J. 2004. The validity of psychomotor vigilance tasks of less than 10-minute duration. *Behavior Research Methods, Instruments, & Computers,* 36:2, 339-346.

Petrilli, Renée M.; Thomas, Matthew J.W.; Dawson, Drew; and Roach, Gregory D. 2006. *The decision-making of commercial airline crews following an international pattern*. Centre for Sleep Research, University of South Australia.

Queensland Government. 2009. *Fatigue Risk Management System Resource Pack.* Queensland Health. ISBN 978-1-921447-57-0.

Roach, G.D.; Rogers, M.; and Dawson, D. 2002. Circadian adaptation of aircrew to transmeridian flight. *Aviation, Space and Environmental Medicine* 73, 1153–1160.

Samn, Sherwood W. and Perelli, Layne P. 1982. *Estimating Aircrew Fatigue: A Technique With Application To Airlift Operations.* USAF School of Aerospace Medicine, Aerospace Medical Division, Brooks Air Force Base, TX Report SAM-TR- 82-21.

Stewart, S. 2009. Human Factors Monitoring Program**:** Fatigue Risk Management Scientific Study Methodology. *easyJet Airline Report*, December 3, 2009.

Tucker A.M.; Basner R.C.; Stern Y.; and Rakitin B.C. 2009. The variable response-stimulus interval effect and sleep deprivation: an unexplored aspect of psychomotor vigilance task performance. *SLEEP,* 32(10): 1393-1395.

Van Dongen, H.P.; Rogers, N.L.; and Dinges, D.F. 2003a. Sleep debt: theoretical and empirical issues. *Sleep and Biological Rhythms* 1, 5–13.

Van Dongen, H.P.; Baynard, M.D.; Maislin, G.; and Dinges, D.F. 2003b. Systematic interindividual differences in neurobehavioral impairment from sleep loss: evidence of trait-like differential vulnerability. *Sleep* 27, 423–433.

Vercoulen, J.H.; Swanink, C.M.; Fennis, J.F.; Galama, J.M.; van der Meer, J.W.; and Bleijenberg, G. 1994. Dimensional assessment of chronic fatigue syndrome. *J Psychosom Res.* 38(5): 383-92

Appendix A: Questionnaire for the Morning/Eveningness Scale

1. Considering only your own 'feeling best' rhythm, at what time would you get up if you were entirely free to plan your day?

Enter an x in the box next to the answer you most agree with

5:00am - 6:30am	
6:30am - 7:45am	
7:45am - 9:45am	
9:45am - 11:00am	
11:00am - midday	

2. Considering only your own 'feeling best' rhythm, at what time would you go to bed if you were entirely free to plan your day?

8:00pm - 9:00pm	
9:00pm - 10:15pm	
10:15pm - 12:30am	
12:30am - 1:45am	
1:45am - 3:00am	

3. If there is a specific time at which you have to get up in the morning, to what extent are you dependent on being woken up by an alarm clock?

Not at all dependent	
Slightly dependent	
Fairly dependent	
Very dependent	

4. Assuming adequate environmental conditions, how easy do you find getting up in the mornings?

Not at all easy	
Not very easy	
Fairly easy	
Very easy	

5. How alert do you feel during the first half hour after having woken in the morning?

Not at all alert	
Slightly alert	
Fairly alert	
Very alert	

6. How is your appetite during the first half hour after having woken in the morning?

Very poor	
Fairly poor	
Fairly good	
Very good	

7. During the first half hour after having woken in the morning, how tired do you feel?

Very tired	
Fairly tired	
Fairly refreshed	
Very refreshed	

8. When you have no commitments the next day, at what time do you go to bed compared to your usual bedtime?

Seldom or never later	
Less than one hour later	
1-2 hours later	
More than 2 hours later	

9. You have decided to engage in some physical exercise. A friend suggests that you do this one hour twice a week and the best time for him is between 7:00am - 8:00am. Bearing in mind nothing else but your own 'feeling best' rhythm, how do you think you would perform?

Would be on good form	
Would be on reasonable form	
Would find it difficult	
Would find it very difficult	

10. At what time in the evening do you feel tired and in need of sleep?

8:00pm - 9:00pm	
9:00pm - 10:15pm	
10:15pm - 12:45am	
12:45am - 2:00am	
2:00am - 3:00am	

11. You wish to be at your peak for a task which you know is going to be mentally exhausting and last for 2 hours. If you are entirely free to plan your day, when would you do this task?

8:00am - 10:00am	
11:00am - 1:00pm	
3:00pm - 5:00pm	
7:00pm - 9:00pm	

12. If you went to bed at 11:00pm, at what level of tiredness would you be?

Not at all tired	
A little tired	
Fairly tired	
Very tired	

13. For some reason you have gone to bed several hours later than usual, but there is no need to get up at any particular time the next morning. Which ONE of the following events are you most likely to experience?

Wake up at the usual time and not go back to sleep	
Wake up at the usual time and doze	
Wake up at the usual time and go back to sleep	
Will NOT wake up later than usual	

14. One night you have to remain awake between 4:00am and 6:00am in order to carry out a night watch. You have no commitments the next day. Would you...

Not go to bed until 6:00am	
Nap before 4:00am and sleep after 6:00am	
Sleep before 4:00am and nap after 6:00am	
Sleep before 4:00am and remain awake after 6:00am	

15. You have to do two hours of hard physical work. Which hours would you prefer to do it between?

8:00am - 10:00am	
11:00am - 1:00pm	
3:00pm - 5:00pm	
7:00pm - 9:00pm	

16. You have decided to engage in hard physical exercise. A friend suggests that you do this for one hour twice a week and the best time for him is between 10:00pm and 11:00pm twice each week. How well do you think you would perform?

Would be on good form	
Would be on reasonable form	
Would find it difficult	
Would find it very difficult	

17. Suppose that you can choose your own work hours, but had to work five hours in the day. Which FIVE consecutive hours would you choose?

