REQUIREMENTS ANALYSIS AND ARCHITECTURE FOR AN OPERATIONAL STUDY OF FATIGUE IN USAF MOBILITY AIRCREW

THESIS

Jonathan F. Mecham, Major, USAF

AFIT-ENV-MS-21-S-076

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT-ENV-MS-21-S-076

REQUIREMENTS ANALYSIS AND ARCHITECTURE FOR AN OPERATIONAL STUDY OF FATIGUE IN USAF MOBILITY AIRCREW

THESIS

Presented to the Faculty

Department of Systems Engineering and Management

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Master of Science in Systems Engineering

Jonathan F. Mecham, BS

Major, USAF

August 2021

DISTRIBUTION STATEMENT A.
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

AFIT-ENV-MS-21-S-076

REQUIREMENTS ANALYSIS AND ARCHITECTURE FOR AN OPERATIONAL STUDY OF FATIGUE IN USAF MOBILITY AIRCREW

Jonathan F. Mecham, BS

Major, USAF

Committee Membership:

Michael E. Miller, Ph.D.
Chair

Lt Col Jeremy R. Geiger, Ph.D.
Member

Maj Joseph P. Kristbaum, Ph.D.
Member

AFIT-ENV-MS-21-S-076

Abstract

Aircrew fatigue in flight operations is a known hazard that has driven the creation of fatigue-reducing regulation and fatigue risk management systems industry wide. In addition, biomathematical models have been created and tested to forecast the effectiveness of aircrew under conditions of time-zone shifts and long duty days. However, limited operational studies exist to validate these models or to help understand how individual factors can affect them. Operational studies have a variety of limitations that make gathering typical data regarding fatigue difficult. This research takes a systems requirement analysis approach to design a study that measures effects of circadian disruption on USAF C-17 Aircrew effectiveness. This study could then aid in understanding the effectiveness of fatigue-related regulation and fatigue risk management systems used in Air Force Mobility operations. The current research develops requirements for such a study through analysis of existing research as well as through a small-scale study to identify limiting factors for conducting such a study in an operational mobility squadron. The research further suggests additional research to explore the inclusion of fatigue monitoring into the Air Force Safety System, particularly for Air Force Mobility Operations.

Acknowledgments

I would like to express appreciation to my faculty advisor, Dr. Miller, for his continual encouragement to finish the course despite the many obstacles and circumstances time have presented. I would also like to express my appreciation for the members of the 6th Airlift Squadron, the Bully Beef, for the inspiration, support, and encouragement with this project, as well as the 469th Flying Training Squadron, the Fighting Bulls, for the providing the support and time to finish this project.

Jonathan F. Mecham

Table of Contents

Page

List of Figures

Page

List of Tables

REQUIREMENTS ANALYSIS AND ARCHITECTURE FOR AN OPERATIONAL STUDY OF FATIGUE IN USAF MOBILITY AIRCREW

I. Introduction

> *"My mind clicks on and off… I try letting one eyelid close at a time while I prop the other with my will. But the effect is too much, sleep is winning, my whole body argues dully that nothing, nothing life can attain is quite so desirable as sleep. My mind is losing resolution and control." – Charles Lindbergh, The Spirit of St Louis*

General Issue

The effects of fatigue on flyers have been studied since the early days of aviation. Physicians who specialized in aviation defined fatigue and discussed its effects on aircrew. This area of study is especially critical in long-haul air mobility operations, where multiple time zones are crossed with long duty days that cause significant shift of circadian rhythms. Decades of controlled studies and aviation experience have greatly increased our understanding of fatigue in flight operations. In response to this, the FAA released new regulations for Part 117 and 121 operations that add additional duty day limitations based on circadian disruption (Lowry, 2012). Despite these updated regulations, however, the National Transportation Safety Board still placed "Reduce Fatigue Related Accidents" in its 2019-2020 Most Wanted List of Transportation Safety Improvements (*2019–2020 NTSB Most Wanted List of Transportation Safety Improvements: Reduce Fatigue - Related Accidents*, 2019). Air Force regulatory guidance in mitigating fatigue, even in long-haul operations, has remained static during this same period.

Within the Aviation Safety Reporting System (ASRS), run by NASA, aircrew can submit voluntary reports providing confidential safety information. During the calendar years of 2017 to 2020, aircrew flying under Part 121 Federal Aviation Regulations - regularly scheduled air carriers - submitted over 10,770 reports to ASRS. Out of those reports, 3% were categorized as having fatigue as a factor (*ASRS Query*, 2021).

While no real conclusions can be drawn from voluntary reporting systems such as this given the differences in reporting incentives, query options, and the wide variety of operations being performed within both communities, the high incidence of self-reports related to fatigue illustrate that aircrew themselves are aware of fatigue being a risk to operations. The incidence of this issue may be higher among Air Force pilots where there is less regulatory protection against fatigue risk.

Prior to departing on most missions, mobility aircrew are provided with a chart produced by the AvORM application. This software uses the Sleep, Activity, Fatigue and Task Effectiveness (SAFTE) applied model to chart the expected cognitive effectiveness of an aircrew throughout a mission and can include variables such as time-zone shift in its modeling (Chaiken, 2005).

Significant research has been done to understand physiological effects of circadian rhythm disruption. Predictive modeling algorithms have been developed to help forecast when an individual may require sleep or be operating at less than desired levels of performance. This could have substantial impact in risk mitigation in aviation, where 24/7 worldwide long-haul and ultra-long-haul operations have become so integrated within commercial and military logistics. However, real world data is required to determine the validity and reliability of these tools (Dinges, 2004). Operational environments bring uncontrolled factors that are not seen in the controlled environments on which these models were based, such as sleep disruptors in the environment, use of caffeine, or naps in crew facilities (Martinez & Quintero, 2015).

Problem Statement

It is unclear if regulation concerned with mitigating aircrew fatigue is effective or sufficient. Naturalistic observation could provide validation or critique of existing regulations as they apply to C-17 aircrew in long-haul operations that cover multiple time zones. The wide variety of C-17 missions flown across time zones as well as fewer regulations in regard to flight duty period and crew rest provide an opportunity to observe a wide spectrum of different scenarios and variation in crew rest, flight duty period, circadian shift, and others.

However, additional restrictions and requirements come with gathering data on operational missions. Air Force aircrew can face challenging or stressful environments where collecting typical data is at best not feasible, and at worse increases risk. Sporadic schedules make collection of data difficult. Combat or other high stress environments can make tools, such as questionnaires or tests of cognitive degradation, which have been used successfully in more controlled environments, impractical. Therefore, robust methods for gathering realistic, real-world data to support improvements in models of pilot fatigue have yet to be successfully demonstrated despite earlier attempts.

Research Objectives

The primary question to explore is whether current Air Force regulation and fatigue risk management systems (FRMS) are sufficient in reducing the risk of aircrew fatigue in mobility long-haul operations. The scope of this question is massive. Beyond the complexity of attempting to measure fatigue and rest in a controlled environment, it introduces additional complexities brought on in an operational environment. With the large number of variables possible to be observed, the difficulty of observing these variables in the operational environment is daunting. This fact is especially true when operations are not to be impeded. Therefore, the goal is not to address the core research objective, but to attempt to find the practices and measures that provide the most information towards providing a possible answer without imposing additional risks on successful operations.

In summary, the objective of this study is to identify a possible system architecture for naturalistic observation of the effects of circadian rhythm disruption on C-17 aircrew effectiveness, comprised of best practices or methods. These practices and methods could then be used in further study to better answer the question of whether current regulations or fatigue management systems effectively address fatigue encountered in air mobility operations.

Methodology

This paper documents the initial steps of creating a system architecture to address the problem, with the goal of creating a system, or study, that can adequately measure the effects of circadian rhythm disruption on mobility aircrew effectiveness while not in and of itself hampering that effectiveness. These initial steps include identifying needs and resources for such a study, translating those needs into requirements and goals, and defining a high-level system architecture with associated interfaces. The initial step of identifying needs and resources involved conducting a small sleep study with air mobility crew members that used actigraphy and light monitoring wearables and questionnaires. The lessons learned from this study provided limitations and requirements on what could be studied at a larger scale. The data is also analyzed to gain insight on possible areas of focus for future study. Finally, requirements are diagrammed, a high level-architecture is proposed, including a class diagram representing data flows, and analysis is performed to support the proposed study.

Assumptions/Limitations

The most significant limitations are based around the subjects: operational aircrew. This removes some options in performing an observational study that does not interfere with normal duties. Included is that polysomnography (PSG) and other more intensive methods of measuring sleep quality, fatigue levels, etc., will not be used as they impose too much workload to be utilized by an active aircrew as they often require multiple electrodes and a professional monitoring their use. These limitations necessitate alternative means to ascertain fatigue mitigation measure effectiveness.

In addition, a general assumption is made that a reduction of sleep and/or sleep efficiency contributes to aircrew fatigue. If a baseline of sleep is obtained and this baseline of sleep is significantly reduced under certain conditions, it is assumed that greater fatigue is experienced. Direct measures of fatigue, such as a Psychomotor Vigilance Tasks, are subject to the same limitations as mentioned above.

Implications or Expected Contributions

This work is meant to establish a foundation for future work in operational naturalistic study of U.S. Air Force mobility aircrew to better understand factors contributing to fatigue.

II. Literature Review

"Indeed, death is the only final lasting escape; nature realizes this, and that is why nature gave us sleep. Sleep is our only surcease from the endless attacks upon our organism; when we sleep we approach the threshold of death; we shut out for a little time the endless assaults of the enemy hordes… the healing power of oblivion." – M.C. Grow, Military Surgeon, 1936 (Porter, 1936)

Chapter Overview

Fatigue as used in this study is defined. Current regulations related to mitigating fatigue in the FAA and Air Force are compared. Common biomathematical models of fatigue are then discussed. Finally, studies that involve operational aircrew are reviewed.

Defining Fatigue

Early human factors aviation research delineated the physical traits which made pilots successful for flight duty. Aircrew were viewed as possessing positive or negative physical traits, similar to design differences which gave aircraft advantages or disadvantages for missions. From this viewpoint, fatigue was often portrayed as an aggravating factor to pre-existing physical issues, for example tiring of physical faculties such as ocular muscles which resulted in pilot deficiencies, such as poor depth perception, despite the pilot overall feeling fine. Discussing aviation casualties from World War I, Dr. Harold Cooper wrote “while many of the fatalities were due largely to poor equipment, physical defects were responsible for a large number”. Physical defects, accentuated by fatigue and minor illnesses, were identified as the root causes of casualties in many cases (Cooper, 1930). The purpose of monthly medical check-ups on flyers was to look for physical deficiencies, psychological trends, and “the fatigues and physical staleness due to either or both of the above and to excessive flying” (Miller, 1930)

Later discussions of fatigue evolved and developed classifications, such as neuro-muscular fatigue, nervous or industrial fatigue, or initiative fatigue. These classifications assumed that fatigue resulted from pilot's continual input into a system and the environment, as well as from the pilot's requirement to acquire input from the system and the environment.

Generally, fatigue can be defined as "a decrease in performance or performance capability as a function of time on task" (Salas, 2010). Within the world of aviation, this general definition is foundational to the current Federal Aviation Regulation definition of fatigue as "a physiological state of reduced mental or physical performance capability resulting from lack of sleep or increased physical activity that can reduce a flight crew member's alertness and ability to safely operate an aircraft or perform safety-related duties" (FAR, 2021). This definition provided by the FAA will be applied in this paper.

