2014-12

Fatigue mitigation and crew endurance management in the Royal Australian Navy and the U.S. Navy: a review of recent efforts and a collaborative path forward

Ryan, Katrina J.

Monterey, California: Naval Postgraduate School

http://hdl.handle.net/10945/44656

NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

FATIGUE MITIGATION AND CREW ENDURANCE MANAGEMENT IN THE ROYAL AUSTRALIAN NAVY AND THE U.S. NAVY: A REVIEW OF RECENT EFFORTS AND A COLLABORATIVE PATH FORWARD

by

Katrina J. Ryan

December 2014

Thesis Advisor:	Nita Lewis Shattuck
Co-Advisor:	Michael Smith

Approved for public release; distribution is unlimited

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	December 2014	Master's Thesis

4. TITLE AND SUBTITLE
FATIGUE MITIGATION AND CREW ENDURANCE MANAGEMENT IN THE ROYAL AUSTRALIAN NAVY AND THE U.S. NAVY: A REVIEW OF RECENT EFFORTS AND A COLLABORATIVE PATH FORWARD

5. FUNDING NUMBERS

6. AUTHOR(S) Katrina J. Ryan

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Postgraduate School
Monterey, CA 93943-5000

8. PERFORMING ORGANIZATION REPORT NUMBER

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

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. IRB Protocol number ____N/A____.

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited

12b. DISTRIBUTION CODE
A

13. ABSTRACT (maximum 200 words)

Formulating naval manpower requirements is a complex problem. The results from workload studies can assist in this endeavor in two ways, by improving endurance and performance for existing crews and by providing manpower planners with the information necessary to determine or validate crew composition and size. Many naval workload studies have been conducted for a variety of reasons. Results from sleep measurement studies in naval operations indicate widespread sleep deprivation. Scientific research shows that insufficient and/or poor quality sleep leads to reduced individual performance and decreased crew endurance, an unacceptable situation for any navy.

This thesis reviews field and simulator studies from the Royal Australian Navy, the United States Navy, the Royal Canadian Navy, and the civilian maritime sector. Major gaps in the research include the assessment of sleep quality onboard ships, the formal design and conduct of evaluative research rather than descriptive efforts, and organizational level fatigue management policy and education. Fatigue risk management systems are still evolving but have not matured to a level guaranteeing that sailors routinely receive adequate, acceptable quality sleep at sea. Best practices for future collaborative studies are suggested, and recommendations for a ten-year Royal Australian Navy research strategy are included.

14. SUBJECT TERMS
actigraphy, crew endurance, crew performance, fatigue management, fatigue mitigation, manpower requirements, sleep, workload studies

15. NUMBER OF PAGES
95

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UU

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

FATIGUE MITIGATION AND CREW ENDURANCE MANAGEMENT IN THE ROYAL AUSTRALIAN NAVY AND THE U.S. NAVY: A REVIEW OF RECENT EFFORTS AND A COLLABORATIVE PATH FORWARD

Katrina J. Ryan
Lieutenant Commander, Royal Australian Navy
B.Sc., University of New South Wales (ADFA), 1993
MBA, Charles Sturt University, 2004

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN MANAGEMENT

from the

NAVAL POSTGRADUATE SCHOOL
December 2014

Author: Katrina J. Ryan

Approved by: Nita Lewis Shattuck
Thesis Advisor

Michael Smith
Co-Advisor

William Gates
Dean, Graduate School of Business and Public Policy

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

Formulating naval manpower requirements is a complex problem. The results from workload studies can assist in this endeavor in two ways, by improving endurance and performance for existing crews and by providing manpower planners with the information necessary to determine or validate crew composition and size. Many naval workload studies have been conducted for a variety of reasons. Results from sleep measurement studies in naval operations indicate widespread sleep deprivation. Scientific research shows that insufficient and/or poor quality sleep leads to reduced individual performance and decreased crew endurance, an unacceptable situation for any navy.

This thesis reviews field and simulator studies from the Royal Australian Navy, the United States Navy, the Royal Canadian Navy, and the civilian maritime sector. Major gaps in the research include the assessment of sleep quality onboard ships, the formal design and conduct of evaluative research rather than descriptive efforts, and organizational level fatigue management policy and education. Fatigue risk management systems are still evolving but have not matured to a level guaranteeing that sailors routinely receive adequate, acceptable quality sleep at sea. Best practices for future collaborative studies are suggested, and recommendations for a ten-year Royal Australian Navy research strategy are included.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

ADO	Australian Defence Organisation
AMI	Ambulatory Monitoring, Incorporated
CEMS	Crew Endurance Management System (USCG)
CF	Canadian Forces
DOD	Department of Defense
ECG	electrocardiogram
EEG	electroencephalogram
EOG	electrooculography (records eyeball movements)
ESS	Epworth Sleepiness Scale
FAA	Federal Aviation Administration
FAST	Fatigue Avoidance Scheduling Tool
FMD	Fleet Management Diary, see also NMD
FRMS	fatigue risk management system
IARC	International Agency for Research on Cancer
IMO	International Maritime Organisation
IMPRINT	Improved Performance Research Integration Tool (U.S. Army Research Laboratory)
KSS	Karolinska Sleepiness Scale (Swedish)
NATOPS	Naval Air Training and Operating Procedures Standardization (USN)
NGN	New Generation Navy (RAN)
NPS	Naval Postgraduate School (USN)
NMD	Navy Management Diary (RAN), formerly Fleet Management Diary (FMD)
NSMRL	Naval Submarine Medical Research Laboratory (USN)
NSWW	Navy Standard Work Week (USN)
ONR	Office of Naval Research (USN)
POMS	Profile of Mood States (questionnaire)
PSQI	Pittsburgh Sleep Quality Index
PVT	psychomotor vigilance test

RAN	Royal Australian Navy
REM	rapid eye movement
SOAP	Special Operations Assessment Profile (CF)
SSS	Stanford Sleepiness Survey
STCW	standards of training, certification and watchkeeping
TEST	Task Effectiveness Scheduling Tool (U.S.)
TST	total sleep time
USCG	United States Coast Guard
USN	United States Navy
WHO	World Health Organization

ACKNOWLEDGMENTS

This thesis would not have existed without the assistance provided by my advisors. Without CAPT Smith, there would be no topic, and without Professor Shattuck, I would have no knowledge of that topic, and without her encouragement the knowledge may not have found its way onto the page. Thank you, and rest assured that I will be doing my part to pass on the importance of sleep to others.

My parents have always encouraged my academic endeavors, and both they and my parents-in-law have supported our family in every way possible over many years. Thank you for all that you have done. As it would be impossible to think of this thesis without thinking of you, Dad, it is dedicated in loving memory of Keith Tunley (9 November 1947–12 November 2014).

Saving the best for last … the most thanks is due to my family, Sean, Sarah and Beth. While Monterey is a wonderful place to live, all of them were somewhat inconvenienced by this relocation, and as a bonus got a grumpier wife and mother, especially when deadlines loomed. I love you all, and appreciate all that you have done to allow me to complete this course of study at NPS.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. PROBLEM STATEMENT

This thesis will focus on two aspects of the seagoing environment with regard to naval workload studies: improving endurance and performance for existing crews and providing manpower planners with the information necessary to determine or validate crew composition and size. Formulating manpower requirements is a complex problem. It requires accurate knowledge of the number of people and types of skills needed to complete a given set of tasks or missions over a designated time period. For navies, the total number of personnel that can be accommodated onboard a platform is just one of the limiting factors in the manpower determination process. Even if space availability is not an issue, government and key decision makers may choose to conserve monetary resources by restricting the size of seagoing crews, thereby leaving more work for fewer crewmembers. The requirement for efficient use of resources means that an optimized solution addressing resource constraints and needs of personnel must be sought. Tradeoffs will always be necessary, but care needs to be taken to ensure that the long-term health and safety of personnel is not compromised.

Consideration of tradeoff issues should utilize an evidence-based decision-making process. Workload and fatigue studies provide information that can assist with manpower requirements decisions. A number of workload studies have been conducted by the Royal Australian Navy (RAN) and the United States Navy (USN) over the last 15 years, complementing work related to maximizing crew endurance in many other areas and industries. These studies, including those by Grech, Roberts, Hamilton, Turner and Cleary (2014), Haynes (2007), Mason (2009), Roberts (2012), Yokeley (2012), and Young (2013), have increased the body of knowledge regarding naval workload and highlight the negative impact of fatigue on performance. Results of RAN and USN fatigue-related workload studies have been used to suggest changes to work routines, particularly alternative watchkeeping routines at sea, to allow personnel to increase their sleep time and reduce fatigue. Where implemented, the recommended changes have had a positive impact (Cordle & Shattuck, 2013; Grech et al., 2014). Other maritime organisations and navies have also undertaken similar studies; for example, a multinational collaborative effort was conducted on fatigue at sea (Lutzhoft, Thorslund, Kircher, & Gillberg, 2007) and studies have examined fatigue management in the Canadian Forces (Paul, Gray, Nesthus, & Miller, 2008; Paul, Ebisuzaki, McHarg, Hursh, & Miller, 2012).