11:00pm - 4:00am	

3:00am - 8:00am	
6:30am - 11:30am	
9:30am - 2:30pm	
12:30pm - 5:30pm	
5:30pm - 10:30pm	

18. At what time of day do you feel your best?

midnight - 5:00am	
5:00am - 9:00am	
9:00am - 11:00am	
11:00am - 5:00pm	
5:00pm - 10:00pm	
10:00pm - midnight	

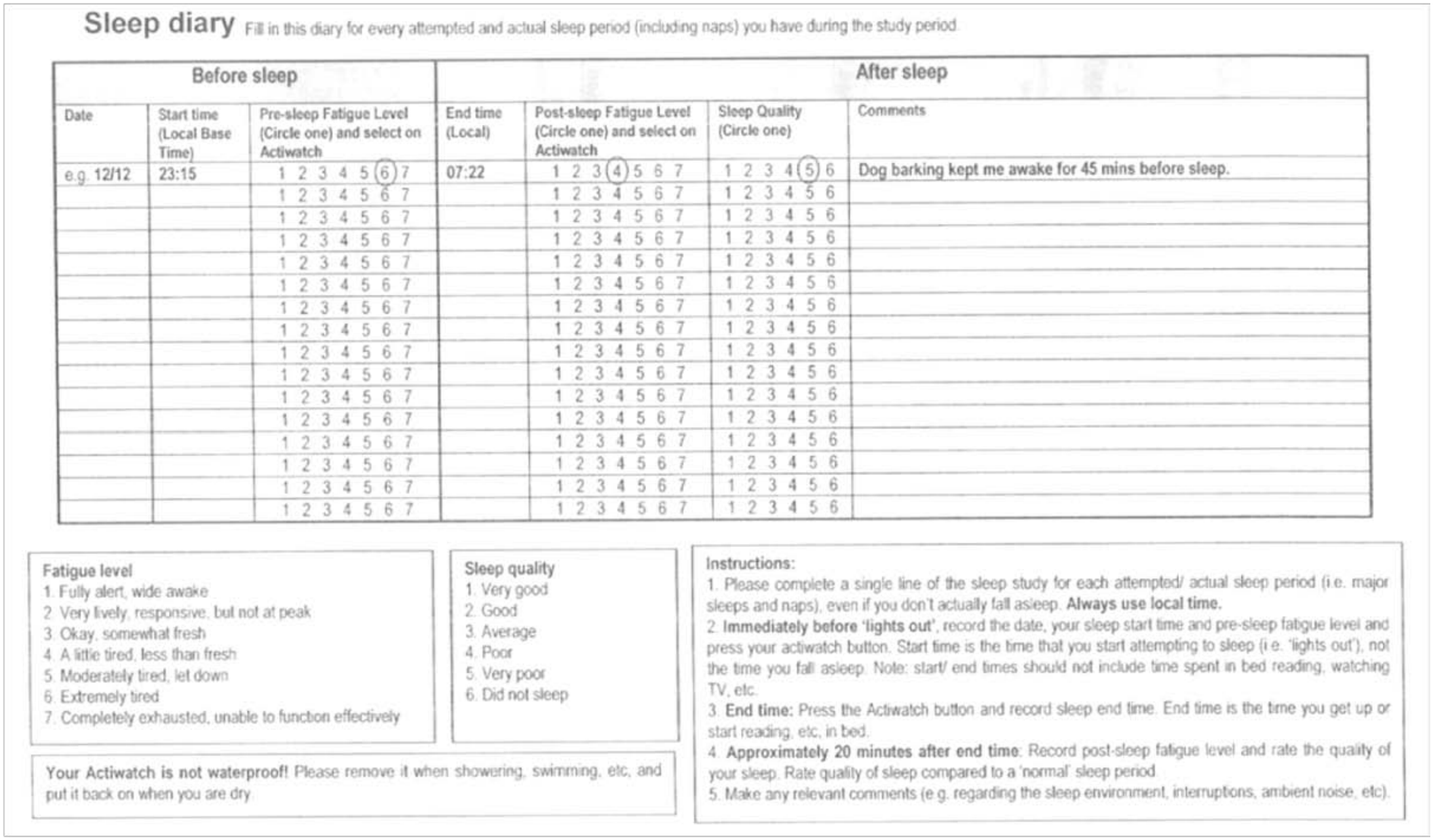
19. Do you consider yourself to be a "morning" or "evening" type of person?

Morning	
More morning than evening	
More evening than morning	
Evening	

Appendix B: Mood/Alertness Scale

SCALES											
MOOD STATES	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
MOTIVATION TO PERFORM: A state of energy or drive that incites or directs an individual's behavior toward achievement of a goal. In this case, motivation to perform this test battery.	BORED: Disinterested, disinclined to effort, lack of enthusiasm.										INTERESTED: Curiosity, fascination, or absorption which holds ones attention.
AROUSAL STATE: Level of physiological activation, state of alertness	SLEEPY: A state of drowsiness, lethargy, inclination to sleep.										ALERT: A state of careful watchfulness marked by ready perception and promptness in perceiving, evaluating and responding, vigilantly attentive.
TENSION LEVEL	TENSE: A state of emotional strain.										RELAXED: Released from nervous tension, at ease, loose.
FEELING	SAD: A state of grief, unhappiness or sorrow.										HAPPY: A state of agreeable emotion, ranging from contentment to intense joy indicative of gladness or delight.
FATIGUE LEVEL: Feelings of bodily discomfort and aversion to effort, a motivated change of attention away from work.	WEARY: A state of being worn out, mentally or physically exhausted.										ENERGETIC: Operating with vigor, active, possessing energy.
EASE OF CONCENTRATION: Bringing all effort and faculties to bear on an activity, exclusive attention.	VERY LOW										VERY HIGH
PHYSICAL DISCOMFORT: A state of physical uneasiness or mild aches or pains.	VERY HIGH										VERY LOW
CONTENTEDNESS: A state of being contented with your situation.	UNPLEASANT										PLEASANT
SLEEP PREVIOUS NIGHT											
TROUBLE FALLING ASLEEP	VERY MUCH										NONE AT ALL
NUMBER OF WAKING EPISODES: Total episodes (0-6)											

Appendix C: Sleep Diary



Sleep diary

Fill in this diary for every attempted and actual sleep period (including naps) you have during the study period.