Current Regulation Regarding Fatigue Management

Air Force regulations regarding Fatigue Management are found in the Air Force Manual (AFMAN) 11-202 Volume 3, Flight Operations, Chapter 3, General Flight Rules. A summary of these rules follows. While the definition of performance provided earlier defines fatigue in terms of alertness and ability to safely operate, it acknowledges that behavior such as lack of sleep or activity are the contributing factors. While reviewing these regulations, it is important to note that current regulations focus on the crew's behavior to avoid human performance degradations which are the measurable effects of fatigue.

According to the regulation, aircrew must be provided a 12-hour rest opportunity prior to beginning a flight duty period. This 12 hour rest opportunity is intended to allow 8 hours of uninterrupted sleep. The flight duty period (FDP) begins when aircrew first report for official duty and ends at final engine shutdown after the final flight of the completed mission. The maximum period of this FDP for a transport aircraft with sleeping provisions, such as the C-17, is 16 hours for a basic crew, or 24 hours for an augmented crew.

In addition, the AFMAN 11-202v3 sets maximum flight times as 56 hours per 7 consecutive days, 125 hours per 30 consecutive days, and 330 hours per 90 consecutive days. These can be waived by MAJCOM/A3, and that waiver authority can be delegated as low as the squadron commander. Other rules included are placing deadhead time within the FDP of an aircrew member, as well as conditions and waiver authorities for when the FDP can be increased, or crew rest decreased (202v3, 2020).

These general flight rules are supplement by Air Mobility Command. This command adds two key elements, Home-Station Pre-Departure Crew Rest (3.1.7) and Post-Mission Crew Rest (PMCR) (3.1.10). For pre-departure crew rest, unit commanders are to enter aircrew members into crew rest 24 hours before the legal for alert time, although it permits the crew to dedicate the first 12 hours of that period for limited non-flying duties like mission planning (202v3 AMC Sup, 2020). This allows aircrew a larger window of time to obtain sufficient rest before a duty day begins. PMCR is 1 hour for each 3 hours off station, up to 96 hours, beginning once all official duties after a mission have concluded. The purpose is to “give aircrew members returning to home base sufficient time to recover from cumulative effects of the mission and tend to personal needs (202v3 AMC Sup, 2020). During the mission itself, the supplement regulation states that mobility planners "should" construct itineraries with longer than the minimum required ground time specifically to provide "opportunities to recover from the cumulative effects of fatigue caused by flying on several consecutive days or due to transiting several time zones." It then states, "if practical, make the enroute ground time 36 hours (maximum) after three consecutive near maximum FDPs" (202v3 AMC Sup, 2020). While this is not required, it does demonstrate an awareness of the possible effects of fatigue on aircrew.

Finally, these general rules and MAJCOM rules are applied specifically to C-17 operations by the AFMAN 11-2C-17 Volume 3, Flight Operations, C-17 Operations Procedures. Chapter 2, Aircrew Complement/Management, does not modify the general rules above with the exception of adding some restrictions to FDP in cases such as the autopilot being inoperative, or limiting the window of the FDP in which events such as air-to-air refueling can be conducted. (Air Force Manual 11-2C-17 v3, Flying Operations, C-17 Operations Procedures, 2019)

There are many similarities between these General Flight Rules used by C-17 aircrew and the Federal Aviation Regulations that govern civilian aviation. In 2012, after decades of discussion and proposed rules around the role of regulation in fatigue management, the FAA issued a final rule that expanded fatigue management regulations beyond hour limit caps and static crew rest requirements. It incorporated requirements "based on the time of day, whether an individual is acclimated to a new time zone, and the likelihood of being able to sleep under difficult circumstances" (CFR Final Rule - Flightcrew Member Duty and Rest Requirements, 2012).

The source of Federal Aviation Regulation concerning Fatigue Management Regulation is Part 117 – Flight and Duty Limitations and Rest Requirements: Flight crew Members. In addition to a series of general hour caps based upon cumulative limitations (i.e., 100 flight hours in any 672 consecutive hours) it adds some key definitions to properly interpret maximum flight duty period times and minimum rest periods. These definitions include:

Acclimated: a condition in which a flight crew member has been in a theater for 72 hours or has been given at least 36 consecutive hours free from duty.

Physiological night's rest: 10 hours of rest that encompasses the hours of 0100 and 0700 at the flight crew member's home base unless the individual has acclimated to a different theater. If the flight crew member has acclimated to a different theater, the rest must encompass the hours of 0100 and 0700 at the acclimated location. Window of circadian low: a period of maximum sleepiness that occurs between 0200 and 0559 during a physiological night. Tables 2 and 3 incorporate these definitions based on a flight duty period’s scheduled start time and the crew composition.

Table 1 Maximum Flight Time Limits for Unaugmented Operations FAR Part 117

Time of report (acclimated)	**Maximum flight time (hours)**
0000-0459	8
0500-1959	9
2000-2359	8

Table 2 Flight Duty Period for Unaugmented Operations Table from FAR Part 117

Scheduled time of start (acclimated time)	Maximum flight duty period (hours) for line holders based on number of flight segments						
	1	**2**	**3**	**4**	**5**	**6**	**7 +**
0000-0359	9	9	9	9	9	9	9
0400-0459	10	10	10	10	9	9	9
0500-0559	12	12	12	12	11.5	11	10.5
0600-0659	13	13	12	12	11.5	11	10.5
0700-1159	14	14	13	13	12.5	12	11.5
1200-1259	13	13	13	13	12.5	12	11.5
1300-1659	12	12	12	12	11.5	11	10.5
1700-2159	12	12	11	11	10	9	9
2200-2259	11	11	10	10	9	9	9
2300-2359	10	10	10	9	9	9	9

Finally, part 117 adds a rest time upon return from a trip, a minimum of 56 consecutive hours if he/she travels more than 60° longitude during a flight duty period or a series of flight duty periods and is away from home base for more than 168 consecutive hours. This rest time must encompass three physiological nights' rest based on the local home station time.

Many other stipulations are provided, and there are some variations provided based on the type of operations being conducted under different sections of the FAR such as 119 Air Carries and Commercial Operations as well as 121, Domestic and Flag carrier operations. For the purposes of this paper, it is sufficient to note that the general FAA regulations have been adopted across operation type regulations with the focus on the disruption of circadian rhythm that can result from the nature of long and ultra-long-haul operations.

Table 3 Flight Duty Period for Augmented Operations Table from FAA Part 117

Scheduled time of start (acclimated time)	Maximum flight duty period (hours) based on rest facility and number of pilots					
	Class 1 rest facility		Class 2 rest facility		Class 3 rest facility	
	3 pilots	4 pilots	3 pilots	4 pilots	3 pilots	4 pilots
0000-0559	15	17	14	15.5	13	13.5
0600-0659	16	18.5	15	16.5	14	14.5
0700-1259	17	19	16.5	18	15	15.5
1300-1659	16	18.5	15	16.5	14	14.5
1700-2359	15	17	14	15.5	13	13.5

With the relatively static limitations of U.S. Air Force regulations, compared to more dynamic Federal regulations, there is relatively little known about the effect of these regulations in mitigating mishap risk due to fatigue. This lack of data has, at times, been used by commercial carriers to push back against the possibility of further rules and regulations (Weir, 2002) and these regulations have not been adopted by the U.S. Air Force.

Biomathematical Models of Fatigue

Significant research has been performed to construct models capable of predicting fatigue in certain conditions. A significant portion of the Fatigue Risk Management

System used by Air Force Mobility aircrew is based off the SAFTE – FAST model and application. This model was developed specifically with DoD application in mind and over the years has incorporated factors such as time zone shifts away from home base into its algorithms. This capability makes it a useful tool for aircrew and schedulers to apply when forecasting air crew risk and effectiveness as a result of fatigue. As a result, these estimates of risk and effectiveness can be included within the criteria for selecting air crew schedules.

Research has been performed to evaluate model effectiveness relative to real world data sets. For example, a 2004 report from the "Fatigue and Performance Modeling Workshop" summarized the results from six modeling teams' algorithms to predict fatigue levels across five different scenarios. Scenario 1 was 88 hours of extended wakefulness with and without naps. Of note for this scenario, "…none of the models predicted the continuing build-up of subjective sleepiness and, in particular, performance impairment across the 14 d of sleep restriction". Scenario 2 was 14 days of partial sleep deprivation for subjects who either got 4 hours of sleep per day, or 6 hours per day. Scenario 2 results were favorable for the models if the probable results of sleep inertia were excluded, as the models did not include this factor. Scenario 3 involved freight locomotive engineers who are on-call and notified of possible work 2 hours prior to report time. Train driving was noted as being "non-vigorous, highly cognitive activity that generates considerable mental workload from continuous mental calculations, spatial memory use, and vigilance monitoring" with noise and variations in light. No significant analysis was performed on this scenario as it was difficult to evaluate the differing schedules of each engineer. Scenario 4 was ultra-long-range flight operations with four crewmembers. Of note in this evaluation was the lack of actual data for the scenario, the results produced by the models were simply compared to illustrate differences between the models. Finally, scenario 5 was another laboratory study with 7 days of restricted sleep, followed by a 3-day recovery period. This scenario gave the models similar difficulty to scenario 2, where they could not predict the time recovered for recovery on waking. (Van Dongen, 2004). The limitations of this study include the presence of factors, such as sleep inertia, that science does not fully understand or know how to model. Both the scenarios where this was an issue could have similarities drawn to missions flown by aircrew where sleep effectiveness is mitigated by a variety of factors and the mission extends to more than a week. In addition, there is limited data available for operational conditions such as long-haul flights to which the models can be compared.

The model used by Air Force mobility aircrew, the SAFTE model, also has a number of limitations. For example, in the SAFTE model structure, one can reach an equilibrium state after so many days of an individual receiving less than optimal sleep each night. At a certain point, sleep accumulation becomes large enough with increased sleep debt, such that the individual maintains a stable level of cognitive performance while continuing to get less than desired sleep (Hursh et al., 2004).

The primary disadvantage to any model is best summarized in the conclusion of a SAFTE study regarding fatigue models for applied research in warfighting which states that: “It may not be possible or desirable to adopt a universally accepted standard for performance measurement, but in the absence of a standard, great care must be taken when applying a model to a performance metric distinct from the one used to design the model. Ultimately, all models will be judged by their ability to make useful predictions of the performance of greatest interest to the user, which is most likely not going to be performance on a standard cognitive test, but rather performance of some job. The greatest challenge facing fatigue modeling is how to bridge this gap between laboratory metrics of performance and performance in the natural environment of work and war" (Hursh et al., 2004).

Fatigue Research in Operational Environments

Operational studies to understand the effects of fatigue have been conducted in many fields. This section summarizes the results of those applicable to the study of mobility Air Force aircrew.

In 2011, the air carrier Finnair conducted a study of 34 pilots over a period of 400 days. The study compared its results of assessing aircrew mental tiredness with that predicted by the Boeing Alertness Model. The correlation was sufficient for the company to incorporate the model into crew scheduling via the Jeppesen fatigue risk management system (Kirby, 2011).

One study intended to understand space crew members' effectiveness on long duration space flights was conducted in the Human Exploration Research Analog. Four crews under different scenarios completed a Psychomotor Vigilance Task (PVT) five times a day for 3 days. The results were then compared with predictions of three sleep-wake models: State-space Model, Unified Model of Performance, and SAFTE-FAST Model. They found significant association of the predictions created by the State space and the SAFTE-FAST models with measured PVT performance (Shin et al., 2018)

NASA Ames conducted a study to understand the effect of short-haul airline operations on operator fatigue. In this study, the schedule was controlled, while the actual amount of sleep taken and use of countermeasures could not be controlled. The study consisted of the aircrew taking a 5-minute PVT test when they woke up, at cruise prior to descent, after the flight, and before they went to bed. Additionally, the crew wore a device capable of performing actigraphy and capturing light levels, as well as responded to questionnaires in sleep diaries. The study found that models generally represented the challenges that aircrew would confront but determined that they did not account for individualized factors, such as countermeasure use and tolerance to those countermeasures, personal circadian clock, age, and other health conditions (Gregory et al., 2017).