The body of knowledge on crew endurance has been built over many years and includes research by the United States Coast Guard (USCG) (Comperatore, Rivera, & Carvalhais, 2005; Miller, Smith, & McCauley, 1998) and studies conducted in the aviation, rail, and food service industries, as well as emergency medicine. These industries have similar complications to those of navies in terms of their requirement for 24/7 shiftwork, transit to and from the workplace, and an inability to replace critical workers immediately. Some of these results can be applied directly to the military environment. Other results are not as easy to apply due to the exigencies of military service, particularly for ships where performance is affected by additional influences, such as motion, compared to land-based workplaces.

Senior decision makers are rarely directly involved in the conduct of workload studies. Consequently, they may not have a detailed understanding of the tradeoffs or assumptions that have been used to inform the recommendations that are made to them. In addition, these tradeoffs and assumptions may not be the same for each workload study, even if the recommendations are similar. The increased risk that one decision maker is happy to accept may not seem reasonable to another decision maker, or may not be appropriate in a different work area. If tradeoffs and assumptions are not well communicated, there is a danger that decision makers may apply a result or recommendation from one study to a new situation where it is not applicable and/or appropriate.

It is difficult to program workload studies into naval operations because they often result in additional tasking for a vessel and its crew and they require resource allocation. Workload studies may be narrowly directed or somewhat fragmented, either by design or necessity. Since these studies involve human subjects, institutional review board approval is required prior to the conduct of a study, which can increase the required lead time. Best practices and recommendations can be collated for each of the studies, but a compilation of best practices is not currently available. In addition, analysis on how these practices could be amalgamated to ensure best practice for future workload studies is not complete. We do not know if improvements could be combined, or if they are supplementary rather than complementary. In short, we are not sure if there might be a better, or a simpler, way to make manpower determinations.

B. OBJECTIVE

This thesis reviews workload studies conducted in the RAN and USN over the last decade and makes recommendations for future workload studies that would utilize best practices in the field, narrowing gaps in the current body of knowledge. These additional studies will assist in achieving more accurate, relevant, empirically based decision support tools for use by manpower requirements analysts and decision makers.

C. SCOPE AND LIMITATIONS

A large amount of research on sleep, fatigue, and crew endurance has been conducted, and more is occurring all the time. Noting the voluminous amount of related research, several boundaries have been imposed to focus this thesis.

1. Scope

This thesis is not a wholesale analysis of workload studies in general, nor a meta-analysis of any kind. The major focus of this thesis is a comparison of RAN and USN studies of crew endurance as indicated by sleep and performance during the period 2001–2014, with an emphasis on the latter studies due to their currency in terms of personnel policies, resource constraints, manning and similar plans. This scope is limiting; many other navies with similar manpower constructs to the RAN and USN may have completed similar works. Exclusion from detailed consideration is not a reflection on the importance of these studies; rather they are outside the tight focus of this thesis.

2. Limitations

Multiple factors, including mental and physical fatigue, influence crew endurance. More sleep of higher quality can serve to reduce fatigue and thereby support crew endurance (Miller, Shattuck, & Matsangas, 2007). Examples of other important factors that support crew endurance are physical fitness, diet and nutrition, use of technology to reduce workload, reasonable living conditions, adequate manning levels that support more generous personnel tempo (for example, planning to have enough personnel for four sections of watchstanders, or shifts, rather than three), and higher levels of training, expertise, or cumulative at-sea experience. Ships also have additional environmental issues such as noise, vibration and/or motion due to sea state, all of which have been and continue to be studied. These other factors will not be considered in detail in this thesis as they are outside the thesis scope.

D. ORGANIZATION OF THE STUDY

Chapter II of this thesis provides background information regarding sleep and crew endurance, particularly in the maritime environment, and reviews relevant literature on these topics. Chapter III outlines a methodology for comparing the design, conduct, and results of workload studies. Chapter IV introduces selected workload studies conducted by the RAN, USN, and others, and provides a comparison utilizing the methodology previously described. Chapter V discusses the comparison and presents the conclusion and recommendations.

II. BACKGROUND AND LITERATURE REVIEW

A. SLEEP AND FATIGUE

Lack of sleep and fatigue are often thought of as similar, and even appear to be used interchangeably in some literature. Yet the Oxford Dictionary (2014b) illustrates that the meanings are quite distinct, defining sleep as “a condition of body and mind which typically recurs for several hours every night, in which the nervous system is inactive, the eyes closed, the postural muscles relaxed, and consciousness practically suspended.”, Sleep is something that one does, a condition in which one exists at certain times. Fatigue is also a condition, one “characterized by a lessened capacity for work and reduced efficiency of accomplishment, usually accompanied by a feeling of weariness and tiredness” (MedicineNet.com, 2014). Given that fatigue, by definition, is said to reduce efficiency and capacity for work, i.e., performance, and is accompanied by a feeling of tiredness, it can be implied that sleep would diminish fatigue, and therefore improve performance. In fact, many studies attest to this relationship. Miller, Shattuck, and Matsangas (2007) provide a concise summary regarding fatigue and its negative effect on performance in military environments. Numerous other works, including many Naval Postgraduate School (NPS) Masters theses, most recently, Roberts (2012), Brown (2012), Yokeley (2012), and Young (2013), discuss sleep and fatigue, and their application to the military environment in some detail. This thesis does not require an extensive knowledge of sleep; therefore, only a summary of the most relevant points will be provided.

The circadian cycle, or rhythm, is a naturally occurring pattern that corresponds roughly to the 24-hour day and predicts periods of increased or decreased alertness, which can also be described as energy levels (Comperatore et al., 2005, pp. 4–5). The circadian cycle is highly resistant to change (Miller et al., 2007). As can be seen in Figure 1, decreased alertness can be expected during the early hours of the morning, with a lesser dip in the early afternoon.

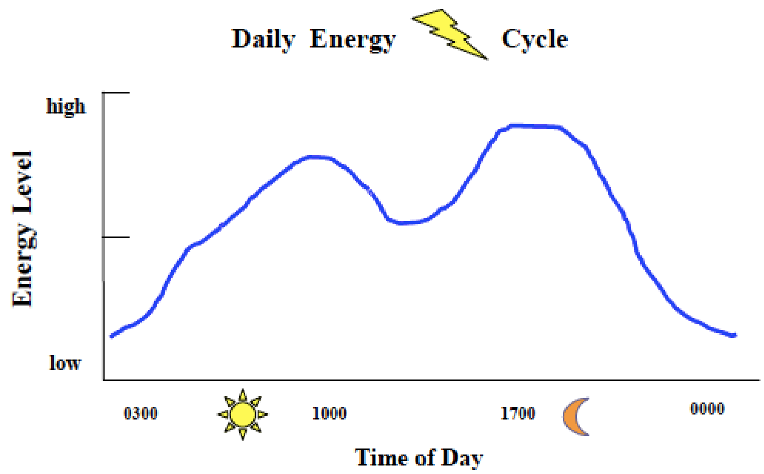

Figure 1. Daily energy level cycle (from Comperatore et al., 2005, pp. 4–5).

The average adult requires about eight hours sleep per night, which generally consist of four to six 90-minute cycles of deeper (Stage 4) or lighter (rapid eye movement or REM) sleep (Miller et al., 2007), as illustrated in Figure 2.

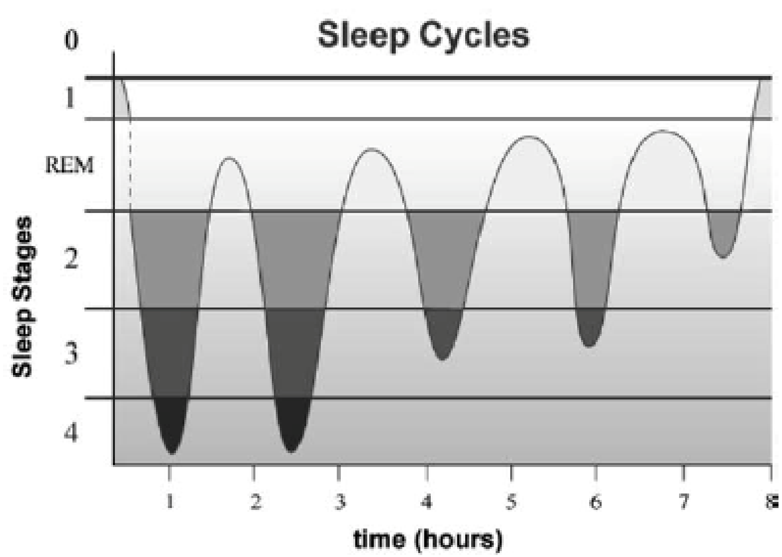

Figure 2. Sleep stages over a typical eight-hour sleep period (from Miller et al., 2007, p. 234).