Before sleep			After sleep			
Date	Start time (Local Base Time)	Pre-sleep Fatigue Level (Circle one) and select on Actiwatch	End time (Local)	Post-sleep Fatigue Level (Circle one) and select on Actiwatch	Sleep Quality (Circle one)	Comments
e.g. 12/12	23:15	1 2 3 4 5 [✓ 6] 7	07:22	1 2 3 [✓ 4] 5 6 7	1 2 3 4 [✓ 5] 6	Dog barking kept me awake for 45 mins before sleep.
		1 2 3 4 5 6 7		1 2 3 4 5 6 7	1 2 3 4 5 6	
		1 2 3 4 5 6 7		1 2 3 4 5 6 7	1 2 3 4 5 6	
		1 2 3 4 5 6 7		1 2 3 4 5 6 7	1 2 3 4 5 6	
		1 2 3 4 5 6 7		1 2 3 4 5 6 7	1 2 3 4 5 6	
		1 2 3 4 5 6 7		1 2 3 4 5 6 7	1 2 3 4 5 6	
		1 2 3 4 5 6 7		1 2 3 4 5 6 7	1 2 3 4 5 6	
		1 2 3 4 5 6 7		1 2 3 4 5 6 7	1 2 3 4 5 6	
		1 2 3 4 5 6 7		1 2 3 4 5 6 7	1 2 3 4 5 6	
		1 2 3 4 5 6 7		1 2 3 4 5 6 7	1 2 3 4 5 6	
		1 2 3 4 5 6 7		1 2 3 4 5 6 7	1 2 3 4 5 6	
		1 2 3 4 5 6 7		1 2 3 4 5 6 7	1 2 3 4 5 6	
		1 2 3 4 5 6 7		1 2 3 4 5 6 7	1 2 3 4 5 6	
		1 2 3 4 5 6 7		1 2 3 4 5 6 7	1 2 3 4 5 6	
		1 2 3 4 5 6 7		1 2 3 4 5 6 7	1 2 3 4 5 6	
		1 2 3 4 5 6 7		1 2 3 4 5 6 7	1 2 3 4 5 6	

Fatigue level

1. Fully alert, wide awake
2. Very lively, responsive, but not at peak
3. Okay, somewhat fresh
4. A little tired, less than fresh
5. Moderately tired, let down
6. Extremely tired
7. Completely exhausted, unable to function effectively

Sleep quality

1. Very good
2. Good
3. Average
4. Poor
5. Very poor
6. Did not sleep

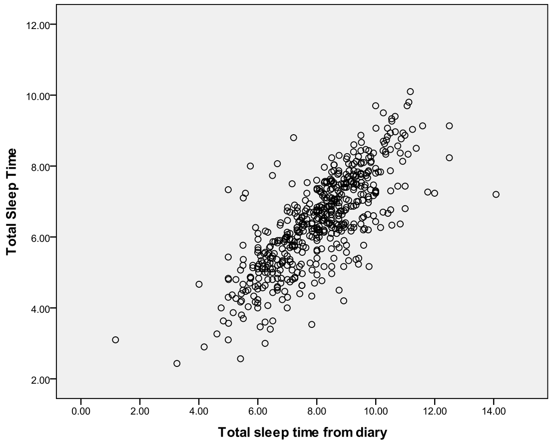
Your Actiwatch is not waterproof! Please remove it when showering, swimming, etc, and put it back on when you are dry.

Instructions:

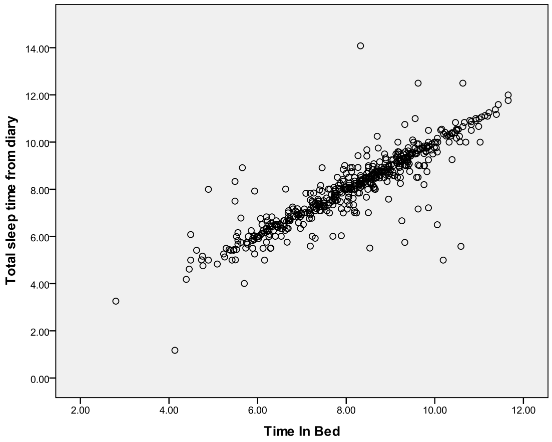
1. Please complete a single line of the sleep study for each attempted/ actual sleep period (i.e. major sleeps and naps), even if you don't actually fall asleep. **Always use local time.**
2. **Immediately before 'lights out'**, record the date, your sleep start time and pre-sleep fatigue level and press your actiwatch button. Start time is the time that you start attempting to sleep (i.e. 'lights out'), not the time you fall asleep. Note: start/ end times should not include time spent in bed reading, watching TV, etc.
3. **End time:** Press the Actiwatch button and record sleep end time. End time is the time you get up or start reading, etc, in bed.
4. **Approximately 20 minutes after end time**: Record post-sleep fatigue level and rate the quality of your sleep. Rate quality of sleep compared to a 'normal' sleep period.
5. Make any relevant comments (e.g. regarding the sleep environment, interruptions, ambient noise, etc).

Appendix D: Correlations between TST and TIB from Actiwatch (TST) and TST from Sleep Diary

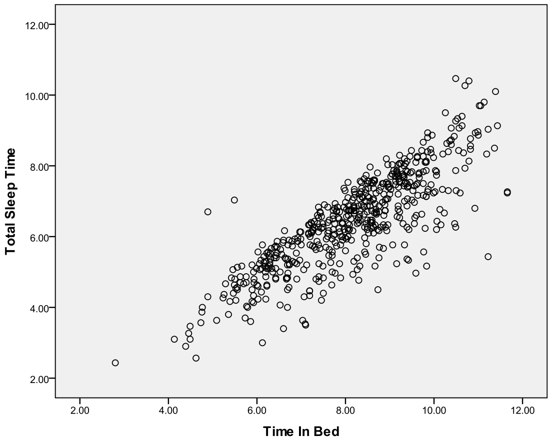
a. Correlation between TST for Actiwatch (TST) and TST from sleep diary;
b. Correlation between TST from sleep diary and TIB from Actiwatch;
c. Correlation between TST from Actiwatch and TIB from Actiwatch