III. Methodology

Chapter Overview

The first significant portion of this research was to look at a variety of options for study to develop possible options for a fatigue study that could work with aircrew. The second portion of the research summarized below was the execution of a small-scale study in an operational squadron to help further identify limiting factors and requirements in an operational environment.

Overview of Research Methodology

Significant research was done to identify not just what needs to be measured for an effective sleep study, but what options were available to measure these variables. More “traditional” sleep study tools such as PSG would be too invasive during actual operational missions, and even non-invasive tools such as PVT or questionnaires can present a time-burden to aircrew that leads to non-participation or incomplete data. Thus, multiple possibilities for measures and tools were researched. Companies with commercial wearable devices that claim to measure fatigue or sleep quality were contacted and provided certain specifications or substantiating research on their products.

Operational Study Description

A small operational study was conducted in a C-17 squadron. This study was performed by providing aircrew with Phillips Respironics Actiware watches for measuring actigraphy as well as light level. The watches were able to measure the amount of time in different states based on movement, the states being: 1 – awake; 2 – asleep; 3 – off wrist. The intention was to provide these watches 2 days prior to the mission to establish a baseline of sleep quality. Aircrew were also asked to fill out a journal daily that included approximate times of any naps taken and estimates of caffeine intake, alcohol intake after duty hours, and if “no-go” pills were used prior to sleep. For mobility aircrew these include Ambien, Restoril, or Sonata.

Upon return from a mission the watches were turned in to a central location, where they were downloaded by a gatekeeper who would adjust date and time stamps to remove any possible association of crew to an actual mission history and provide an associated masked itinerary of locations designated only by their time zone shift from home station.

Subjects were volunteer aircrew between the ages of 23 and 35, and included not just pilots, but loadmasters as well, and in one case a crew chief that was attached to the crew. No medical histories were obtained, as Air Force aircrew have been previously screened for issues that would affect sleep quality such as sleep apnea.

The only criteria used to select a mission for study was that the planned mission was to go at least three time zones from home station. Otherwise, circadian rhythm disruption would be minimal unless it was to depart during normal sleeping hours. However, other factors were added by operations officers, such as evaluation missions where aircrew were under additional stress and any additional tasks to perform would be seen as a possible hinderance to their performance. Unfortunately, many of the most difficult missions in terms of schedule, airfields transited, and time away from home were thereby excluded from possible data collection.

Data was obtained from only two missions. For a third, questionnaires were completed but data was corrupted for an unknown reason from the watches. Within the two missions, four subjects wore the device for a sufficient duration to record usable data in addition to completing questionnaires.

IV. Analysis and Results

Chapter Overview

This chapter begins with presenting results of the small-scale operational study conducted and an associated data analysis. It then moves to analysis of possible measurements and measurement devices. It concludes with the creation of a requirements diagram and domain model class diagram for a study system that could answer the question of the effectiveness of current fatigue regulation and models.

Results of Small-Scale Study

As the study was only able to generate limited data, the most valuable lessons learned were the barriers to data collection. Barriers included physical downloading of the data, lack of motivation to wear the device, and perceived added workload as will be discussed in detail within this section. Further, we can explore the limited available data to understand whether certain assumptions within the current models may be sufficient.

Regarding the physical downloading of data, the actigraphy devices required the researcher to have the physical watch, plug it into a particular computer that had certain software installed with an associated key, and download. The device then had to be cleared and reset. To collect the watches, aircrew were directed to leave them at a central location after returning from their mission. This was often forgot, as the end of a mission includes unloading and loading personal equipment, completing post-mission paperwork, and debriefing at the end of what was often a 24-hour duty day at a time when no other personnel, including any researcher who could collect the watches, was present at the squadron. Those who could download the data were aircrew as well and were likely to be on missions themselves once watches were available for download. The combination of these factors and especially the data being collected by other operators in the squadron led to multi-month delays in turning a watch from one mission to the next.

With regards aircrew to motivation to wear the device, the device was often viewed as a nuisance by aircrew. While PSG is the "gold-standard" of sleep scoring, this device is considered top of the line for actigraphy, which is considered a less reliable but alternative method for sleep scoring. In many studies looking at consumer grade wearables, the benchmark used for accuracy is the Actiwatch Spectrum Pro. It is a clinical device and has a clinical look. In the words of one aircrew, "wearing it made me feel like an outpatient, or a parolee with a tracker." It's the size of a normal watch but lacks the functionality or style of one. This meant that aircrew would wear their own watch, typically specifically selected and/or setup to help in their specific duties and wear the Actiwatch on an opposing wrist. After duty hours and before sleeping it was common for aircrew to remove the watch for things such as going to dinner, as it looked significantly different from something they would wear or use. Many times, this happened, and no data was logged. For three subjects, this happened relatively early in the mission and, as they had failed to follow the experimental instructions, they perceived

Figure 1 Actiwatch Spectrum Pro - usa.philips.com

that their data may no longer be valuable and therefore, there was little use from wearing the watch the remainder of the mission.

Additionally, a review of the data shows that device collection did not typically begin until just prior to the crew rest leading up to the first “alert” or start of the mission. This limited the ability to establish a baseline against which to compare. The intension was for subjects to wear the watch at the start of pre-mission crew rest (PMCR), but whether due to last minute crew changes, or the motivational concerns discussed earlier, crews would typically not begin wearing the watch until the one rest period prior to their mission alert. This period of sleep is often at an irregular time and a “no-go” sleep aid was typically taken as well, making it unsuitable for establishing a benchmark of sleep quantity or efficiency.

Finally, it was surprising that the most complete part of data collection came from the questionnaires submitted; however, aircrew often noted that they had forgotten to fill out the questionnaire near the period of sleep being described and were guessing at some questions, like which sleep periods they had used no-go sleep aids, by the time they remembered to complete the questionnaire.

Multiple lessons were learned from these trials with the following requirements derived for an operational study of aircrew: the sleep study system should be capable of transmitting data remotely; the sleep study system should be wearable and usable, able to integrate into the aircrew’s patterns, or possibly already integrated with some wearable solutions discussed below; the sleep study system should be used long enough to establish a baseline of sleep quality; the sleep study system should remind aircrew to complete necessary survey data periodically.

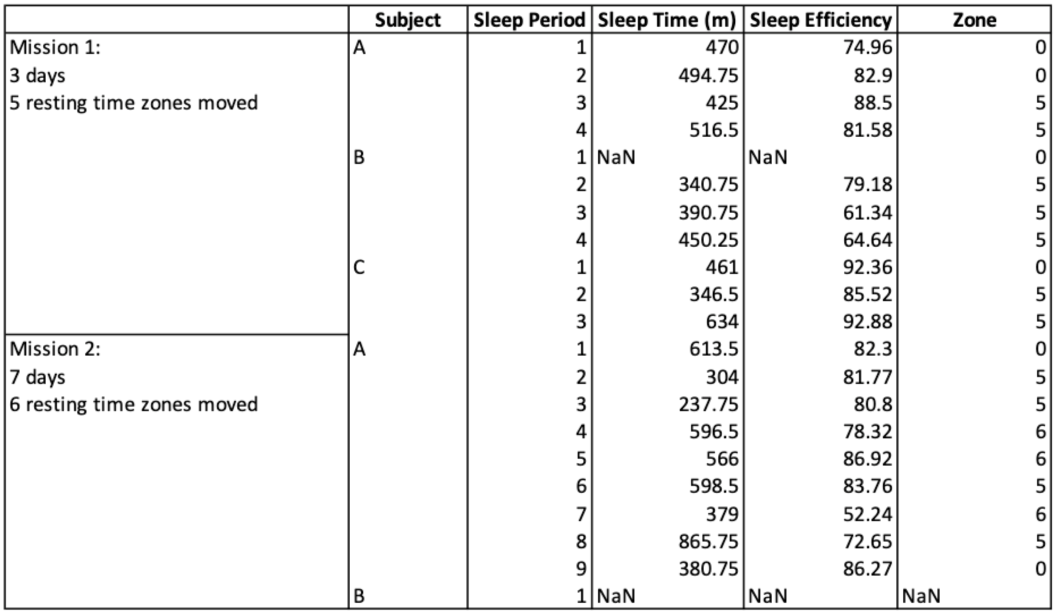
Small-Scale Study Data Analysis

As discussed in Chapter 2, the primary purpose for this study was to provide a ground truth database against which the regulations and models discussed in Chapter 2 can be validated. However, from a different perspective, we can examine the assumptions upon which the models are based and examine our data for consistencies with these assumptions. An assumption made by US Air Force regulation is that that sleep, regardless of when it happens, provides equal benefit in overcoming fatigue. Recent models employed by the FAA, however, consider disturbances in circadian rhythms, as these disturbances may decrease the ability to achieve restful sleep. Therefore, we can use the data from our small-scale study to explore whether sleep during missions is comparable to sleep prior to missions.

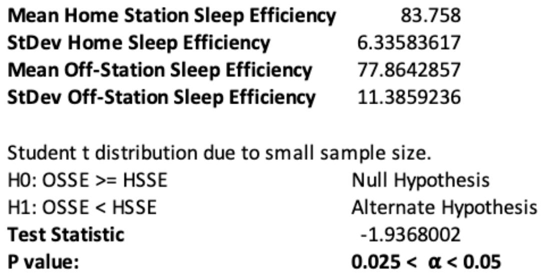
For the small-scale study, we gathered two measures which provide insight to sleep quality prior to or during missions. These include a quantitative measure of sleep efficiency as calculated from the Actiwatch and participant's rated quality of sleep. In its simplest form sleep efficiency is a ratio of the amount time an individual is actually asleep compared to the amount of time the individual is in bed. The Actiware software uses a threshold of movement to determine if a given period is spent awake or asleep.

The periods of sleep in which all aircrew were at home station with no disruption of circadian rhythm were taken as a normal population. The null hypothesis was that the mean of the off-station sleep efficiency for those same aircrew when they were greater than three time zones from where they were sleep adjusted would be equal to or greater than the home-station sleep efficiency. A Student's t distribution was used due to the small sample size of only fourteen off-station sleep periods across four subjects. The table below gives the sleep efficiency values for home station and off-station across all subjects and their associated means and standard deviations. The appendices contain a summary of sleep period sample data along with other values collected in the study. The result was a p value being between 0.025 and 0.05, so there is evidence that the sleep efficiency decrease was statistically significant.

Table 4 Data Summary



	Subject	Sleep Period	Sleep Time (m)	Sleep Efficiency	Zone
Mission 1: 3 days 5 resting time zones moved	A	1	470	74.96	0
		2	494.75	82.9	0
		3	425	88.5	5
		4	516.5	81.58	5
	B	1	NaN	NaN	0
		2	340.75	79.18	5
		3	390.75	61.34	5
		4	450.25	64.64	5
	C	1	461	92.36	0
		2	346.5	85.52	5
		3	634	92.88	5
Mission 2: 7 days 6 resting time zones moved	A	1	613.5	82.3	0
		2	304	81.77	5
		3	237.75	80.8	5
		4	596.5	78.32	6
		5	566	86.92	6
		6	598.5	83.76	5
		7	379	52.24	6
		8	865.75	72.65	5
		9	380.75	86.27	0
	B	1	NaN	NaN	NaN



Mean Home Station Sleep Efficiency	83.758
StDev Home Sleep Efficiency	6.33583617
Mean Off-Station Sleep Efficiency	77.8642857
StDev Off-Station Sleep Efficiency	11.3859236

Student t distribution due to small sample size.