Sleep requirements change over the human lifespan. Adolescents and young adults, those in their late teens and early to mid-20s, require approximately 0.5 to 1.25 hours more sleep per night. They also have a delayed sleep and wake time compared to other adults (Miller et al., 2007), as illustrated in Figure 3.

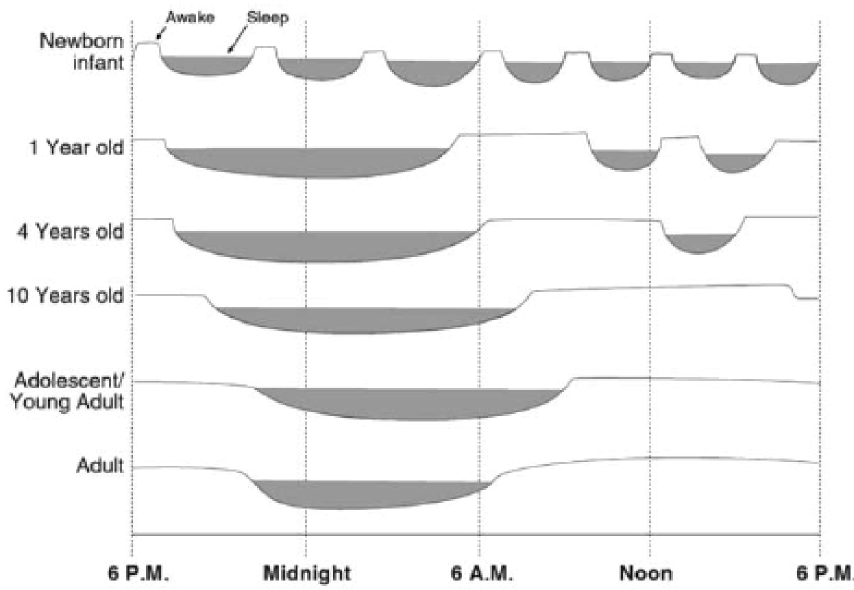

Figure 3. Sleep patterns over the human lifespan (from Miller et al., 2007, p. 233).

So far, the discussion has been about sleep occurring during the night, but there are many people who do not sleep during the night because of their work commitments, such as emergency services and after-hours healthcare personnel. There are many other examples of industries that employ shift workers; transportation, hospitality and other services, and of course, the military. Shiftwork often means that people are not able to achieve eight hours of sleep in a 24-hour period. Even if night shift workers intend to sleep during the day, it can be quite difficult for them to get adequate sleep. Emergencies or unplanned events may require some workers to remain at work after the end of their shift; distractions and sleep disruptions occur, for example, sleeping during the day may be difficult if you live next door to a school. Furthermore, family or other commitments that require you to be awake impact on planned sleep. When we consider shiftwork, and then superimpose the knowledge that the circadian cycle does not predispose personnel toward being at their most alert during the night, it is not surprising to discover that shift workers do not get eight hours sleep per day, and that the performance of shift workers may be impaired. This is especially true for those on night shift, shifts longer than eight hours, and shifts with insufficient breaks (Folkard, Lombardi, & Tucker, 2005; Folkard & Tucker, 2003; Huey & Wickens, 1993).

Having established that people do not always get enough sleep, it is important to understand why this might be the case. Shiftwork may be one reason, or increased work hours—either out of desire or necessity. The "can-do" attitude of the military culture and/or personnel shortages may contribute to insufficient sleep. Further, why do we care? First, the risk of errors or accidents increases when people are sleep deprived or if there are insufficient personnel to complete a job (Lazzaretti, 2008). Not having the right number of people can lead to sleep deprivation for those who are there—setting up a vicious circle. As individuals, most would have to admit to operating at less than their optimal performance level when they have not had enough sleep. Mandated aviation crew rest cycles were formulated to increase safety for aircraft and their passengers. Nurses driving home after the end of their shifts were found to have impaired driving performance, especially after night shift (Ftouni et al., 2013). Lazzaretti (2008) reported a negative relationship between manning levels and mishaps in the USN Oliver Hazard Perry Class Frigates, and there is no reason to suppose that this result is not generalizable. A second factor that must be considered is the longer term health and duty of care. Employers have a responsibility to not cause long-term poor health in their employees, and people are not as likely to take, or retain, jobs where this impact is likely. A recent study on retired shift workers by Monk, Buysse, Billy, Fletcher, & Kennedy (2013) found that their "subjective and objective measures of sleep showed a detrimental effect of shift work exposure." Gallicchio and Kalesan (2009) completed a meta-analysis regarding sleep duration and mortality, finding a link between all-cause mortality and shorter (less than seven hours' sleep per night) sleep duration. Advice from the World Health Organization (WHO) is that sleep deprivation is a Level 2A carcinogen (International Agency for Research on Cancer (IARC) Working Group, 2010). Any one of these factors should be reason for concern, but when they are combined, there is a compelling need to do all that is possible to ensure sleep deprivation does not occur.

B. CREW ENDURANCE

Should one search online for "crew endurance," there are likely to be hits relating to the crew of the 1914 Imperial Trans-Antarctic Expedition ship *Endurance*. While this association is merely a linguistic coincidence, it is worthwhile to consider this story of survival against the odds, described by Browning (2007) in a case study as follows: "Ernest Shackleton led a crew on the Endurance that would attempt to be the first group of individuals to cross the Antarctic continent overland. Only one day's sail away from the land, the Endurance became "iced in" and eventually sank, leaving the men with limited supplies. Shackleton's new goal quickly became getting every man home alive." There is perhaps no better way to define crew endurance than to think of Shackleton's trip, together with the Oxford Dictionary (2014a) definition of endurance as "the capacity of something to last or to withstand wear and tear." In any organization there is a requirement to complete tasks, and in militaries in particular, to get everyone home alive. In addition, we want our people, our crews, to be able to last, to withstand wear and tear.

The phrase crew endurance may thus be a measure of the ability of a group to get things done without lasting detriment to themselves, to complete required tasks, or meet goals (either original or amended), within the constraints imposed by the number and/or skills of the people available. In a ship, there is seldom a capacity to provide immediate replacements, so some additional capacity in terms of personnel is needed to ensure that missions can be safely completed. Being able to predict that amount accurately, translate it into a finite number of people and then fit all the required people into the confines of a warship, is the task of those who formulate naval manpower requirements.

Once a manpower requirement has been determined, the only way to improve the endurance of a crew within that determination is to find a means of improved performance, enhanced productivity, greater efficiency, or some similar process. There are many ways these effects can be achieved. For example, better use of technology, offloading tasks to another department or a shore-based organization, additional training or a combination of all of these methods can contribute to crew endurance. Performance can also be improved by minimizing fatigue (Miller et al., 2007), through ensuring that each individual gets a sufficient amount of quality sleep at an appropriate time each day. Some ways that this could be achieved are by changing routines, enforcing rest times, or providing incentives for being well rested. Workload studies are one means of measuring sleep, and sometimes also performance, in order to demonstrate increased productivity, or make recommendations regarding crew endurance.

C. WORKLOAD STUDIES

This thesis concentrates on the workload studies method of quantifying crew endurance through the objective measurement of individual sleep and, where possible, performance in varying situations. Once data is obtained regarding sleep patterns, analysis can show which groups of personnel are more likely to be under greater stress in each type of anticipated situation. Simulation tools can be used to predict failure points and alternative scheduling arrangements that would improve endurance times.

1. Objective Measurement of Sleep

Sleep is often objectively measured using wrist actigraphy. In this method, the motion of a person is monitored, via wrist monitors similar in size to a wristwatch, such as the Ambulatory Monitoring, Incorporated (AMI)[1] or Philips Respironics motionloggers used in recent RAN and USN studies (Davey, 2013; Grech et al., 2014). When using actigraphy, periods of inactivity can be categorized as sleep. For greater accuracy, particularly for personnel who may have a job that entails them remaining stationary for long periods of time, the data taken from the actigraphy device can be compared with a self-reported diary of sleep and wake times. The standard actigraphy report for workload studies provides an estimate of the amount of sleep per 24-hour period, which can be averaged across groups and/or time periods. The amount (or lack) of sleep is then used as an indicator for fatigue. Sleep may also be entirely self-reported; care must be taken using this method as self-reported sleep may overestimate the amount of time slept, as consistently demonstrated in a study by Mason (2009), probably due to the time it takes to fall asleep (sleep onset). Generally, actigraphy is viewed to be a more accurate measure of sleep than a self-reported, or even observed, sleep diary (Ancoli-Israel et al., 2003; Lauderdale, Knutson, Yan, Liu, & Rathouz, 2008; Tryon, 2004). Unlike actigraphy, a sleep-wake diary does not allow for an objective measurement of sleep quality.