a



b

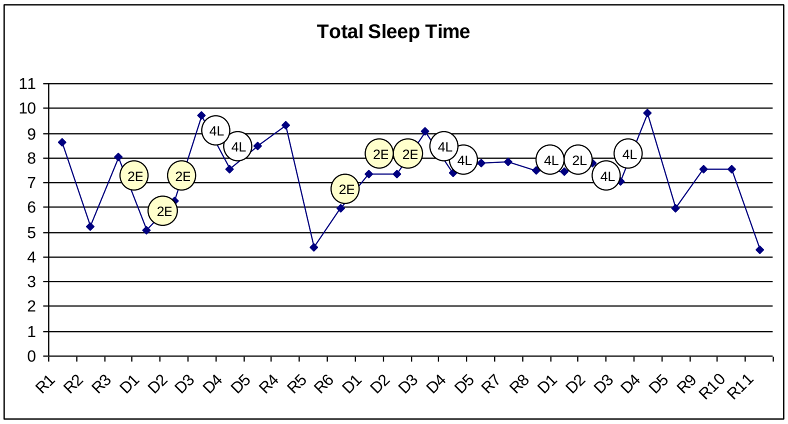


c

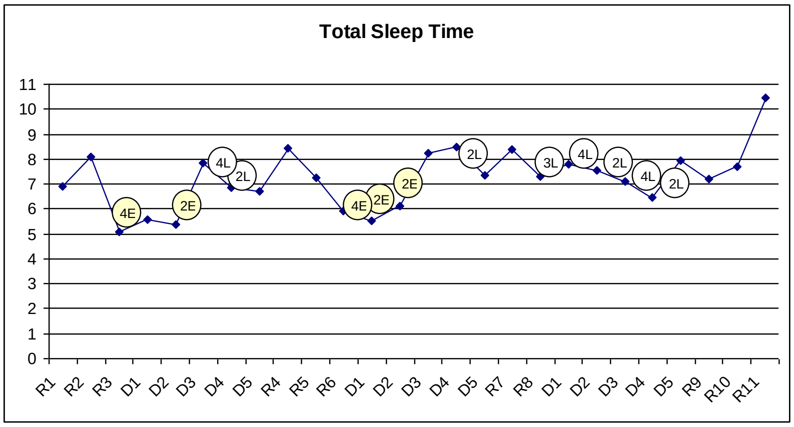
Appendix E: TST for each pilot across the participation period

Indicated in each figure are early start days (E) and late finish days (L) together with number of sectors flown that day. For example, 2E means an early start day with 2 sectors or 4L is a late finish day with 4 sectors. The TST is shown in hours of sleep.