H0: OSSE >= HSSE	Null Hypothesis
H1: OSSE < HSSE	Alternate Hypothesis
Test Statistic	-1.9368002
P value:	**0.025 < α < 0.05**

A large issue with this data beyond the small sample size is that it violates the assumption of equality of variance. There are two different crew on two different missions for two different lengths of time. The aircrew that was on a longer mission participated in more sleep periods, and thus produce a larger number of samples, which implies the sample is biased towards representing their experience more than the aircrew on the shorter mission. Further, some aircrew wore the devices for more periods of time that were used as samples, which provides further potential bias.

In addition, this test was not paired for each aircrew. The on-station and off-station sleep periods were put all together. Taking a mean of aircrew sleep efficiency averaging all aircrew together is reminiscent of Lieutenant Gilbert S. Daniel’s technical note published after studying the physical measurements of 4,000 Air Force aircrew. He wrote “The tendency to think in terms of the ‘average man’ is a pitfall into which many persons blunder when attempting to apply human body size data to design problems. Actually, it is virtually impossible to find an ‘average man’ in the Air Force population. This is not because of any unique traits of this group of men, but because of the great variability of bodily dimensions which is characteristic of all men” (Daniels, 1952). This same description of physical body dimensions applies to our individual differences in how we sleep, as described in the literature review.

With that in mind, it could potentially be more helpful to look at each individual’s response in sleep efficiency as a result of mission that includes time-zone change and any other number of factors to be studied. This would be better accomplished with a longer period of observation to better establish a baseline level of sleep. This would provide a large-sample normal distribution for each individual. The number of those individuals who show a statistically significant decrease in sleep efficiency for a given factor could then be discussed. The result of this with even few aircrew subjects would be likely be more interesting than a small number of mission and sleep samples for a large number of aircrew. Therefore, it is recommended that future studies, consider monitoring over a longer period, even if this resulted in fewer individuals being studied. This would necessitate a requirement for the study that uses a wearable device the subject is willing to wear for a long period of time without inconvenience.

In addition, a larger data set that could measure a wide variety of options would then open up similar tests to above, but possibly with a fractional design to incorporate multiple factors that even if they could not have hypothesis tests performed on them, could give strong indications of large two-way interactions between different factors.

Finally, a larger data set across individuals could also help balance a naturalistic study such as this. It would allow the selection of an equal number of samples under a given set of conditions, compared to this study where every piece of data was used, even if one subject had a two-week mission compared to a four-day mission.

Proposed Study Analysis

Measures

Movement

If polysomnography is the “gold-standard” of measuring sleep time and quality in controlled laboratory studies, the measurement of movement – usually referred to “actigraphy” – is the gold-standard of wearable and less intrusive devices to measure the same. This is especially true for the Philips Actiwatch Spectrum Pro, which was used in the small study mentioned before (Roomkham et al., 2019). Significant research has been performed to compare different actigraphy devices and their algorithms, showing that actigraphy alone is relatively effective in predicting sleep time and efficiency.

Limitations are typically seen around the time of waking, which actigraphy will often designate as sleep while the subject is not in a state of rest (Russell et al., n.d.). In all consumer devices considered and researched here, actigraphy was used in combination with other sensor input such as heart rate and analyzed using proprietary algorithms to output sleep measures. These capabilities are discussed later for individual devices.

For the proposed operational naturalistic sleep study in aircrew, the ability to measure movement is required.

Heart Rate and Heart Rate Variability

As wearable devices have increased in popularity, two measures have been adopted across many as an indicator of physical work performed, a physical predictor of performance, and an indicator of recovery: heart rate and heart rate variability (HRV). HRV is the measure in variation in time between each heartbeat and is controlled by the autonomic nervous system (ANS) that sends signals to a variety of functions across the body. In fact, the ANS varies based on different stages of sleep, and this can be measured not just through electroencephalography, but through heart rate and its variance. This has led to the use of heart rate and heart rate variability as a supplement to actigraphy when making assessments of sleep-wake states (Roberts et al., 2020). HRV can be understood as the body’s ability to “change gears” where a high variability allows increased flexibility or resilience, while a low variation means systems are more stressed. As the measure has increased in use, research has seen relationships over the long term between low HRV and negative health outcomes such as depression, anxiety, and increased risk of cardiovascular disease (Campos, 2019).

In addition, research has illustrated a link between HRV and common measurements of fatigue. One study followed ten members of a wildfire service management team and tracked subjective fatigue measures, total sleep time using a wrist worn device for recording actigraphy and HRV, and reaction time tests. In this study there was significant inverse association between HRV and sleepiness and fatigue, as well as a positive association between HRV and sleep time. There was no significant association found between HRV and reaction times (Jecklin et al., 2021).

For the proposed operational naturalistic sleep study in aircrew, the ability to measure heart rate and heart rate variability is required. In addition to aiding sleep-wake state determination, it could be desirable to provide further insight between possible heart rate variability and aircrew fatigue and/or effectiveness.

Respiratory Rate

Many devices offer measurement of respiratory rate. This is not from any direct method, but rather through variations of the heartbeat that indicate a breath in and breath out. As a result, while it may be of some interest in the future to look at respiratory rate, it can be somewhat measured through the heart rate measurement requirement and does not need to be added as its own requirement.

Body Temperature

One easy and reliable measure for endogenous, or internal, circadian rhythms is through the measurement of core body temperature. “A person’s body temperature is highest in the late evening and will drop steadily until early morning when it begins to rise. This cycle persists even when the person is in an environment without time cues, such as the day/night light cycle, though it will drift toward a slightly longer period of

24.1 h.” This cyclical change is positively correlated to human performance of tasks that require manual dexterity, inspection, or monitoring. Performance of tasks can also be based on whether one classifies oneself as a morning or evening type. Morning types may have peak periods for their endogenous circadian rhythm earlier than an evening type (Proctor & Van Zandt, 2008).

While body temperature could be relevant to understanding of endogenous circadian rhythm for the aircrew, it remains difficult to measure in a unintrusive way. Skin temperature varies significantly compared to core temperature as the body regulates itself and addition skin temperature is affected by environmental temperature. To reliably measure core temperature would require wearable adhesives or sensor arrays (Dias & Paulo Silva Cunha, 2018). These types of sensors could impede aircrew during operations or place an extra burden outside of duty hours to configure or wear. Proprietary algorithms due exist to calculate body temperature based off skin temperature and heart rate. For the proposed operational naturalistic sleep study in aircrew, the ability to skin temperature and possibly thereby estimate core temperature could provide some value to estimate endogenous circadian cycles but is not required.

Light and Sound

Finally, to help understand possible sources of aircrew fatigue other than circadian rhythm disruption, it would be helpful to look at factors that could contribute to reduced rest, such as light and or sound exposure during sleep periods. While not required, it would be desirable to collect these variables in this operational study, so that if a sleep period appears significantly disturbed it could be looked at with more lenses than simply being a result of a long duty day across time zones.

Drug Use

Throughout the course of a mission aircrew may use caffeine during a duty day to aid in alertness, and use “no-go” pills or alcohol to aid in changing or establishing a circadian rhythm. For this reason, it would be useful to survey aircrew in the small-scale study. The ability to survey aircrew for factors such as this and any other factors that the researcher would want to analyze without the aid of a sensor is a requirement for the operational study proposed.

Wearable Device Feasibility Analysis

The number of consumer wearable devices that measure aspects of health have become prolific, with thousands of models currently being used by hundreds of millions people (Chinoy et al., 2021)   To make a comparison of such a wide variety of products feasible, many were eliminated based on straightforward requirements of aircrew. For example, one device that has become popular is a ring. While it’s low profile and ease of wear during sleeping would be ideal, Air Force aircrew are not permitted to wear hard rings and other types of jewelry while performing flight duties, so capturing data from naps during flight are not practical with any hard ring-shaped device.

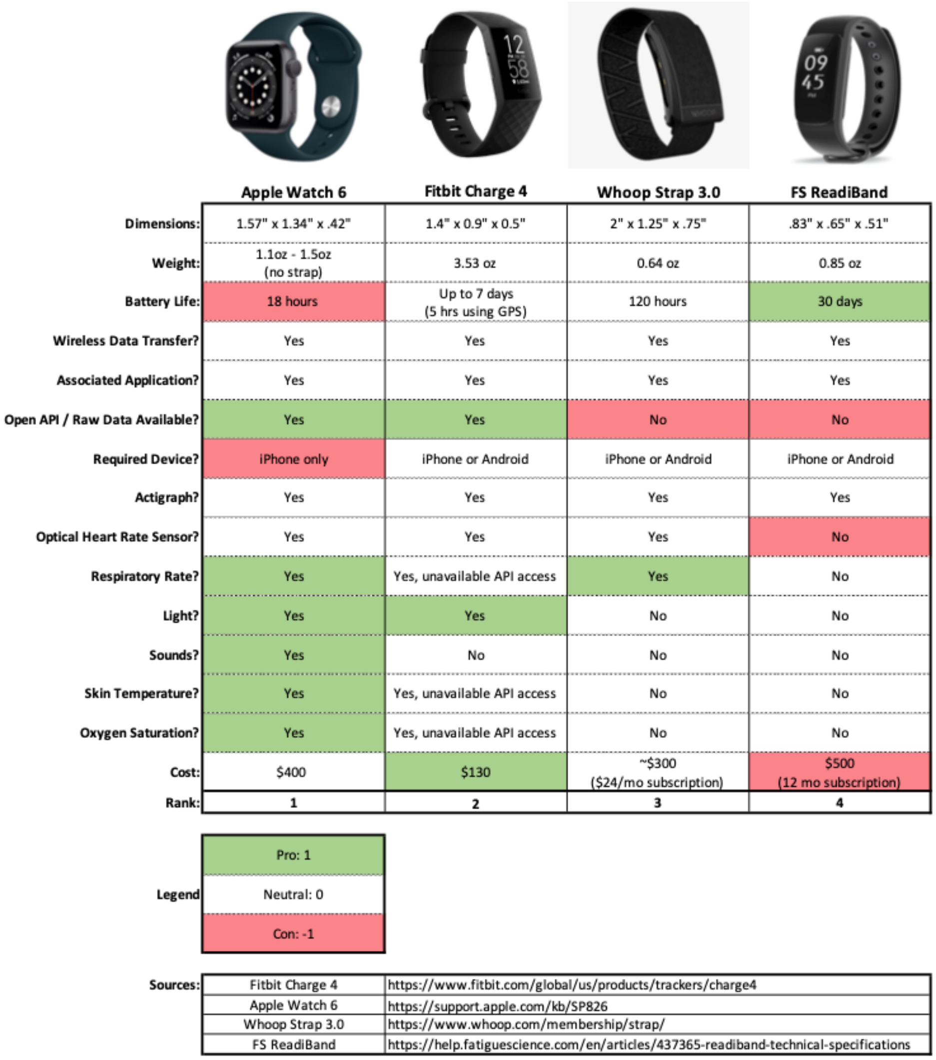
Another factor for deciding which wearables may be options for an operational sleep study is their performance as compared to polysomnography in previously accomplished studies. While actigraphy in almost all cases is better than self-reported sleep times, it can at times lack specificity specifically with how much time a subject is actually awake but perhaps in bed and not moving, as opposed to sleeping and getting actual rest when compared to polysomnography. The cumulative effects of overestimating actual sleep time could yield dramatically different results over a period of days or weeks (Russell et al., n.d.), which is the timeframe of typical mobility missions. Devices found lacking compared to others in their ability to classify time asleep versus time awake compared to basic actigraphy were removed from consideration.