Current actigraphic devices are a step up, in terms of ease of use, from the early studies into human sleep, including those done by the USCG, which utilized traditional scalp electrodes with additional physiologic metrics to approximate polysomnograph methods for monitoring responses including brain function (EEG) and cardiac rhythm (ECG). While these methods remain appropriate for use in sleep laboratories and specific studies, the flexibility gained by use of a wrist monitor allows a mobile workforce to be monitored without detriment to their job or changes in behavior caused by participation in the study. Continual advances in technology mean that data collection for future studies is very likely to become even simpler and more efficient than it is now. Personnel could be fitted with a wristwatch that has actigraphy as a secondary functionality,

[1] "AMI provides unique instruments to objectively document long-term sleep, hyperactivity, daytime activity levels, fatigue, circadian rhythm, vigilance, and respiration as well as environmental light, temperature and sound measurements in ambulatory subjects" (AMI, 2014).

batteries are likely to last for longer, and data may be available to analysts on a continuous basis, perhaps even remotely, via wireless download.

In addition to measuring sleep, sleepiness information can be collected. This is important in terms of performance, with a link between reduced sleep at night and sleepiness the following day shown to exist in both lab and field conditions (Åkerstedt, Axelsson, Lekander, Orsini, & Kecklund, 2013). Sleepiness measurements can be objective, but in the field it is simpler to use a subjective measurement, usually using one of a number of validated sleepiness scales (or variations thereof) such as the Epworth Sleepiness Scale (ESS), Pittsburgh Sleep Quality Index (PSQI), or Stanford Sleepiness Survey (SSS). Lutzhoft et al. (2007) used electrooculography (EOG) in a maritime workload study, but concluded that this method, which records eyeball movements and uses the number of eye blinks as an objective sleepiness measure, did not add much value to the study. This finding is supported by another study, where the measurement of subjective sleepiness using the Karolinska Sleepiness Scale was shown to be as sensitive an indicator of sleepiness as objective measures (Åkerstedt, Anund, Axelsson, & Kecklund, 2014).

Other tools used to look at sleep, sleepiness and performance also take into account general mood or happiness. These include the Canadian Forces Special Operations Assessment Profile (SOAP) which has six parameters (difficulty concentrating, level of depression, level of irritability, level of fatigue, work frustration and physical discomfort) (Paul et al., 2007), the Profile of Mood States (POMS), and surveys such as the Lark and Owl survey which categorize people according to their “morningness” or “eveningness” (Nguyen, 2002). The sleep measures discussed here are summarized in Table 1.

Table 1. Sleep measures (after Miller, Crowson, & Narkevicius, 2003, p. 718).

Name	Method	Where conducted	Expense	Ease of Use
Actigraphy	Motionlogger worn on wrist	Lab/field	$$	*
Polysomnograph	Scalp electrodes	Lab/field	$$$	***
Sleep and mood scales	Subjective rating (e.g. ESS, PQSI, SSS, KSS, SOAP, POMS, Lark/Owl)	Lab/field	$	*

$	Inexpensive	*	Minimal training required
$$	Somewhat expensive	**	Specialized training required
$$$	Expensive	***	Highly specialized technical training required

2. Measuring Performance

Reporting performance is more complex than reporting sleep, and potentially more subjective. For an assembly line worker, performance might be measured in terms of the number of items completed in a certain time. For other roles it is more difficult to measure performance but one method is to use a Psychomotor Vigilance Test (PVT). The PVT is a standard test that is completed at specified intervals during a workload study. The standard PVT is ten minutes in length, but field versions have been shortened to only three minutes, significantly simplifying participation, with no degradation in performance prediction (Basner, Mollicone, & Dinges, 2011). The AMI actigraphic device used in some RAN and USN studies has a built in PVT, which can be used to measure performance, in this case via response time. Participants must press a button when they see a stimulus—a light flash or a color change—with multiple stimuli offered at random intervals making up one test.

3. Industry Studies

Military organizations are very concerned about monitoring fatigue and increasing crew endurance, but they are certainly not the only industries that have an interest in this area of study. All of the industries previously mentioned as having shift workers, especially those in areas where they are dealing with members of the public or paying customers, are motivated to ensure that the performance of their personnel is at its best, and that employee fatigue does not impact on the safety of the individual worker, now or in the future, or those in their care.

Recent studies regarding the performance of shift workers are numerous; a few with results that may be broadly applicable are mentioned here. Brachet, David, & Duseja (2012) investigated the performance of paramedics in Mississippi, concluding that fatigue was the mediating factor in a deterioration in performance towards the end of long shifts. Their quantitative analysis suggested a 0.76 percent increase in 30-day mortality, or one person in every 132 treated by paramedics who did not survive due to paramedic fatigue. Many military ships or units have more than 132 personnel onboard. Even if only one of them is impacted by the fatigue of others it can have a large follow on impact, and if everyone onboard is fatigued then the impact is magnified.

The aviation industry has had regulations covering fatigue countermeasures since the 1930s (Miller & Strohl, 2011), but experts continue to recommend updates and refinements (Caldwell et al., 2009) arising from evolving technologies, significant increases in air traffic volume, both military and civilian, etc. Studies cover civilian and military sectors, multiple job roles including pilots, air crew, and flight attendants; and the duration/routing of flights, including how many time zones have been transited (Avers & Johnson, 2011; Banks, Avers, Nesthus, & Hauck, 2012; Caldwell et al., 2009; Gander, 1986; Gander et al., 2014; O'Connor, Buttrey, O'Dea, & Kennedy, 2011).

4. Naval Workload Studies

Similar to the civilian sector, military aviators have been compelled to adopt rule based fatigue countermeasures, and are literally unable to participate in a routine evolution if crew rest parameters are violated (United States Navy, 2009, pp. 8–15). This is similar to a situation that the reader may have experienced with some personal frustration where an airline crew for a delayed flight will have to be replaced if the delay means that they will "run out" of hours prior to the revised end time of the flight. The strict application of fatigue policy measures for naval sea-going personnel is not currently enforced, although this thesis will demonstrate support for that idea.

The USCG conducted much of the early maritime-based sleep research in the United States in the latter half of the 20th Century. Their findings have been refined over time and implemented into USCG day-to-day routines. It is now common practice, or business as usual, for the USCG to consider the effect of fatigue via the use of the *U.S. Coast Guard Guide for Managing Crew Endurance Risk Factors* (Comperatore et al., 2005). This is not yet the case for the RAN and the USN, although workload studies are being conducted and the results are being used to minimize fatigue.

D. USING WORKLOAD STUDY RESULTS

Workload study results have two major uses—improving current crew endurance, and predicting future manpower requirements. Improvements to current crew endurance result when recommendations are made and changes implemented that allow sleep-deprived personnel to get more quality sleep each day. This improvement could result from something as simple as changing a regular meeting time, or reducing the number of phone calls or announcements made during particular hours (Grech et al., 2014). Improvements of this nature are an example of increased productivity, a factor in many workforce negotiations. Unfortunately, improvements can be difficult to implement, even the aforementioned simple change to a meeting time can be difficult to accomplish due to organizational inertia or change resistance.

Predicting manpower requirements for future tasks or work areas is more complex; available information may be incomplete, and immediate verification regarding success cannot be obtained. Some indicators can be obtained through the use of bio-mathematical models which are in common use. Tools such as the Fatigue Avoidance Scheduling Tool (FAST), in extensive use in the U.S. Department of Defense (DOD), Australian Defence Organisation (ADO), and Canadian Forces (CF), use intended work-sleep schedules to forecast predicted task effectiveness, and the Task Effectiveness Scheduling Tool (TEST) suggests an optimal schedule for a group of personnel to follow given a required task list and a specified level of effectiveness (Tvaryanas & Miller, 2010). The RAN Navy Management Diary (NMD) is not yet fully operational, but is intended to fill a similar role (S. Hamilton, personal communication, July 30, 2014). The Army Research Laboratory Improved Performance Research Integration Tool (IMPRINT), and other similar simulation based tools, can provide a probability of success for a schedule of planned and unplanned events to be completed by a specified group of workers with particular skills.