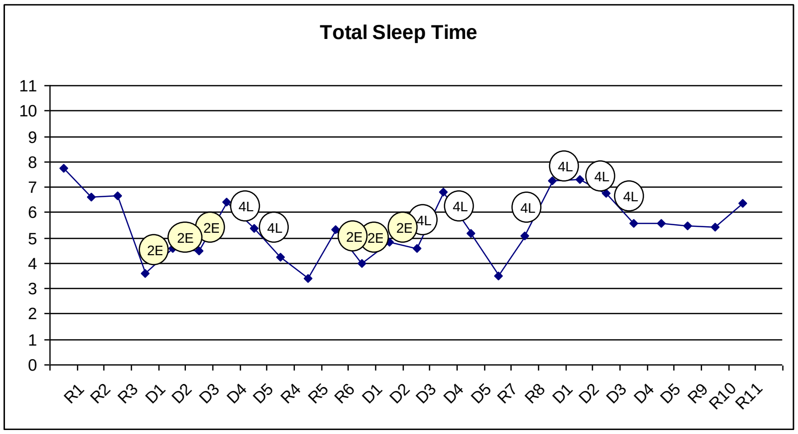
Pilot 1



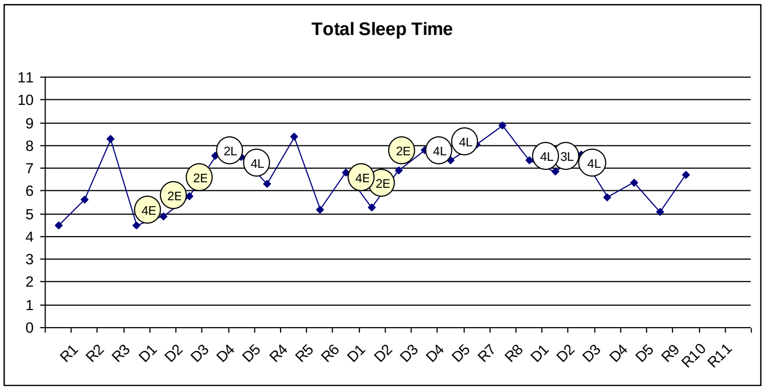
Pilot 2



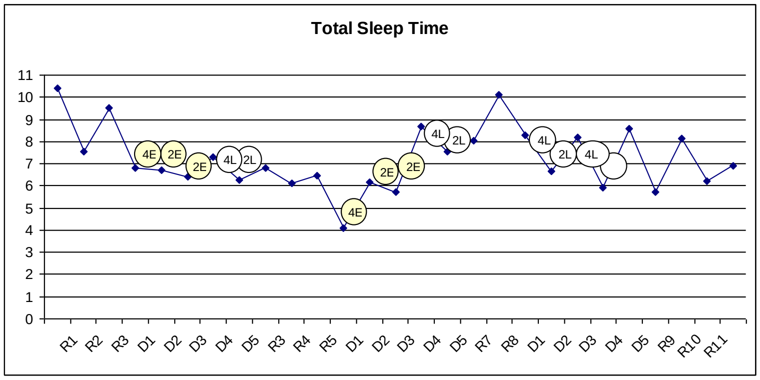
Pilot 3



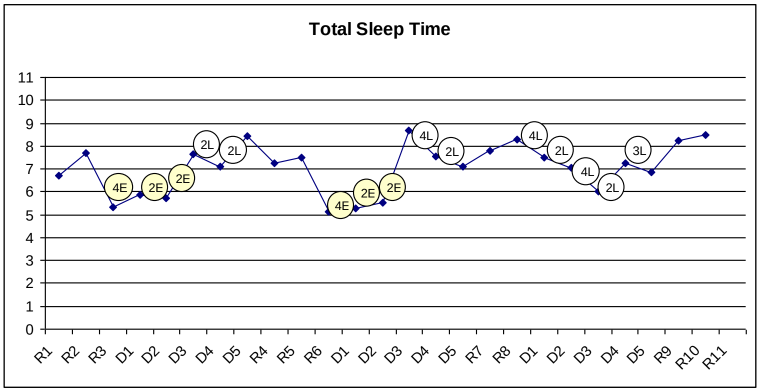
Pilot 4



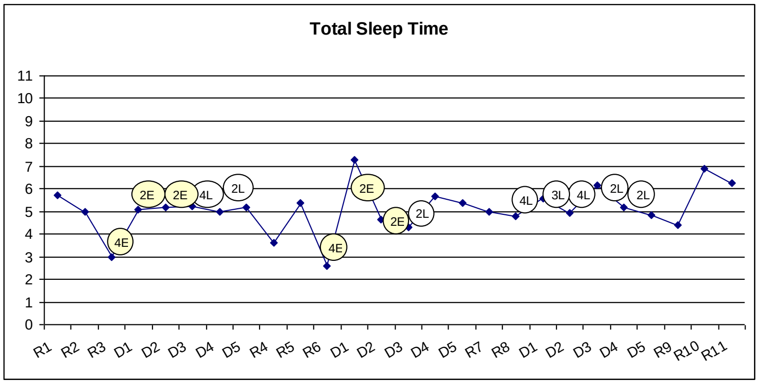
Pilot 5



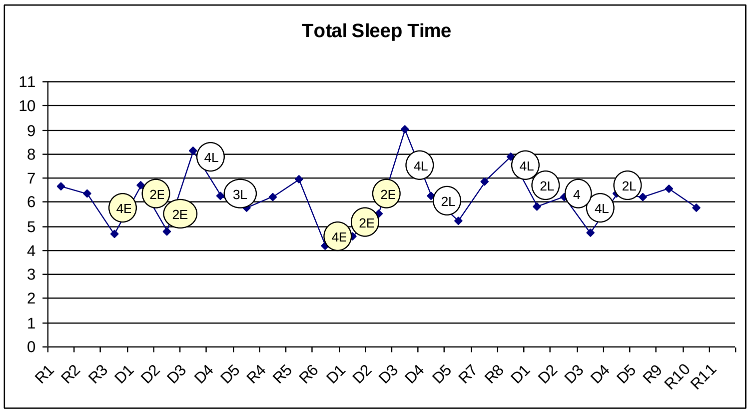
Pilot 6



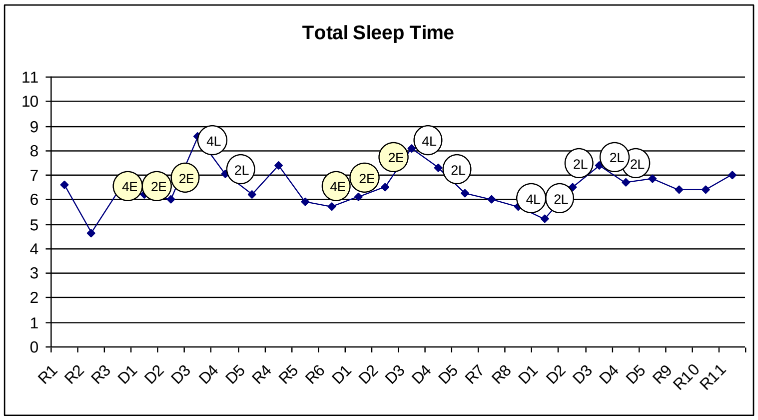
Pilot 7



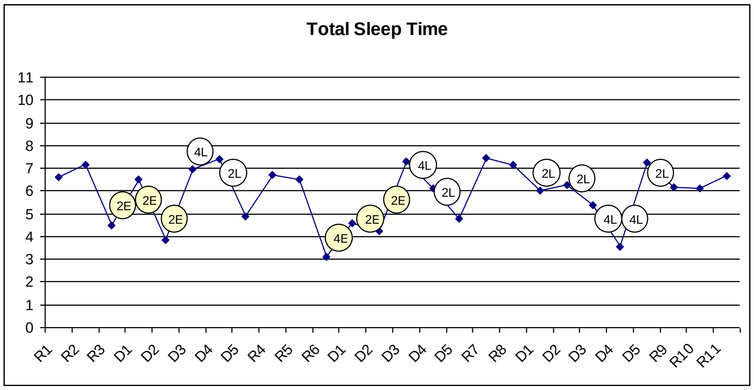
Pilot 8



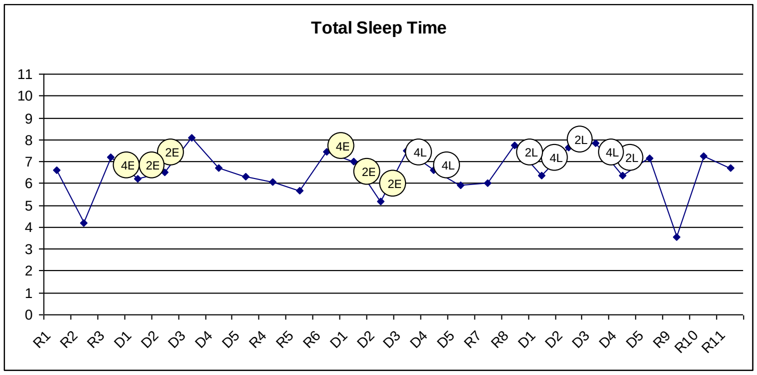
Pilot 9



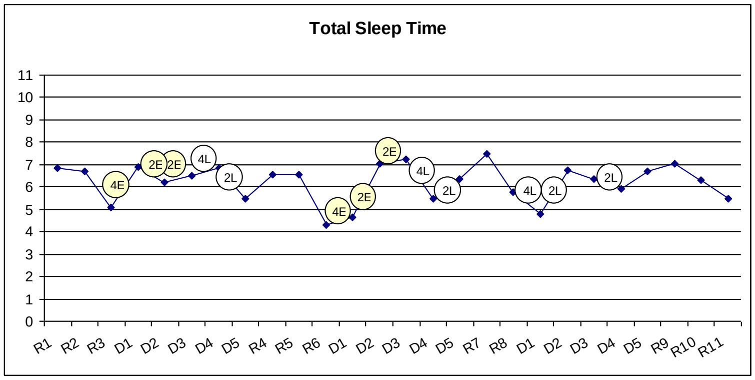
Pilot 10



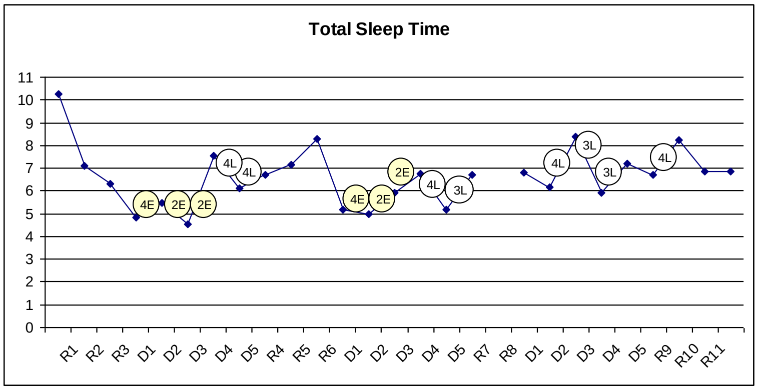
Pilot 12



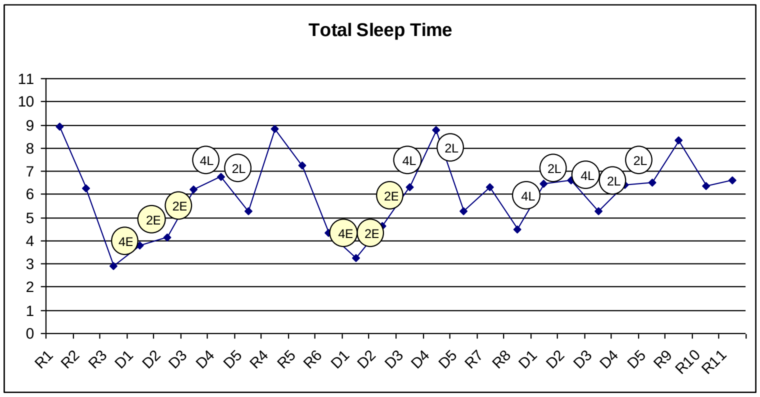
Pilot 13



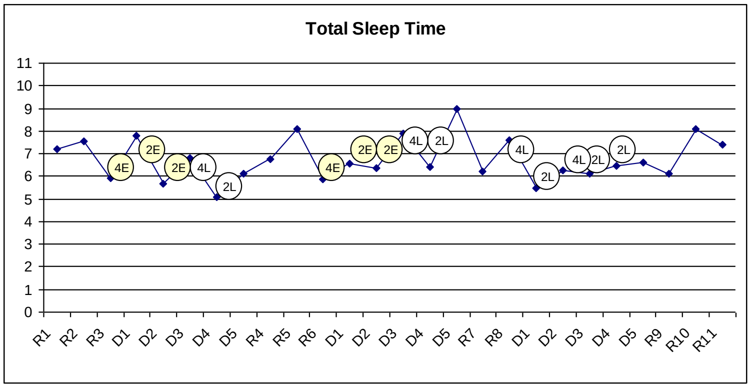
Pilot 14



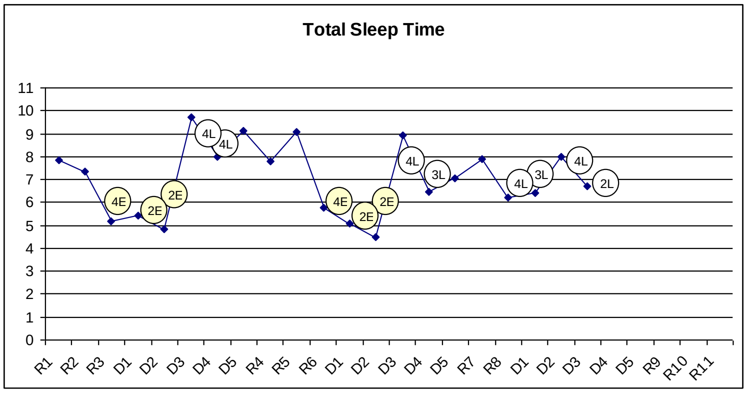
Pilot 15



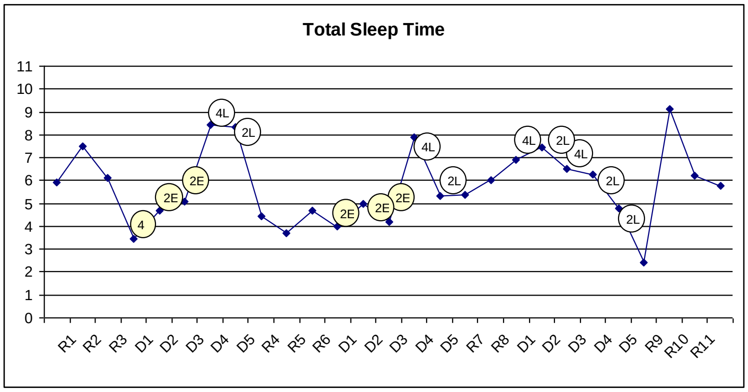
Pilot 16



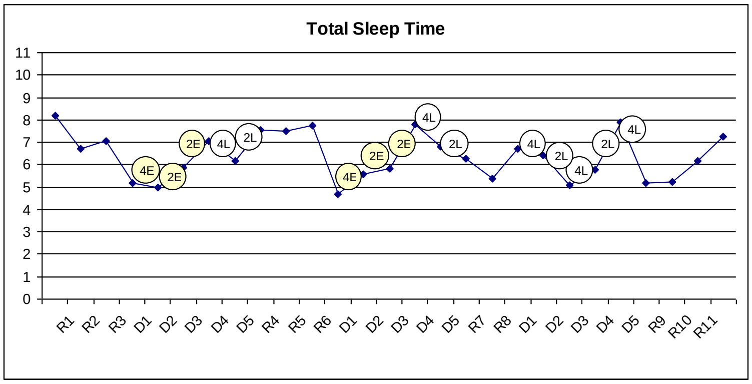
Pilot 17



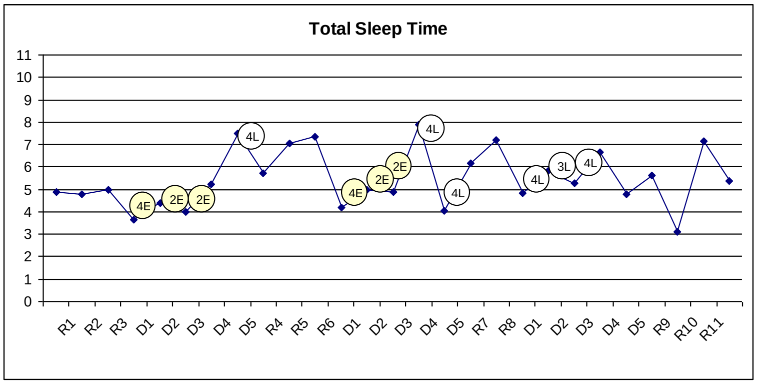
Pilot 18



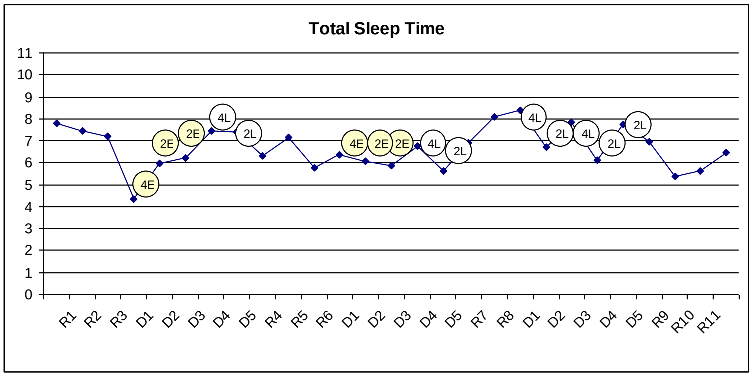
Pilot 19



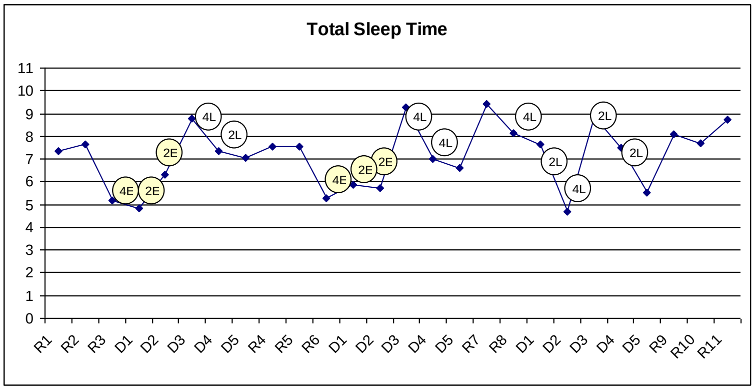
Pilot 20



Pilot 21



Pilot 22



Appendix F: Descriptive Statistics for Sleep Diary Parameters for each Pilot during Rest Days and Duty Days (Blocks A, B, C)

Pilot	Block		Pre-sleep Fatigue Level	Post-sleep Fatigue Level	Sleep Quality	Total Sleep Time
1	**Rest Days**	M	5.08	3.25	3.00	7.97
		SD	1.00	1.48	1.13	1.86
	Block A	M	5.40	2.80	2.40	9.02
		SD	.55	.84	1.14	2.30
	Block B	M	5.20	2.40	2.00	9.20
		SD	.84	.55	.71	.83
	Block C	M	5.20	2.60	1.80	9.47
		SD	.84	.89	1.10	1.06