One study looked at seven consumer wearables compared to polysomnography, and found the devices to be highly sensitive, but lacking in specificity. This means that they accurately detected sleep compared to PSG but were less accurate in detecting "wake." The current standard for mobile sleep detection of actigraphy only, such as those used in the small study described above with the Respironics Actiware, have been validated to be in line with PSG, but newer devices incorporating additional measures into sleep/wake states have outperformed actigraphy alone. The only two devices that performed worse than actigraphy in terms of specificity were the Garmin Fenix and Vivosmart. As a result, they were not considered as possible options in the proposed study. All devices tested were highly variable in their predictions of amount of time in different stages of sleep relative to PSG. (Chinoy et al., 2021)

In all cases with consumer wearable devices, there is a lack of transparency in algorithms used to calculate key measures, whether it's total sleep time, sleep efficiency, or heart rate. The raw data of the instruments are not provided, and the algorithms transforming that raw data into these measures are proprietary. A possible way to mitigate this issue would be to select one device to use in the study, however it is still possible that the algorithm used by the same device changes (Wetsman, 2021).

Figure 2 captures many of the key specifications of the wearable devices studied, and a more in-depth summary of each is done below. Each of the devices uses Bluetooth transmission to communicate with a smart phone device. All are permitted to wear in-flight but would need to be removed briefly during ground duties when aircrew receive classified briefs. In addition to the many capabilities of booth the Apple Watch and Fitbit discussed below, their widespread popularity would make them excellent choices for a future naturalistic study. Their APIs could also allow both to be used for the study if sufficient work was done to standardize algorithms used between them for calculating measures such as sleep state versus wake state.

Table 5 Wearable Device Comparison



	Apple Watch 6	Fitbit Charge 4	Whoop Strap 3.0	FS ReadiBand
Dimensions:	1.57" x 1.34" x .42"	1.4" x 0.9" x 0.5"	2" x 1.25" x .75"	.83" x .65" x .51"
Weight:	1.1oz - 1.5oz (no strap)	3.53 oz	0.64 oz	0.85 oz
Battery Life:	🟥 18 hours	Up to 7 days (5 hrs using GPS)	120 hours	🟩 30 days
Wireless Data Transfer?	Yes	Yes	Yes	Yes
Associated Application?	Yes	Yes	Yes	Yes
Open API / Raw Data Available?	🟩 Yes	🟩 Yes	🟥 No	🟥 No
Required Device?	🟥 iPhone only	iPhone or Android	iPhone or Android	iPhone or Android
Actigraph?	Yes	Yes	Yes	Yes
Optical Heart Rate Sensor?	Yes	Yes	Yes	🟥 No
Respiratory Rate?	🟩 Yes	Yes, unavailable API access	🟩 Yes	No
Light?	🟩 Yes	🟩 Yes	No	No
Sounds?	🟩 Yes	No	No	No
Skin Temperature?	🟩 Yes	Yes, unavailable API access	No	No
Oxygen Saturation?	🟩 Yes	Yes, unavailable API access	No	No
Cost:	$400	🟩 $130	~$300 ($24/mo subscription)	🟥 $500 (12 mo subscription)
Rank:	1	2	3	4

Legend	
🟩	Pro: 1
⬜	Neutral: 0
🟥	Con: -1

Sources:		
	Fitbit Charge 4	https://www.fitbit.com/global/us/products/trackers/charge4
	Apple Watch 6	https://support.apple.com/kb/SP826
	Whoop Strap 3.0	https://www.whoop.com/membership/strap/
	FS ReadiBand	https://help.fatiguescience.com/en/articles/437365-readiband-technical-specifications

Apple Watch 6

The Apple Watch is prolific in use and provides a wider variety of sensors than other wearables mentioned here. While not necessarily made to be worn overnight to monitor sleep, it has the functionality. In fact, when used for sleep monitoring, it performed with high accuracy (97%) and sensitivity (99%) in detecting sleep, and strong specificity (79%) in detecting wakefulness when compared to the Philips Actiwatch Spectrum Pro (Roomkham et al., 2019).

Documentation on how to develop for the Apple Watch is significant and would allow for custom app development for the project. In fact, much of the raw sensor data is available if a native app were to be developed for the Apple Watch and could be used for research (Walch et al., 2019), (Roberts et al., 2020). This could mitigate the previous issue mentioned with the lack of transparency in algorithms used to calculate certain measures. It could be possible to use public domain algorithms on the raw data or use machine learning techniques to create a new algorithm. While it may be less accurate relative to a proprietary algorithm, there would be value in consistency and transparency.

The primary disadvantage of the Apple Watch is its reduced battery life. While it does have a native sleep tracking application, that is not a primary use case and typically it is charged while the user is sleeping. This would require a subject to charge the watch during breaks during the day for it to be available at night for recording data. In addition, its dimensions are larger than other options, which could possibly increase interference with sleep or other activities.

Fitbit Charge 4

The Fitbit line of products are an extremely popular option, having 3% of wearable shipments in 2020 and over 30 million active users (Alsop, 2020). In addition Fitbit hardware and algorithms were a standout success in one independent study, performing best out of seven common consumer devices for measuring sleep and wake times, including the Actiwatch and Fatigue Science Readiband (Chinoy et al., 2021). Another study in 2018 using only the Fitbit Versa and comparing it to polysomnography in a naturalistic environment determined that it could be useful in measuring sleep duration in "longitudinal epidemiologic naturalistic studies albeit with some limitations in specificity." These limitations were noted to be particularly in defining the amount of sleep in specific sleep stages (Svensson et al., 2019).

It's open access API for developers is an inviting option for research, despite all its features not being available through the API like skin temperature. Actigraphy and heart rate data are available with minimal processing from onboard hardware (Roberts et al., 2020).

ReadiBand Actigraph

The Readiband Actigraph is made by Fatigue Safety, the same organization responsible for the SAFTE model used in many DoD applications, in particular the AvORM application that generates a graph for every mission of expected crew performance based on the schedule. The ReadiBand was mentioned earlier in the same independent study as the Fitbit, and while it did not perform as well overall, it performed better than simple actigraphy in predicting total sleep time and efficiency (Chinoy et al., 2021).

Separating the ReadiBand from other wearables, is the fact that it is part of the Readi platform that is designed as part of a system intended as a comprehensive fatigue risk management system. The system generates fatigue predictions based on modeling, and these predictions can be enhanced through the wearable ReadiBand device. A large part of its system is an application that allows supervisors to monitor expected operator performance and delivers alerts when the operator experiences significant fatigue. Through the ReadiBand, an operator can receive personal alerts on their fatigue levels. This can also be done with the operator using other wearable devices such as Fitbit or Garmin (*Fatigue Science,* n.d.).

The ReadiBand and its associated algorithm performed well in a non-independent study compared to other actigraphy pairs of wearables/algorithms: Actiwatch L20, Actiwatch L40, the Cole-Kripke algorithm and the Lötjönen algorithm. This was true specifically in specificity of wake states, with 55% specificity compared to PSG. The study points to the cumulative issues that can arise with devices and algorithms of lower specificity over a period of days in predicted fatigue levels of a subject, when they consistently overestimate how much sleep a subject receives due to lack of specificity (Russell et al., n.d.).

The ReadiBand has a version specifically for researchers. However, the band itself seems less robust than others. While it may compare favorably with its actigraphy, it does not collect any other type of data that could be useful in developing insights beyond actigraphy such as heart rate or heart rate variability. In addition, the manufacturer notes that the band is not suitable for showers or swimming (*Fatigue Science,* n.d.). For a wearable device to be suitable in a naturalistic operational study, it would be useful for the device to be able to be less restrictive in the environments allowed for the wearer to not be made aware of its presence with multiple restrictions. More useful than the device by itself in research could be the underlying algorithm that appears to do better in predicting actual wake times than other pure actigraphy options.

Whoop Strap 3.0

The Whoop Strap is a small low-profile band that is meant to be worn continuously, with a focus on athletic performance and recovery. In one study that measured the ability of wearables to improve sleep habits, polysomnography was used to validate that the wearable data was valid for the subjects. The Whoop strap was found to measure sleep duration, measured dream sleep, and slow wave sleep accurately. In addition, error was found to be low for heart rate and respiratory rate, 1.5% and 6.7% respectively (Berryhill et al., 2020)

Whoop strap is a popular choice among athletes, with a focus on "recovery" vs "strain." Its proprietary algorithms combine heart rate variability and actigraphy to produce scores for how rested and prepared for strain your body is, and then during the day how much strain or work performed.

It has a minimalist low profile look that works with what people already wear and is easy for a wearer to get used to and sleep with. It is simply a small black strap, and not much else. It has pop culture appeal as well as it is advertised by large influencers.

With a minimalist appeal comes minimalist sensor ability. Its measurement is limited to heart rate, heart rate variability, and movement. Finally, it does not appear to have an open API that could be used if an app was developed to use the Whoop strap.

Domain Model Class Diagram

Figure 2 depicts a possible domain model for the operational naturalistic sleep study, illustrating the entities involved, with associated functions and data. Some data sources included but not previously discussed include the use of the Mobility Air Forces (MAF) Scheduling System, Global Decision Support System (GDSS), as well as the Air Force Safety Center system for aircraft data analysis, Military Flight Operations Quality Assurance (MFOQA).

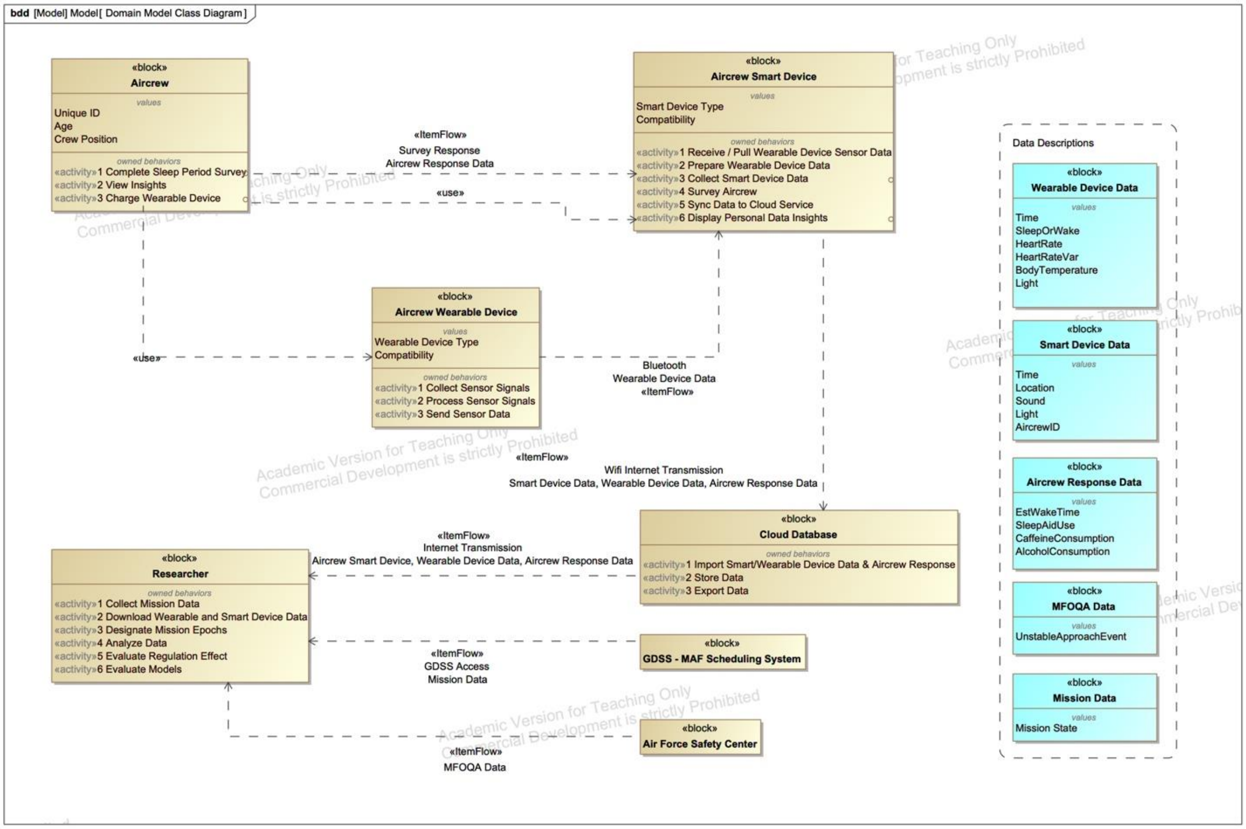


Figure 2 Domain Model Class Diagram

Using GDSS, a researcher would be able to associate data epochs with the state of an aircrew, whether they were on a mission or not. This could also include more specific states, such as whether the aircrew is in a designated crew rest time either before, during, or after a mission, or set on alert but not actually conducting flight operations. This information would support parsing the data based upon the state of the aircrew to permit sleep baselines to be established prior to a mission, understanding of crew rest both flights, understanding crew rest while in designated crew rest time during or after a mission, as well as understanding fluctuations in circadian cycle throughout the mission.