Whether improving current endurance or predicting future endurance, there will be tradeoffs of many kinds required. One tradeoff relates to the amount of risk that is considered acceptable by an organization, or by different individuals within an organization. A person can be required to work for four hours longer and sleep for four hours less each day, but only if his or her employer is willing to accept the risk that the person may be fatigued to the point of being ineffective in the workplace or even a danger to themselves or others. One employer may be happy to accept a higher level of risk than another employer placed in the same situation. The tools mentioned previously are extremely useful, but their output may be completely wrong if the inputs, such as the level of acceptable risk, were inaccurate or inappropriate for a given situation. This discrepancy could arise simply due to different individuals having different ideas about what an appropriate level of risk is, or one crew having worked together for longer than another and therefore being slightly more proficient, or one team having many experienced personnel and another team having few. No matter what the reason is for any differences, or their magnitude, they exist in every situation and are applicable to factors other than risk. Tradeoffs and assumptions were made, constraints imposed, and the results impacted by them.

Every naval workload study considered here was conducted for a particular reason, under certain conditions, within certain constraints. The researchers made a large number of decisions when the study was formulated and conducted, and they may not have recorded each assumption, or the background behind it, exactly. The results are useful in that situation but cannot be relied upon for accurate predictions under different conditions unless the assumptions and limitations remain appropriate for that new situation. The only way to tell if workload study results can be generalized is to look closely at that study and the new scenario. Ideally, this task would involve an individual familiar with the original study in order to avoid any issues with non-documented assumptions or facts about the study. This use of subject matter experts is not always possible, and not everyone realizes that it is necessary; consequently there is a danger that workload study results could be, or have been, wrongly utilized or applied. This thesis will mitigate that risk by analyzing several workload studies, documenting the background to each, suggesting where gaps in the existing body of knowledge exist, and recommending the best practices to use for future workload studies. Understanding what the gaps are, and what assumptions have informed previous results, will allow results to be used to formulate policy founded on empirically based decision support mechanisms.

THIS PAGE INTENTIONALLY LEFT BLANK

III. METHODOLOGY

Military workload studies have been conducted for multiple reasons, and using different variables and methods. Despite these differences, all of the studies can be compared if the assumptions and limitations of each study are recognized and considered. Results from different studies may be generalizable if a new situation is similar to the original study environment or if those areas where it is different are not relevant to the comparison. This chapter describes the methodology that will be used to compare selected workload studies in order to identify areas worthy of further study and suggest best practice for future workload study involvement.

A. STUDIES COMPARED

This thesis considers only a small fraction of the workload studies that have been conducted within military environments. Within the studies that will be considered in detail, many different approaches have been used. Some studies were conducted to validate a crew composition, others to try and improve a situation that was considered less than optimal or to try and identify the best of a number of options, such as deciding between different watch standing routines. Each study was constrained by the availability of ships able to accommodate a study team, the equipment available, and the willingness of the crew to participate.

Every study compared includes the amount of time underway, either at sea or under simulated voyage conditions. Remembering that the scope of this thesis is to compare RAN and USN studies, these form the majority of the studies chosen, with a deliberate emphasis on more recent studies (since 2001) to ensure the most up-to-date methods are considered. Information from recent studies (where results have not yet been published) is included where possible. A cross section of aims and analysis types is desired, although there is some benefit in looking at studies with similar aims, conducted in differing conditions, to enable a recommendation regarding generalizability of results.

B. DESCRIPTIVE INFORMATION COMPARISON

A comparative tabular analysis methodology is used to present selected information for each study. The studies are arranged by platform type, and then described across three areas—study description (the “who, what, when, where”), study methods and variables (the “how”), and study conclusions (the “so what,” and “what next”). This method entails a substantial amount of simplification, and does not allow for all of the information about each study to be included in the table. The major benefits of this method are that it allows easy comparison of studies, with potential gaps being easier to recognize.

1. Platform Type

Different types of platforms have different capabilities, and the roles and responsibilities dictated by their operating concept may change over time. Platforms also differ in terms of operational tempo and crew size. All of these factors influence onboard workload; for example, both small and large ships may be required to launch and recover small boats for various reasons. In a ship with a small crew this evolution may involve most of the crew, including those who are off watch. With a larger crew it could be accomplished using only personnel who were already on watch, therefore, having less impact on the workload of off-watch personnel. The platform types used here are somewhat dictated by the literature available for comparison, but do consider similarities in crew size, and likely or actual role/tasking and operational tempo. Platform types are listed in Table 2, with indicative crew size provided as a means of classifying studies. Any shore-based simulation of time at sea is classified in the same way, with the addition of (Simulator) at the end of platform type, e.g., Civilian Ship (Simulator).

Table 2. Platform types and criteria for classification.

Platform Type (or Simulator)	Crew Size / Classification Criteria
Aircraft Carrier	Any aircraft carrier
Submarine	Any submarine
Large Warship	Warship with crew size greater than 70
Small Warship	Warship with crew size less than 70
Auxiliary Warship	Any naval auxiliary with civilian ship type role
Civilian Ship	Any civilian ship

Aircraft carriers and submarines are considered as separate groups despite the similarities that some sections within each would have with similar sections in a large warship, such as engine room staff (assuming similar power and propulsion arrangements). All groups could contain diverse platforms, and care must still be taken with generalization of results. For example, submarines can be very different; nuclear versus conventional propulsion particularly has an impact on crew size and work requirements.

This thesis does not consider any studies involving a platform that would be classified as an auxiliary warship, but the criteria is included for completeness. Aircraft and shore-based commands have been, and will continue to be, the subject of workload studies; however, these are outside the scope of this thesis.

2. Study Descriptions

The first part of the comparison looks at identifying data for the study. This is the "who" section, and includes information about the country or countries leading the study, the group/s leading the study, e.g., the RAN or the USN, the author/s of any published reports or results from the study, and the actual platform/s used during the study. The "what" section of the description is a short summary of the purpose of the study, often taken directly from the executive summary of the published report. The "when" section provides the year in which the study was conducted, and the month/day, if available. This section also includes the total duration of the study, and details any phased approach such as an in-port phase followed by a sea phase. Details regarding where (geographically) a study occurred are not always available, when provided they are included in this section, together with information regarding platform tasking.

3. Study Methods and Variables

This section is the most complex and variable in the comparison table, but arguably the most important as it includes the scientific methodology followed during the study, details about the data collection and analysis, and any known limitations of the study. The method or approach itself is listed in a few words describing the type of study that was conducted, such as observational or experimental, and what type of data was collected (e.g., time series).

This section includes descriptive data about the number of personnel on board the platform (either from the report or the generic crew size for the class of ship), the number of participants in the study, and the number of participants whose data was used for analysis. Participants are further described by gender, job role (officer or sailor), department, and watch routine; including the number from each department or in each watch routine. If all of this information is available then a participation rate can be calculated (data may not be valid if only a very small percentage of a crew was studied), as well as a reject or dropout rate. This is most likely to be important if the dropout rate was high as that could indicate that participant workload was so large that only above average performers had time to participate, which could mean that fatigue results are understated. If departmental and/or watch routine information is available then it is easier to check that a representative sample has been obtained for analysis, and also allows for analysis by department and/or routine. Other variables used in the study are also listed here.

Equipment and software used during the study are recorded with as much detail as possible, including version numbers, to enable comparison across studies. If surveys or questionnaires of any type, such as a sleep and activity log, are used, they are listed here as equipment. Validation of data ensures that it is reasonable to use the information collected, and analysis of data provides us with information about what happened during a study, details regarding both of these factors are included here if known. When analysis involves statistical analysis, individual techniques are not usually listed.

4. Study Conclusions

The conclusions section lists the major conclusions for the study, which are usually related back to the purpose of the study listed early in the table. The pros and cons, or challenges and limitations of a study are included here if mentioned by the researcher. In addition, if it is obvious when reading a report that a limitation existed or that a particular aspect of a study was or was not considered, then that is also included. This section also contains any relevant recommendations for further study, and other notes or information not already included that could be important or useful, including any information about a study made available following publication of the initial results.

C. CREW ENDURANCE RISK FACTORS COMPARISON

Each study is also discussed in terms of crew endurance risk factors, with each factor given a “yes” or “no” depending on whether a study considered the impact of a risk factor, either by directly reporting that the factor was considered, or if it can be inferred that the factor was considered. In some cases a numeric value is added. This additional enhancement recognizes that for some factors there are a large number of subject areas that are relevant to that factor, and allows differentiation between the studies. A higher value is awarded if the factor was quantitatively analyzed. The scale used is shown in Table 3:

Table 3. Crew endurance risk factors use and analysis scale.