2	**Rest Days**	M	3.80	2.80	2.60	8.94
		SD	.63	1.40	1.51	1.47
	Block A	M	4.80	3.20	1.40	7.72
		SD	.84	.45	.55	1.50
	Block B	M	4.80	4.00	1.00	8.84
		SD	.45	.71	.00	2.60
	Block C	M	2.50	3.40	2.40	8.10
		SD	.71	.55	.89	1.85
3	**Rest Days**	M	4.91	2.45	1.91	8.99
		SD	.83	.52	1.04	1.88
	Block A	M	4.20	2.20	1.20	7.92
		SD	.45	.45	.45	1.65
	Block B	M	5.20	2.20	1.60	8.52
		SD	.84	.45	.55	2.21
	Block C	M	4.75	1.67	2.00	10.01
		SD	.50	.58	.00	.87
4	**Rest Days**	M	5.20	3.00	2.60	7.73
		SD	.63	.94	.97	1.46
	Block A	M	4.80	2.60	2.40	7.42
		SD	.84	.55	.55	1.28
	Block B	M	5.20	3.00	2.40	7.75
		SD	.84	.00	.89	1.06
	Block C	M	5.20	2.60	2.20	7.60
		SD	.45	.55	.45	.70
5	**Rest Days**	M	5.45	5.42	3.08	8.37
		SD	.69	.90	1.00	1.53
	Block A	M	5.40	4.80	3.40	7.38
		SD	.89	.84	1.52	1.20
	Block B	M	5.80	6.00	2.20	7.84
		SD	.45	.00	.84	1.65
	Block C	M	6.00	5.20	2.20	7.51
		SD	.00	.84	1.10	1.57
6	**Rest Days**	M	4.45	2.90	1.90	8.49
		SD	.52	.74	.57	1.39
	Block A	M	5.40	2.80	2.20	7.99
		SD	.55	.45	.45	1.50
	Block B	M	5.60	2.60	2.00	7.82
		SD	.55	.55	.71	1.38
	Block C	M	6.20	2.80	2.40	8.15
		SD	.45	.84	1.14	.70
7	**Rest Days**	M	4.00	2.64	2.55	7.76
		SD	.63	.50	.93	1.27
	Block A	M	4.40	2.40	2.00	8.00
		SD	1.34	.89	.00	.94
	Block B	M	4.00	2.40	1.80	8.47

		SD	1.22	.55	.45	1.24
	Block C	M	3.80	2.40	2.40	7.97
		SD	.84	.55	.55	.62
8	**Rest Days**	M	4.55	3.30	2.50	8.70
		SD	.69	.67	1.08	1.37
	Block A	M	5.80	4.20	2.80	8.90
		SD	1.10	.84	.45	1.58
	Block B	M	6.00	4.40	2.80	8.20