In addition, the Air Force Safety Center (AFSEC) uses data obtained from the aircraft itself and processes the data using gatekeepers to ensure it is not associated with individual aircrew members. One key measure provided by AFSEC is whether an aircraft complies with stabilized approach criteria. These criteria are used primarily to examine times or locations where these criteria are less likely to be met to mitigate hazards imposed by procedures or aspects of the airfield. However, these criteria also provide possible measures of aircrew effectiveness. If these criteria were made available for individual aircraft in an operational sleep study, it would be possible to compare sleep efficiency and other measures to how likely that an aircraft was to adhere to stabilized approach criteria, providing a direct link between the measures of sleep and circadian cycle to aircrew performance during a critical phase of flight. It should be noted, however, that AFSEC serves as a gatekeeper for this data to permit it to be used to reduce potential hazards without penalizing aircrew for lapses in performance induced by environmental or system factors. Therefore, access to this data may be restricted outside of AFSEC, which implies that it may be necessary for the safety center to be provided access to the sleep study data so that they can merge and disassociate specific aircrew members information from the data before releasing results from the data analysis. In fact, having AFSEC serve as the gatekeeper for the data from the sleep study may reduce aircrew concerns with participation in the study and the potential misuse of the data.

Requirements Diagram

Figure 3 below compiles lessons learned from the small-scale operational study, as well as the data gathered from the analysis of measures utilized in existing devices for the sleep study. Requirements derived from the study focus on enabling gathering of data, how to get a large enough sample size through participation and wearability of the device. Requirements derived from analysis of possible measures focus on obtaining those measures that could provide value or insight in answering the question of whether current fatigue-related regulation are effective in mitigating fatigue risk.

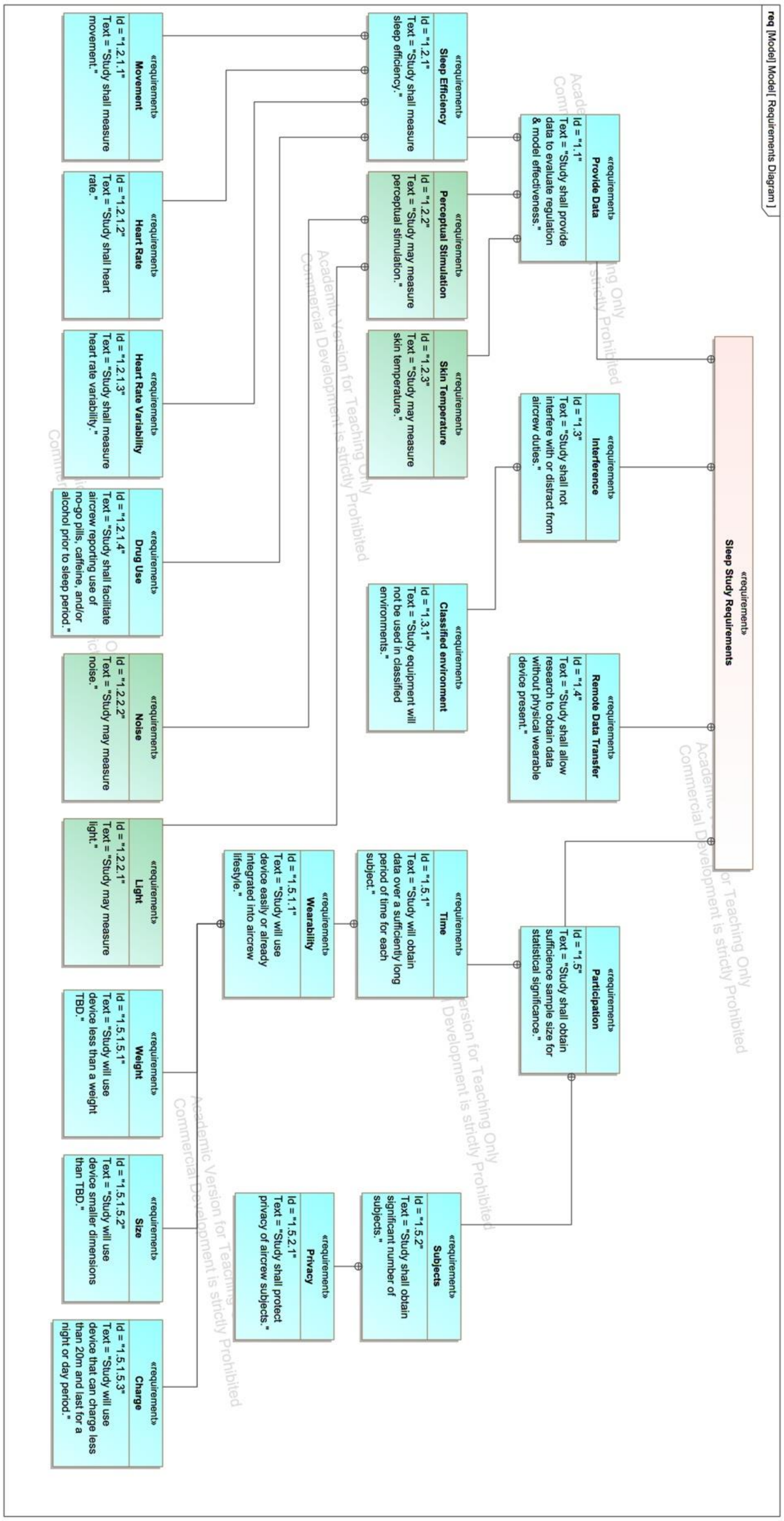


Figure 3 Requirements Diagram

V. Conclusions and Recommendations

Conclusions of Research

This research determined that an operational naturalistic sleep study is practical. Lessons learned from conducting a small-scale study, as well as analysis of alternative means of measuring fatigue-related indicators, were used to inform the development of a high-level system architecture of a potential sleep study.

In conducting the small-scale study, the collection of data in the operational environment was difficult. This was a result of a variety factors, as indicated. First, aircrew did not want to wear a measurement device that had no apparent function, sometime requiring wearing of the device together with a watch. Therefore, they often removed the device when in public settings and believed this corrupted their data enough that they were less motivated to wear the device afterwards. Secondly, it was difficult to collect the data from the device as the aircrew and the experimenters had to coordinate handoffs. In the operational environment this sometimes led to months of delay between aircrew returning from a mission, the devices being returned, data downloaded, and assigning the devices to another aircrew on an appropriate mission. Finally, the additional duties required of the aircrew often resulted in the aircrew failing to complete experimental tasks. In the small-scale study, aircrew sometimes did not fill out the brief survey after a sleep period, often ignoring this task until the end of a mission when other matters were not seen as more pressing. Unfortunately, this delay caused them to struggle to remember details and therefore they felt their responses were not very accurate. Finally, this study illustrated that assigning devices to crews for single missions results in data sets with significantly different durations due to the variability in missions. This variability makes it difficult to accurately integrate or compare data across these individual missions.

This study informed the development of future study requirements. Analysis of options for wearable devices was performed which illustrated that many consumer-grade options provide sufficient accuracy and specificity for measuring wake and sleep states and other measures to provide valuable data. In addition, these wearable devices are often already used by aircrew and provide functionality and style that they desire to wear, rather than wear out of a sense of obligation. The presence of these devices could enable a study to be performed across months or longer for individuals and scale relatively easily across larger numbers of participants. The devices discussed all pair to a smart phone device through Bluetooth, and this smart phone when connected to a data signal could synchronize data remotely, removing the barrier of the measuring device being physically present to transfer data.

Significance of Research

By applying analysis and lessons learned from this research, a study could be designed and conducted that provides valuable insight on the effectiveness of fatigue-related regulation, especially as it relates to long-haul mobility operations. In addition, it could provide general and individualized data from an operational environment that can be lacking in fatigue-risk management systems, as the models employed by these systems are often based on generalized, controlled, laboratory studies.

As previously mentioned, the consumer-grade measurement devices can be paired with the use of smart phone devices already used by aircrew. In implementing the study system described in this research, a software application would need to be constructed to send data from the device to the researcher or safety personnel. However, small further investments in this application could provide large returns in the quality of data obtained and the ease in obtaining it.

There was difficulty in obtaining reliable survey data from aircrew, as they would often forget to fill out the small card provided. A simple application that knows when a sleep period has ended could provide a push notification to aircrew that prompts them to fill out the small survey on their smart phone. Provided that the notification is provided in a thoughtful way, this could provide a less intrusive way to collect valuable survey data that does not require aircrew to carry extra material or remember to take the survey after sleep.

It would also be possible for the application to retrieve and transform raw sensor data, applying a standardized algorithm across different devices and enabling the possibility that data could be obtained even if aircrew use their own personal and different types of wearable devices and smart phones. This could decrease hardware cost of the study, increase aircrew participation, and remove the need to rely on proprietary platform-specific algorithms that are not transparent to researchers.

Finally, such an application could increase the value of participation to aircrew and the Air Force by providing personalized insights. Once an individual baseline is established, the application could give basic analysis of how different factors related to the survey effect something such as sleep efficiency. This feedback could return value to the operator in providing an understanding of how they as an individual may respond to different fatigue countermeasures and implement practices in response that mitigate the risk of fatigue during operations.

Recommendations for Action

Although the original intent of this research was to explore methods for conducting individual human factors studies against which to validate fatigue models, this research led to the awareness that the available technology has matured to the level that it is now possible to integrate fatigue monitoring into the larger Air Force safety system. Thus, the system illustrated in the class diagram shown in Figure 2 could be further integrated with AFSEC systems and used as a new pillar of proactive safety measures. This capability would further augment AFSEC's programs which analyze trends across multiple platforms using aircraft data to identify high risk practices or types of approaches, to include fatigue monitoring and measurement within its capabilities. This augmented system would permit AFSEC to analyze data across aircrew themselves to identify the areas where there is the highest risk for a mishap due to fatigue. The coupling of fatigue measures discussed as part of this research with aircraft flight safety data being used within AFSEC could yield further insights, as the aircraft data is truly the measure of an aircrew's effectiveness and the sleep and fatigue data would aid the understanding of root causes of performance issues, particularly for mobility air crew. This coupling is also important as the pilot community generally trusts AFSEC to provide feedback on performance without attribution and this trust will be important in incentivizing the pilots to wear the devices necessary to support this function.