0	not considered
1	partial consideration (1–2 subject areas from question list considered)
2	comprehensive consideration (3 or more subject areas considered) OR partial consideration and quantitative analysis
3	comprehensive consideration and quantitative analysis

The list of factors used is in Table 4, with the questions related to each factor being examples of the types of subject areas addressed or questions a study might provide answers to if that factor was considered as part of the study. The list of questions is not exhaustive, and there may be some overlap between categories, such as between sleep quality and environment.

Table 4. Crew endurance risk factors

Factor	
Sleep quantity	How much sleep is obtained in a 24-hour period? Is historical sleep data collected/available (for baselining)? Is napping possible? Is major sleep episode at same time each day?
Sleep quality	How many sleep episodes per 24-hour period? What time of day did participants sleep? Where did participants sleep? Usual location or different? New time zone? Were there any factor/s that inhibited sleep quality?
Work schedule	How many hours work? Are historical work schedules available? Are fatigue predicting tools updated with actual schedules? Rest breaks? Shift worker? What hours did you work (identify night workers)? What type of watch routine is being used? Do the watch sections rotate? If so, how often, which direction? Commuting time
Work type	Work location How physically or mentally demanding is the work? How stressful is the work? How repetitive is the work? Team or individual? If team based, how long has the team been working together
Organization & Culture	Work related stress e.g. from inflexible or un-empowering culture Rules-based organization? Clear procedures and policies? Information and training provided?

Factor	
Individual & Lifestyle	Health, including existing sleep disorders Diet/nutrition Use of caffeine/tobacco Consumption of alcohol/drugs (including legal medication such as sleeping pills) Opportunity for exercise/amount of exercise Personal stress e.g. from family, or from being away from home Social life Personal responsibilities Second job? Susceptible to motion sickness? Historical performance during continuous/sustained operations
Environment	Have environmentals such as weather, motion, noise, temperature been considered? Were any extreme? Were bunking arrangements considered e.g. type of bedding, private/shared rooms/bunks? Was light exposure noted, particularly just prior to sleep time?

This list is a simplified compilation of factors taken from official fatigue management/crew endurance/safety documents produced by several organizations. The original documentation includes the RAN factors that contribute to fatigue list, USN guidelines for aviation fatigue management, USCG crew endurance risk factors assessment form, Canadian forces fatigue risk management information, and Federal Aviation Administration (FAA) fatigue risk assessment tool and fatigue risk management system sample fatigue report. Further information on each of these contributing publications is provided in the following sections.

1. RAN Fatigue Factors

The Navy Safety Management System for the RAN is detailed in publication ABR 6303 (Commonwealth of Australia, 2014), the document includes a chapter relating to fatigue. The ABR 6303 list of factors that contribute to fatigue, and a checklist for investigating officers of factors to consider when ascertaining if fatigue was a causal factor in an incident, have been used in compiling the fatigue factors list used in this thesis. The original lists are not reproduced here due to distribution limitations.

2. USN Factors Affecting Aircrew Performance

Military aviation has a history of being more advanced than the surface fleet in terms of mandated crew rest, probably due to linkages to civil aviation standards. *The Naval Air Training and Operating Procedures Standardization (NATOPS) General Flight and Operating Instructions* (United States Navy, 2009) lists factors affecting aircrew performance. Those factors that are relevant for personnel in other work areas are presented in Table 5 (anthropometric and personal traits, for example, beards, corrective lenses for vision, have been omitted):

Table 5. Selected NATOPS factors affecting aircrew performance (United States Navy, 2009, pp. 8-15–8-23).

crew rest and sleep
circadian rhythm
flight time
nutrition, including nutritional supplements
exercise
drugs
illness
dental care (with respect to injectable drugs or intravenous sedatives)
pregnancy
emotional upset/excessive stress
immunizations and injections
blood donation
hypobaric or hyperbaric exposure
dehydration
simulator sickness (this could be expanded to motion sickness)
performance maintenance during continuous and sustained operations (this is a very individual factor, but of interest due to the links it has to resilience)

3. USCG Crew Endurance Risk Factors

The USCG Crew Endurance Risk Factors Assessment Form in Figure 4 is part of the USCG Crew Endurance Management System (CEMS). This system is considered to be the most mature of the fatigue management systems included here.

CREW ENDURANCE RISK FACTORS ASSESSMENT FORM

Write in the number of days per week (0-7) that each risk factor (Insufficient Daily Sleep, for example) occurs in your unit. Also note: (1) to which people or departments they apply; and (2) under what conditions they occur.

__ Insufficient daily sleep duration (less than 7-8 hours of *uninterrupted* sleep; **Ch. 4**)

__ Poor sleep quality (awakenings during main sleep period due to work-related disruptions, ship motion, or noisy environment; **Ch. 4**)

__ Sleep fragmentation (breaking sleep into multiple rest periods–"naps"– because unable to take a single, 7-8 hour sleep; **Ch. 4**)

__ Scheduling main sleep period during the day (the human body is designed to sleep at night; **Ch. 4**)

__ Changing work/rest schedules (rotating between working days and working nights one or more times per week; **Ch. 5**)

__ Long work days (exceeding 12 hours; **Ch. 5**)

__ No opportunities to make up lost sleep (napping during the day is not possible; **Ch. 4**)

__ Poor diet (menu includes frequent fried foods, high fat and sugar content, frequent caffeine consumption; **Ch. 9**, **Ch. 6**)

__ High workload (high physical and-or mental effort requirements; **Ch. 7**)

__ High work stress (caused by extreme environment, high sustained physical or mental workload, rotating work schedules, and-or authoritarian leadership style; **Ch. 7**)

__ Lack of control over work environment or decisions (workers are isolated and not allowed to contribute in problem identification and resolution; **Ch. 7**)

__ Excessive exposure to extreme environmental conditions (cold, heat, high seas; **Ch. 8**)

__ No opportunity for exercise (not enough time or no equipment/facilities; **Ch. 7**, **Ch. 9**)

__ High Family stress (child and parent care, divorce, finances; **Ch. 7**)

__ Isolation from family (need to know how family is doing; **Ch. 7**)

Figure 4. USCG Crew Endurance Risk Factors Assessment Form (from Comperatore et al., 2005, pp. 2-4).

CEMS is used throughout the USCG, with anecdotal evidence suggesting that, unlike the RAN and USN, personal sleep hygiene and command consideration of fatigue-related issues is the standard rather than the exception.

CEMS has been demonstrated to have successful application in other maritime industry areas such as the towing industry (United States Coast Guard, 2005). The earlier implementation of maritime fatigue management processes as a mandated system, and translation to similar industries, may be related to the roles of the USCG, including safety and rescue. The USCG has a vested interest in ensuring that other mariners are not impacted by fatigue, low numbers of accidents, investigation or rescues are good if the reason is that there is no need for them due to education and widespread implementation of fatigue management systems.

4. Canadian Forces Fatigue Risk Management

Defence Research and Development Canada provides general recommendations on fatigue risk management for the Canadian Forces under the headings listed in Table 6. Guidance is also provided on several topics with a notation that further research is required; these topics include nutrition, over-the-counter preparations (such as melatonin), bright light exposure, and exercise (Cheung, Vatanian, Hofer, & Bouak, 2010).

Table 6. Canadian Forces fatigue risk management recommendations (Cheung et al., 2010, pp. 19–35).

Identify and treat physiological sleep disorders
Minimize sleep loss by maintaining good sleep hygiene
Implement strategic naps or short sleeps
Anchor sleep (regular sleep period at least four hours long obtained at the same time each day)
Judicious use of caffeine
Guidelines for duty/rest scheduling
Guidelines on shift lag management
Guidelines for trans-meridian travel: how to manage jet lag
Guidelines for using sedatives
Guidelines for pharmacological stimulants

5. FAA Fatigue Risk Assessment Tool by Pulsar Informatics

The FAA Fatigue Risk Assessment Tool produced in conjunction with Pulsar Informatics (Pulsar Informatics & FAA, 2014) is shown in Figure 5. This is an online tool that uses historical sleep and work information for the preceding three days to produce a fatigue assessment for a potential work shift of interest. This tool uses very few of the fatigue factors in the assessment list produced for this thesis, but many of the other factors are mitigated by the requirement for each pilot or air crew member to advise if they are *not* fit for duty at any time; in other words they agree that they *are* fit for duty prior to the commencement of each duty period (United States Department of Transportation, 2011).