		SD	.71	.89	.84	2.35
	Block C	M	5.40	4.00	3.40	8.76
		SD	.55	.71	.55	1.03
9	**Rest Days**	M	4.50	2.17	2.00	8.10
		SD	.80	.39	.95	.99
	Block A	M	4.80	2.20	2.40	8.91
		SD	.45	.45	1.52	1.36
	Block B	M	4.60	2.00	2.60	8.59
		SD	.55	.00	.55	.72
	Block C	M	4.20	2.20	2.20	8.49
		SD	.84	.45	.45	.54
10	**Rest Days**	M	4.91	1.82	2.18	8.33
		SD	.54	.40	.60	1.68
	Block A	M	5.60	2.60	2.60	7.88
		SD	.55	.55	1.14	1.78
	Block B	M	5.80	3.20	2.80	7.48
		SD	.45	.45	.45	1.66
	Block C	M	6.00	3.00	2.60	7.75
		SD	.00	.71	.55	1.78
11	**Rest Days**	M	5.09	3.58	2.67	8.21
		SD	0.54	0.67	1.07	1.42
	Block A	M	4.80	4.20	2.40	8.00
		SD	1.30	1.30	0.89	1.30
	Block B	M	5.60	4.20	2.40	7.46
		SD	1.14	0.45	0.89	1.25
	Block C	M	5.60	4.00	2.60	8.59
		SD	0.89	0.00	0.89	1.47
12	**Rest Days**	M	4.45	2.73	2.36	8.57
		SD	0.93	0.65	0.67	2.42
	Block A	M	5.20	3.00	3.20	8.06
		SD	0.45	0.00	1.10	0.89
	Block B	M	5.60	2.60	3.00	8.38
		SD	0.55	0.89	0.00	0.72
	Block C	M	4.80	2.00	2.60	9.31
		SD	0.45	0.00	0.55	0.65
13	**Rest Days**	M	5.36	3.20	3.33	7.39
		SD	0.50	0.63	0.71	1.24
	Block A	M	5.80	3.60	3.40	7.92
		SD	0.45	0.55	0.55	0.55
	Block B	M	5.60	3.80	3.60	7.65
		SD	0.89	0.45	0.55	0.98
	Block C	M	5.60	3.40	3.25	7.13
		SD	0.89	0.55	0.50	1.24
14	**Rest Days**	M	4.80	3.60	2.60	7.68
		SD	0.42	0.52	0.84	1.20
	Block A	M	5.00	4.00	3.20	7.40
		SD	0.00	0.00	0.45	1.27
	Block B	M	4.40	3.40	2.60	7.26
		SD	0.55	0.55	0.55	1.00
	Block C	M	5.80	3.40	2.80	8.17
		SD	0.84	0.55	0.45	1.06