Additionally, work should be conducted within the pilot community to understand the features that such a system could provide to the pilot community so that the system would provide direct value for the pilots themselves. For example, functions which provide the flight commander information on each crew member’s circadian cycle and sleep history may provide information useful in managing crew rest. Further, the crew may benefit from feedback on practices which permit easier circadian entrainment throughout a mission.

Recommendations for Future Research

While use of an application described above could be of high value to the Air Force in regard to risk mitigation, there is room for further research to determine what insights could be drawn from adopting this system across the crew force, generating a much larger sample and stronger conclusions about the effectiveness of current fatigue-related regulation.

The focus of this study was on long-haul operations and how time-zone change can affect sleep efficiency. However, Air Force mobility aircrew are subjected to wide variety of different schedules and conditions that a non-intrusive study such as the one proposed here could examine. While it is common for mobility aircrew to be assigned missions that last one to two weeks and extend across three continents, it is just as likely for that same crew to perform small 3-day missions, hopping between multiple bases within the continental United States. Then a week later fly on a normal day schedule back home, followed by not showing up to work until evening for a night tactical training event that lasts until the early hours of the morning, showing up to work again once their twelve-hour crew rest time is complete and getting back to a normal day cycle.

This wide variety of missions present a large opportunity to learn about how different scenarios affect aircrew. In addition, it was previously discussed that fatigue risk management system models do not consider individual factors. Further research is required with a period of sufficient length to establish a strong baseline to look at how individuals respond to different conditions. This would be extremely valuable in providing feedback to existing models, as well as possibly providing insight to individual crew on the method of least resistance necessary for them to mitigate the risk of fatigue.

Finally, while actigraphy has been significantly validated in measuring sleep efficiency, other measures are not as well established in how they relate to real-world outcomes. Further research in an operational environment could attempt to look at heart rate variability or respiratory rate to determine if there is a significant link between those factors and other outcomes.

Summary or Significance of Research

Fatigue in aircrew is a significant risk to flight operations. Research has been performed to understand what causes fatigue as well as what it effects, and mathematical models have been developed to attempt to predict those effects under certain conditions. However, this research and these models have little operational data available to validate or refine them for real-world application, especially at an individual level. In addition, it is not clear that existing Air Force or FAA regulations regarding fatigue management are effective or sufficient, especially when it comes to long-haul operations across multiple time-zones. These reasons create a significant need for an operational naturalistic study of aircrew.

Lessons learned from a small-scale study conducted in this research, as well as analysis of different data collection devices led to the development of a possible architecture for an operational naturalistic sleep study. Such research could mitigate issues seen with data collection and lack of aircrew participation to build a long-term data set with a large subject population that could be used to better understand general and individual fatigue-related factors, such as circadian rhythm shift. This system could be scaled and integrated with existing AFSEC systems to act as a proactive safety measure from the individual aircrew level to the mobility Air Force community.

Bibliography

2019–2020 NTSB Most Wanted List of Transportation Safety Improvements: Reduce Fatigue—Related Accidents. (2019). National Transportation Safety Board. https://www.ntsb.gov/safety/mwl/Pages/mwlfs-19-20/mwl2-fsa.aspx

AFMAN 11-202 Volume 3, Air Mobility Command Supplement, Flying Operations. (2020). HQ USAF/A3.

AFMAN 11-202 Volume 3, Flight Operations. (2020). HQ USAF/A3.

Air Force Manual 11-2C-17 v3, Flying Operations, C-17 Operations Procedures. (2019). HQ AMC/A3V.

Alsop, T. (2020, October 28). *Fitbit—Statistics and facts*. Statista. https://www.statista.com/topics/2595/fitbit/

ASAP Data Query CY17-20. (2021).

Aviation Safety Reporting System Database Query. (2021). https://asrs.arc.nasa.gov/overview/database.html

Berryhill, S., Morton, C. J., Dean, A., Berryhill, A., Provencio-Dean, N., Patel, S. I., Estep, L., Combs, D., Mashaqi, S., Gerald, L. B., Krishnan, J. A., & Parthasarathy, S. (2020). Effect of wearables on sleep in healthy individuals: A randomized crossover trial and validation study. *Journal of Clinical Sleep Medicine*, *16*(5), 775–783. https://doi.org/10.5664/jcsm.8356

Campos, M., MD. (2019, October 22). Heart Rate Varaibility: A New Way to Track Well Being. *Harvard Health Blog*. https://www.health.harvard.edu/blog/heart-rate-variability-new-way-track-well-2017112212789

CFR Final Rule—Flightcrew Member Duty and Rest Requirements (Volume 77, Number 2). (2012). FAA, DOT; Federal Register. https://rgl.faa.gov/Regulatory_and_Guidance_Library/rgFinalRule.nsf/861ae0b1f7efc3ee85256453007b0e8a/681787ac6e53df238625797c005321de

Chaiken, S. (2005). *A Verification and Analysis of the USAF/DoD Fatigue Model and Fatigue Management Technology:* Defense Technical Information Center. https://doi.org/10.21236/ADA443545

Chinoy, E. D., Cuellar, J. A., Huwa, K. E., Jameson, J. T., Watson, C. H., Bessman, S. C., Hirsch, D. A., Cooper, A. D., Drummond, S. P. A., & Markwald, R. R. (2021). Performance of seven consumer sleep-tracking devices compared with polysomnography. *Sleep*, *44*(5), zsaa291. https://doi.org/10.1093/sleep/zsaa291

Cooper, H. J., M. D. (1930). *The Relation Between Physical Deficiencies and Decreased Performance*. Department of Commerce.

Daniels, G. S. (1952). *The “Average Man”?* [Technical Note].

Dias, D., & Paulo Silva Cunha, J. (2018). Wearable Health Devices—Vital Sign Monitoring, Systems and Technologies. *Sensors*, *18*(8), 2414. https://doi.org/10.3390/s18082414

Dinges, D. F. (2004). Critical Research Issues in Development of Biomathematical Models of Fatigue and Performance. *Aviation, Space, and Environmental Medicine*, *75*(3), A181–A191.

Fatigue Science. (n.d.). Fatigue Science. Retrieved July 4, 2021, from https://www.fatiguescience.com/

Federal Aviation Regulations. (2021). U.S. Department of Transportation.

Gregory, K., Martinez, S., & Flynn-Evans, E. (2017). *Validation of Fatigue Modeling Predictions in Aviation Operations*. 22.

Hursh, S. R., Redmond, D. P., Johnson, M. L., Thorne, D. R., Belenky, G., Balkin, T. J., Storm, W. F., Miller, J. C., & Eddy, D. R. (2004). *Fatigue Models for Applied Research in Warfighting*. *00*(0), 10.

Jecklin, A. T., Perrotta, A. S., Davies, H. W., Bredin, S. S. D., Paul, D. A., & Warburton, D. E. R. (2021). The association between heart rate variability, reaction time, and indicators of workplace fatigue in wildland firefighters. *International Archives of Occupational and Environmental Health*, *94*(5), 823–831.

Kirby, M. (2011). Finnair First with Pilot Fatigue Test. *Flight International*, *Jun 21-Jun 27*, 179.

Lowry, P. (2012). Are you Tired of Fatigue? *Mobility Forum: Journal of the Military Airlift Command*, *Spring 2012*, 4–7.

Martinez, S. B., & Quintero, L. O. (2015, August 6). *Validating and Verifying Biomathematical Models of Human Fatigue*. NASA Ames Research Center Internship Poster Session.

Miller, W. H., M. D. (1930). *Medical Problems in Air Transport Operations*. Department of Commerce.

Porter, H. B. (1936). Fatigue Factors in Special Relation to Flying. *Aviation Medicine*, 120–126.

Proctor, R. W., & Van Zandt, T. (2008). *Human factors in simple and complex systems* (2. ed). CRC Press.

Roberts, D. M., Schade, M. M., Mathew, G. M., Gartenberg, D., & Buxton, O. M. (2020). Detecting sleep using heart rate and motion data from multisensor consumer-grade wearables, relative to wrist actigraphy and polysomnography. *Sleep*, *43*(7), zsaa045. https://doi.org/10.1093/sleep/zsaa045

Roomkham, S., Hittle, M., Cheung, J., Lovell, D., Mignot, E., & Perrin, D. (2019). Sleep monitoring with the Apple Watch: Comparison to a clinically validated actigraph. *F1000Research*, *8*, 754. https://doi.org/10.12688/f1000research.19020.1

Russell, C. A., Caldwell, J. A., Arand, D., Myers, L. J., Wubbels, P., & Downs, H. (n.d.). *Validation of the Fatigue Science ReadibandTM Actigraph and Associated Sleep/Wake Classification Algorithms*. Archinoetics, LLC.

Salas, E. (Ed.). (2010). *Human factors in aviation /[ed. By] Eduardo Salas* (2. ed). Academic Press.

Shin, O., Gregory, K., Kirkley, C., Costedoat, G., & Flynn-Evans, E. (2018). *Evaluation of the Validity of Bio-Mathematical Models in Predicting Fatigue in an Operational Environment*. 1.

Svensson, T., Chung, U., Tokuno, S., Nakamura, M., & Svensson, A. K. (2019). A validation study of a consumer wearable sleep tracker compared to a portable EEG system in naturalistic conditions. *Journal of Psychosomatic Research*, *126*. https://www.sciencedirect.com/science/article/pii/S0022399919306804

Van Dongen, H. P. A. (2004). Comparison of Mathematical Model Predictions to Experimental Data of Fatigue and Performance. *Aviation, Space, and Environmental Medicine*, *75*(3), A15–A36.

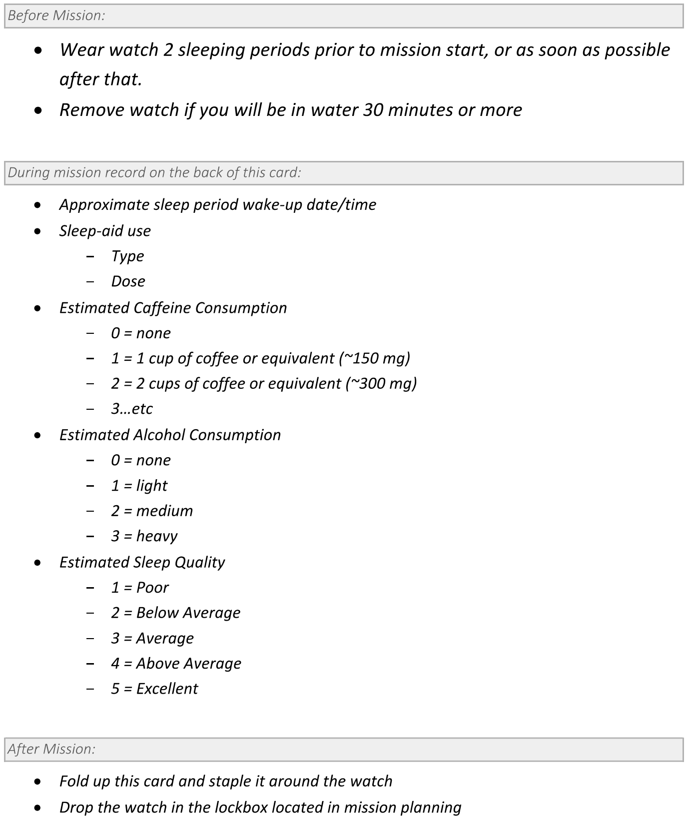
Walch, O., Huang, Y., Forger, D., & Goldstein, C. (2019). Sleep stage prediction with raw acceleration and photoplethysmography heart rate data derived from a consumer wearable device. *Sleep*, *42*(12), zsz180. https://doi.org/10.1093/sleep/zsz180

Weir, J. (2002). *A Three Phase Approach to Solving the Bidline Generation Problem with an Emphasis on Mitigtating Pilot Fatigue Through Circadian Rule Enforcement*. Georgia Institute of Technology.