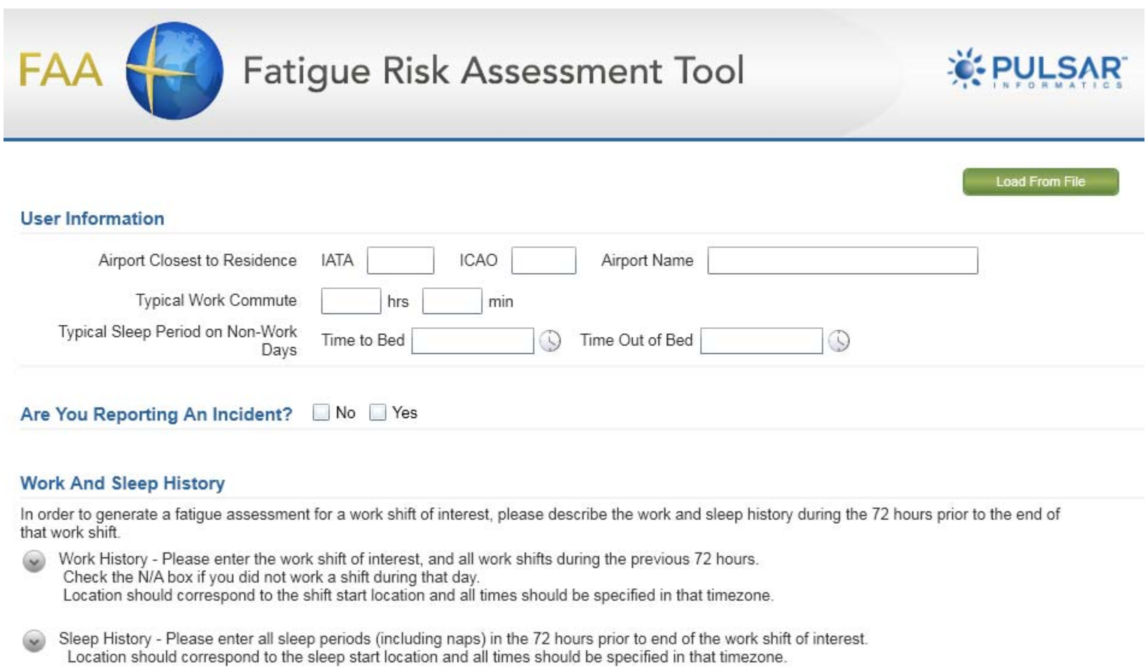

Figure 5. FAA (USA) Fatigue Risk Assessment Tool.

The FAA does not have a mandatory fatigue risk management system, but does have an advisory circular that provides details of an acceptable method for development of a fatigue risk management system (United States Department of Transportation, 2013). This document includes a sample fatigue report that suggests consideration of the topics listed in Table 7.

Table 7. FAA fatigue report considerations (United States Department of Transportation, 2013, pp. 19–23).

fatigue occurrence	rest (previous 72 hours)
commuter	sleep opportunities
reserve pilot assignment	circadian issues
operational issues	nutrition and hydration
augmented crew	personal factors
hotel/suitable accommodation	pre-duty activities

D. STUDY COMPARISON TABLE

The selected data for each study is tabulated in an Excel spreadsheet containing all of the items discussed. Each study is detailed across one row, with each of the elements listed in Table 8 having a separate column. The table is not reproduced here due to size and distribution limitations, but relevant excerpts are included in the analysis discussion in Chapter IV as appropriate.

Table 8. Data included in comparison table.

Description	
"who"	• country • group • author • platform
"what"	• purpose of study
"when"	• year • month/day • duration
"where"	• geographical location • tasking/role
Methods and Variables	
"how"	• method or approach • number of personnel onboard • number of participants • number of participants data analyzed • gender, age, job role, department, watch routine • other variables • equipment • software • data validation techniques • data analysis techniques
Conclusions	
	• conclusions • recommendations for further study • challenges • limitations • other notes
Crew Endurance Risk Factors	
	• sleep quantity • sleep quality • work schedule • work type • individual & lifestyle • organization & culture • environment

E. IDENTIFYING GAPS AND BEST PRACTICES

Gaps in knowledge can be identified in three ways using the comparison table. These are gaps in rows, gaps or queries in columns, and gaps provided by previous researchers. First, if it is not possible to fully complete a row due to the study not considering some areas of interest, a lack of information about a study, or because the study is not yet complete, then that may indicate a gap. (This does not invalidate a study; it could merely indicate that the scope of a study was limited to begin with or that the study had a different focus from the aspects considered here.) Similarly, if a factor from the risk management guides is not or only partially considered in a study, then that could also indicate a gap. Next, if a particular section of the table or even just one column has limited data, then this could mean that studies have not yet addressed this topic in detail. Finally, most researchers provide suggested areas for further research, which are very likely to be gaps suitable for further study, particularly if more than one report recommends a particular type of research be conducted in the future.

A list of best practices for workload studies can be assembled in a similar manner. If every study conducted uses a technique, then it is likely to be worthy of use. If a researcher comments on an aspect of their study in favorable terms, then it is likely that it will also be of benefit to others to utilize.

IV. ANALYSIS AND DISCUSSION

The 17 studies selected for comparison were conducted over the period from 1997 to 2013. Not all studies conducted by the RAN and USN were included in the final comparison. Some studies were precluded due to distribution limitations, while others were omitted due to their close similarity. If information from the same studies has been published in multiple locations then all documents are listed. Some studies from countries other than Australia and the United States (and from the civilian maritime sector) were included to allow generalizability of results to be discussed in greater depth. In addition, research into naval and maritime workload and fatigue is ongoing. Therefore, there were studies where results have not yet been published (or were not readily available) but enough is known about the study or the findings to warrant inclusion. A few of these studies were mentioned in the analysis, but were not included in the comparison table as the paucity of knowledge in some areas could lead to erroneous gap analysis. Table 9 provides a list of the studies selected for comparison in this thesis. The list is arranged by country and in chronological order from earliest to most recent. Platform type designation, as described in Chapter III, is also provided in the Country/Group column.

Table 9. Workload studies comparison list.

Country/Group	Year	Study Title
Australia RAN Submarine	2000	*The management of stress and fatigue amongst Royal Australian Navy submariners: A strategic, operational and financial imperative* (Chapman, 2001)
Australia RAN Large Warship	2011	*Crew endurance at sea: An analysis of sleep, work-hours and fatigue across a deployment period (HMAS WARRAMUNGA)* (Grech et al., 2014)
USA USCG	1997-1998	*Crew fatigue and performance on U.S. Coast Guard cutters* (Miller et al., 1998)

Country/Group	Year	Study Title
USA USN/NPS Aircraft Carrier	2002	*The effects of reversing sleep-wake cycles on sleep and fatigue on the crew of USS John C. Stennis* (Nguyen, 2002) and *Working the nightshift on the USS John C. Stennis: Implications for enhancing warfighter effectiveness* (Miller & Nguyen, 2003)
USA USN/NPS/other DOD Submarine	2003	*An analysis of the effectiveness of a new watchstanding schedule for U.S. submariners* (Osborn, 2004) *& Submarine watch schedules: Underway evaluation of rotating (contemporary) and compressed (alternative) schedules* (Duplessis, Miller, Crepeau, Osborn, & Dyche, 2007)
USA USN/NPS Small Warship	2004	*Effects of noise, temperature, humidity, motion and light on the sleep patterns of the crew of HSV-2 SWIFT* (Archibald, 2005)
USA USN/NPS Large Warship	2007	*A comparison between the Navy standard workweek and the actual work and rest patterns of U.S. Navy sailors* (Haynes, 2007)
USA USN/NPS Large Warship	2008	*A comparative analysis between the Navy standard workweek and the work/rest patterns of sailors aboard U.S. Navy cruisers* (Mason, 2009)
USA USN/NPS Large Warship	2009	*A comparative analysis between the Navy standard workweek and the actual work/rest patterns of sailors aboard US Navy frigates* (Green, 2009)
USA USN/NPS/Other Small Warship	2011	*Maritime platform sleep and performance study: Evaluating the SAFTE model for maritime workplace application* (Brown, 2012)
USA USN/NPS Large Warship	2012	*A comparison of sleep and performance of sailors on an operationally deployed US Navy warship* (Young, 2013)
USA USN/NPS/ONR Small Warship	2013	*Effects of sleep deprivation on US Navy watchstander performance onboard the independence class littoral combat ship (LCS-2)* (Davey, 2013)
Canada RCN/DRDC Submarine	2007	*An assessment of the CF submarine watch schedule variants for impact on modeled crew performance* (Paul et al., 2008) *& Alternative submarine watch schedules: Recommendations for a new CF watch schedule* (Paul, Hursh, & Miller, 2010)

Country/Group	Year	Study Title
Canada RCN/DRDC Large Warship	2011	*An assessment of some watch schedule variants used in CDN patrol frigates: OP Nanook 2011* (Paul et al., 2012)
Sweden Swedish National Road and Transport Research Institute (VTI)	2007	*Fatigue at sea* (Lutzhoft et al., 2007)
Finland Finnish Institute of Occupational Health/Others Civilian Ships	approx. 2007	*Effects of 6/6 and 4/8 watch systems on sleepiness among bridge officers* (Härmä, Partinen, Repo, Sorsa, & Siivonen, 2008)
Multinational (Sweden/UK/others) Project Horizon Civilian Ship (Simulators)	2011	*Project Horizon - a wake-up call: Research into the effects of sleepiness on the cognitive performance of maritime watchkeepers under different watch patterns, using ships' bridge, engine and liquid cargo handling simulators* (Project Horizon, 2012) and *Sleep, sleepiness and neurobehavioral performance while on watch in a simulated 4 hours on/8 hours off maritime watch system* (van Leeuwen et al., 2013)

A. COMPARISON OF STUDIES

Studies are compared under the headings described in Chapter III. Examples from outside the list of compared studies and from non-maritime industries (such as aviation) are used to highlight similarities, or demonstrate uniqueness, as appropriate.