15	**Rest Days**	M	5.36	4.27	2.82	9.14
		SD	.50	.47	1.08	2.31
	Block A	M	5.40	4.40	2.80	6.96
		SD	.89	.89	.84	1.49
	Block B	M	5.00	4.60	3.00	6.78
		SD	.00	.89	1.22	1.51
	Block C	M	5.60	4.40	2.60	8.74
		SD	.55	.55	.89	.85
16	**Rest Days**	M	4.55	3.36	2.60	8.25
		SD	.52	1.12	1.26	.82
	Block A	M	5.00	3.80	2.50	7.65
		SD	.71	.84	.58	1.14
	Block B	M	5.80	3.80	3.20	8.60
		SD	.84	.45	.45	1.47
	Block C	M	5.00	3.40	3.00	7.10
		SD	.00	.55	.00	.59
17	**Rest Days**	M	4.50	3.00	3.13	8.41
		SD	.93	.76	1.13	1.88
	Block A	M	6.00	4.00	3.00	8.15
		SD	.71	1.00	1.00	3.25
	Block B	M	5.40	4.00	2.80	8.00
		SD	1.14	.71	1.30	2.00
	Block C	M	4.00	3.25	3.50	8.68
		SD	1.15	1.50	1.73	2.84
18	**Rest Days**	M	4.91	3.82	2.09	7.06
		SD	1.45	1.17	.83	2.13
	Block A	M	5.80	3.20	2.00	7.82
		SD	1.10	.84	1.15	2.30
	Block B	M	4.60	5.00	1.80	7.12
		SD	1.14	.71	.84	1.61
	Block C	M	5.20	3.80	2.20	7.26
		SD	.84	1.30	.45	2.57
19	**Rest Days**	M	5.55	2.18	2.82	7.69
		SD	.69	.98	1.08	1.20
	Block A	M	5.40	2.20	1.40	7.52
		SD	.55	.84	.55	1.07
	Block B	M	5.60	2.60	2.20	7.45
		SD	.89	1.14	.45	1.20
	Block C	M	6.00	2.80	2.00	7.32
		SD	.00	1.64	.00	1.09
20	**Rest Days**	M	5.42	3.75	3.10	6.13
		SD	1.08	.97	.88	2.14
	Block A	M	6.60	4.20	3.00	6.97
		SD	.55	1.30	.00	1.48
	Block B	M	6.20	3.40	3.20	6.85
		SD	1.10	1.52	1.10	1.50
	Block C	M	5.80	3.20	3.40	7.16
		SD	.45	.45	.55	1.18
21	**Rest Days**	M	5.08	2.75	2.75	8.61
		SD	.90	1.42	1.06	1.21
	Block A	M	5.40	2.60	2.40	8.58
		SD	.55	1.14	.55	.64
	Block B	M	5.60	3.20	2.80	8.01
		SD	.55	1.30	.84	.80
	Block C	M	5.60	1.80	2.80	9.17
		SD	.55	.84	.84	.61
22	**Rest Days**	M	4.82	3.91	2.82	8.88
		SD	.40	.54	.75	1.38
	Block A	M	4.60	3.80	2.80	8.92

		SD	.55	.45	.84	1.69
	Block B	M	5.00	3.80	2.60	8.58
		SD	.71	.84	.55	1.51
	Block C	M	5.20	4.00	3.00	8.84
		SD	.45	.71	1.41	1.69

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)* 05-10-2012	**2. REPORT TYPE** Technical Memorandum	**3. DATES COVERED** *(From - To)* 12/2009 - 8/2012

4. TITLE AND SUBTITLE

Second Interim Report
NASA – easyJet Collaboration on the Human Factors Monitoring Program (HFMP) Study

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Ashok N. Srivistava and Phil Barton

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER
534723.02.01.01.10

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

NASA Ames Research Center
Moffet Field, California 94035-1000

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

National Aeronautics and Space Administration
Washington, DC 20546-0001

10. SPONSORING/MONITOR'S ACRONYM(S)
NASA

11. SPONSORING/MONITORING REPORT NUMBER
NASA/TM-2012-216053

12. DISTRIBUTION/AVAILABILITY STATEMENT

Unclassified-Unlimited
Subject Category 03
Availability: NASA CASI (443) 757-5802

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This is the second interim report jointly prepared by NASA and easyJet on the work performed under the agreement to collaborate on a study of the factors entailed in flight and cabin-crew fatigue, and decreases in performance associated with fatigue. The objective of this Agreement is to generate reliable procedures that aid in understanding the levels and characteristics of flight and cabin-crew fatigue factors, both latent and proximate, whose confluence will likely result in unacceptable crew performance. This study entails the analyses of numerical and textual data collected during operational flights. NASA and easyJet are both interested in assessing and testing NASA's automated capabilities for extracting operationally significant information from very large, diverse (textual and numerical) databases; much larger than can be handled practically by human experts.

15. SUBJECT TERMS

data mining, safety management systems, human factor monitoring

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19b. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			STI Help Desk (email: help@sti.nasa.gov)
U	U	U	UU	103	**19b. TELEPHONE NUMBER** *(Include area code)* (443) 757-5802