Wetsman, N. (2021, July 27). *Apple Watch’s data ‘black box’ poses research problems*. The Verge. https://www.theverge.com/2021/7/27/22594178/apple-watch-data-research-heart-rate-reliability

Appendix

A1: Aircrew Survey Data Card



Before Mission:

- *Wear watch 2 sleeping periods prior to mission start, or as soon as possible after that.*
- *Remove watch if you will be in water 30 minutes or more*

During mission record on the back of this card:

- *Approximate sleep period wake-up date/time*
- *Sleep-aid use*
  - *Type*
  - *Dose*
- *Estimated Caffeine Consumption*
  - *0 = none*
  - *1 = 1 cup of coffee or equivalent (~150 mg)*
  - *2 = 2 cups of coffee or equivalent (~300 mg)*
  - *3…etc*
- *Estimated Alcohol Consumption*
  - *0 = none*
  - *1 = light*
  - *2 = medium*
  - *3 = heavy*
- *Estimated Sleep Quality*
  - *1 = Poor*
  - *2 = Below Average*
  - *3 = Average*
  - *4 = Above Average*
  - *5 = Excellent*

After Mission:

- *Fold up this card and staple it around the watch*
- *Drop the watch in the lockbox located in mission planning*

A2: Small-Scale Study Data Summary

Data Summary

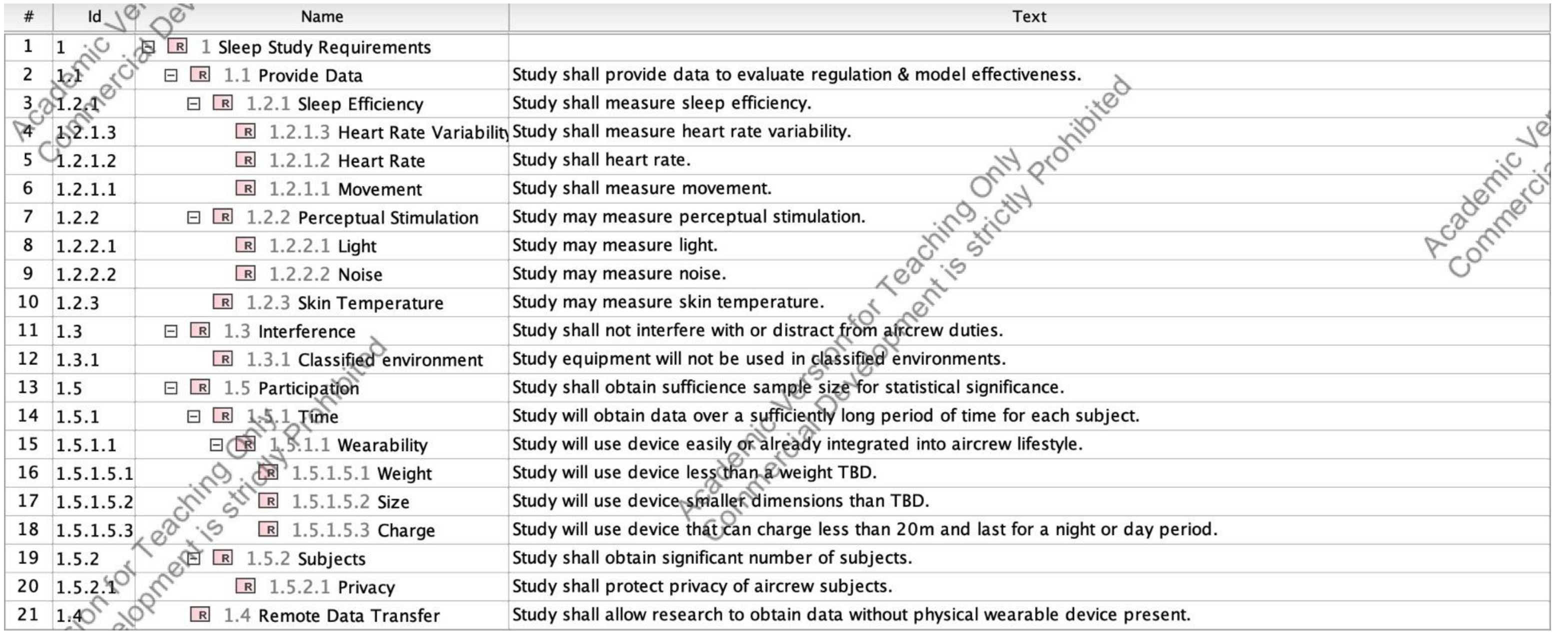
Mission 1:									
3 days	Subject	Sleep Period	Sleep Time (m)	Sleep Efficiency	Zone	Sleep Aids	Caffeine Prior	Alcohol Prior	Rated Quality
5 resting time zones moved	A	1	470	74.96	0	0	0	0	3
		2	494.75	82.9	0	0	2	1	3
		3	425	88.5	5	0	2	1	2
		4	516.5	81.58	5	0	2	3	4
	B	1	NaN	NaN	0	0	0	0	3
		2	340.75	79.18	5	0	0	1	2
		3	390.75	61.34	5	0	0	2	1
		4	450.25	64.64	5	0	0	0	3
	C	1	461	92.36	0	NaN	NaN	NaN	NaN
		2	346.5	85.52	5	NaN	NaN	NaN	NaN
		3	634	92.88	5	NaN	NaN	NaN	NaN
Mission 2:	A	1	613.5	82.3	0	0	0	2	4
7 days		2	304	81.77	5	0	0.5	2	4
6 resting time zones moved		3	237.75	80.8	5	0	0.5	3	4
		4	596.5	78.32	6	0	0	1	3
		5	566	86.92	6	0	0	0	4
		6	598.5	83.76	5	0	0	2	5
		7	379	52.24	6	0	0.5	0	4
		8	865.75	72.65	5	NaN	NaN	NaN	NaN
		9	380.75	86.27	0	NaN	NaN	NaN	NaN
	B	1	NaN	NaN	NaN	NaN	NaN	NaN	NaN

Mean Home Station Sleep Efficiency	83.758
StDev Home Sleep Efficiency	6.33583617
Mean Off-Station Sleep Efficiency	77.8642857
StDev Off-Station Sleep Efficiency	11.3859236

Student t distribution due to small sample size. Sample standard deviation significantly different than population standard deviation

H0: OSSE >= HSSE	Null Hypothesis
H1: OSSE < HSSE	Alternate Hypothesis
Test Statistic	-1.9368002
P value:	0.025 < alpha < 0.05

A3 Requirements Table



#	Id	Name	Text
1	1	⊟ R 1 Sleep Study Requirements	
2	1.1	⊟ R 1.1 Provide Data	Study shall provide data to evaluate regulation & model effectiveness.
3	1.2.1	⊟ R 1.2.1 Sleep Efficiency	Study shall measure sleep efficiency.
4	1.2.1.3	R 1.2.1.3 Heart Rate Variabilit	Study shall measure heart rate variability.
5	1.2.1.2	R 1.2.1.2 Heart Rate	Study shall heart rate.
6	1.2.1.1	R 1.2.1.1 Movement	Study shall measure movement.
7	1.2.2	⊟ R 1.2.2 Perceptual Stimulation	Study may measure perceptual stimulation.
8	1.2.2.1	R 1.2.2.1 Light	Study may measure light.
9	1.2.2.2	R 1.2.2.2 Noise	Study may measure noise.
10	1.2.3	R 1.2.3 Skin Temperature	Study may measure skin temperature.
11	1.3	⊟ R 1.3 Interference	Study shall not interfere with or distract from aircrew duties.
12	1.3.1	R 1.3.1 Classified environment	Study equipment will not be used in classified environments.
13	1.5	⊟ R 1.5 Participation	Study shall obtain sufficience sample size for statistical significance.
14	1.5.1	⊟ R 1.5.1 Time	Study will obtain data over a sufficiently long period of time for each subject.
15	1.5.1.1	⊟ R 1.5.1.1 Wearability	Study will use device easily or already integrated into aircrew lifestyle.
16	1.5.1.5.1	R 1.5.1.5.1 Weight	Study will use device less than a weight TBD.
17	1.5.1.5.2	R 1.5.1.5.2 Size	Study will use device smaller dimensions than TBD.
18	1.5.1.5.3	R 1.5.1.5.3 Charge	Study will use device that can charge less than 20m and last for a night or day period.
19	1.5.2	⊟ R 1.5.2 Subjects	Study shall obtain significant number of subjects.
20	1.5.2.1	R 1.5.2.1 Privacy	Study shall protect privacy of aircrew subjects.
21	1.4	R 1.4 Remote Data Transfer	Study shall allow research to obtain data without physical wearable device present.

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 074-0188*

1. REPORT DATE *(DD-MM-YYYY)* 19-08-2021	**2. REPORT TYPE** Master's Thesis	August 2017 – August 2021

TITLE AND SUBTITLE REQUIREMENTS ANALYSIS AND ARCHITECTURE FOR AN OPERATIONAL STUDY OF FATIGUE IN USAF MOBILITY AIRCREW	**5a. CONTRACT NUMBER** **5b. GRANT NUMBER** **5c. PROGRAM ELEMENT NUMBER**
6. AUTHOR(S) Mecham, Jonathan F., Major, USAF	**5d. PROJECT NUMBER** **5e. TASK NUMBER** **5f. WORK UNIT NUMBER**
7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S) Air Force Institute of Technology Graduate School of Engineering and Management (AFIT/EN) 2950 Hobson Way, Building 640 WPAFB OH 45433-7765	**8. PERFORMING ORGANIZATION REPORT NUMBER** AFIT-ENV-MS-21-S-076
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Lt Col Kiser, 469 FTS/CC 835 20th Ave Bld 2320 Sheppard AFB, TX 76311	**10. SPONSOR/MONITOR'S ACRONYM(S)** **11. SPONSOR/MONITOR'S REPORT NUMBER(S)**

12. DISTRIBUTION/AVAILABILITY STATEMENT
DISTRUBTION STATEMENT A. APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

13. SUPPLEMENTARY NOTES
This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States.

14. ABSTRACT

Aircrew fatigue in flight operations is a known hazard that has driven the creation of fatigue-reducing regulation and fatigue risk management systems industry wide. In addition, biomathematical models have been created and tested to forecast the effectiveness of aircrew under conditions of time-zone shifts and long duty days. However, limited operational studies exist to validate these models or to help understand how individual factors can affect them. Operational studies have a variety of limitations that make gathering typical data regarding fatigue or sleep difficult. This research takes systems requirement analysis approach to design a study that measures effects of circadian disruption on USAF C-17 Aircrew effectiveness. This study could then aid in understanding the effectiveness of fatigue-related regulation and fatigue risk management systems used in Air Force Mobility operations. The current research develops requirements of such a study through analysis of existing research as well as through a small-scale study to identify limiting factors for conducting such a study in an operational mobility squadron. The research further suggests additional research to explore the inclusion of fatigue monitoring into the Air Force Safety System, particularly for Air Force Mobility Operations.

15. SUBJECT TERMS
Fatigue, sleep, circadian rhythm, circadian disruption, fatigue risk management system

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Dr. Michael Miller, AFIT/ENV
U	U	U	UU	56	**19b. TELEPHONE NUMBER** *(Include area code)* (937) 255-6565, ext 4651 (Michael.miller@afit.edu)

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39-18