1. Platform Type

The comparison considered one aircraft carrier study, three submarine studies, three small warship studies, seven large warship studies, and three civilian maritime industry studies (one completed in simulators). These are shown by type and country in Figure 6, with a study represented by a solid dot. RAN and USN studies that were not considered in the comparison but which are relevant to this thesis are listed in Table 10, and shown in Figure 6 as open circles.

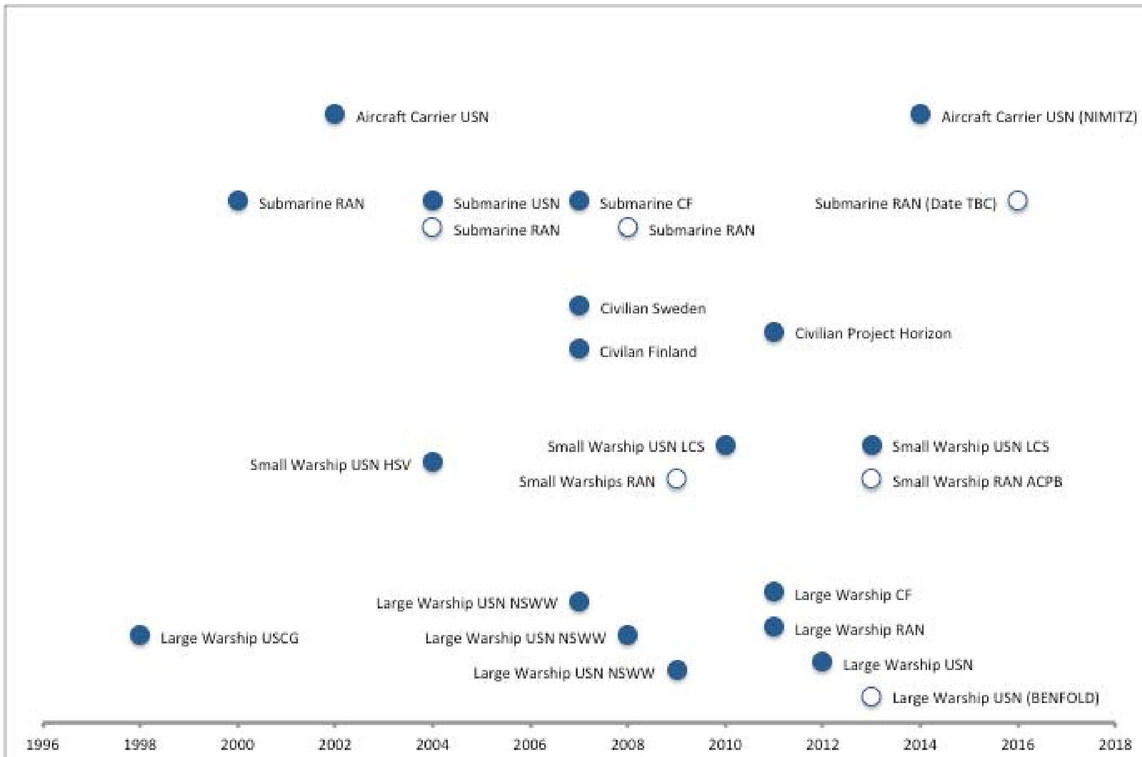

Figure 6. Studies by platform type and country/research group.

Table 10. Additional RAN and USN studies.

Country/Group	Year	Study Topic
Australia/RAN Submarine	2004	Two-section watch alternative trial HMAS RANKIN (Hussey, 2004)
Australia/RAN Submarine	2008	Three-section trial HMAS COLLINS 2008 (Buckley, full report not available)
Australia/RAN/DSTO Small Warships	2009	Several surveys (McLean, Grech, & Elischer, 2009, report not available for distribution)
Australia/RAN Small Warship	2013	Armidale Class Patrol Boat study
USA/USN/NPS Large Warship	2013	USS BENFOLD study
USA/USN/NPS Aircraft Carrier	2014	USS NIMITZ study
Australia/RAN Submarine	2016	Planned study

A large number of the studies recommend that further examination of the same topic be completed in different platform types, recognizing that each type of ship has factors that make it unique. Having said that, the author also considers that there might be some similarities across groups that have not been investigated. Submarine participants might have some similar results as those for aircraft carrier below-deck night workers (both have limited, but not necessarily zero, access to sunlight); or there could be similarities between surface ships and aircraft carriers for workers in similar roles such as bridge workers. These similarities were investigated, but not confirmed, in this review; each study was tailored toward its own aim and did not include enough general information to draw any conclusions regarding the groups. In the case of the surface ship bridge workers, we do not know if the study participants worked on the bridge, if topside aircraft carrier and surface ship workers have similar routines, or what the bunking arrangements for each platform are in terms of disturbances when off watch. Information limitations aside, certainly each platform type has unique factors, but the overall results are not markedly different in terms of measurements of amounts of sleep.[2]

All three civilian maritime industry studies, including one completed in simulators rather than underway, looked at a comparison between 6/6 (two-section six-hours on six-hours off) and 4/8 (three-section four-hours on eight-hours off) watch schedules, and came to similar conclusions. These happen to be the same conclusions reached by naval studies comparing these two watch systems: the 6/6 (two-section) watch system is more tiring than the 4/8 (three-section) (Härmä et al., 2008; Lutzhoft et al., 2007; Paul et al., 2012; Project Horizon, 2012). This finding may be coincidental, or it may be that there is only one real conclusion when the 6/6 and 4/8 watch systems are compared. Regardless of the reason, we cannot conclude that all simulator-based studies are able to predict underway fatigue accurately and/or that all civilian maritime studies are generalizable to navies. However, the result does imply that people

[2] A notable exception to this assertion is the amount of total sleep per day for topside aircraft carrier participants working the night shift and sleeping during the day. Nguyen (2002) found that topside workers working nights and sleeping days received much less sleep per day than below-deck night shift workers. This result will be discussed later in this chapter.

perform and react in similar ways in similar situations, and that watch system comparisons may be generalized across platform types (including carefully designed use of simulators) and industries, or at the very least that the results found in one area should be an assumed starting position for another platform or work area.

2. Study Descriptions

The 17 studies reviewed took place all over the world, in all types of situations, from simulators to warlike operations, and in calm to rough seas. Geographical and environmental data was not provided in all cases. The number of days of data collection was between three (Nguyen, 2002) and a prodigious 125 days (Grech et al., 2014), with one to three weeks being the most common duration. The number of platforms involved in each study was most commonly one; although six studies did have more than one vessel involved. The largest representation came from the Finnish study conducted by Härmä et al. (2008), with receipt of 92 separate survey responses from individuals, providing details of the respondents' most recent continuous seven days at sea.

In terms of what the studies aimed to achieve, they can be broadly separated into two groups. The first group appears to be more observational in nature, or focused on baseline data collection. Studies in this group measured sleep and either measured or predicted performance, with predictions based on the actual amount of sleep received, as collected via actigraphy or self-reported. Additional variables hypothesized to have an impact on sleep and/or performance, such as motion or light exposure, were included in some cases. The studies in this group are listed in Table 11.

Table 11. Studies measuring sleep and/or performance.

Group/Type	Year	Study Focus
RAN Submarine	2000	Stress and fatigue amongst Royal Australian Navy submariners (Chapman, 2001)
RAN Large Warship	2011	Analysis of sleep, work-hours and fatigue (RAN frigate) (Grech et al., 2014)
USA USCG	1997-1998	Crew fatigue and performance (USCG cutters) (Miller et al., 1998)
USN/NPS Aircraft Carrier	2002	Effect of reversing sleep-wake cycles on sleep and fatigue (Miller & Nguyen, 2003; Nguyen, 2002)
USN/NPS Small Warship	2004	Effect of noise, temperature, humidity, motion and light on sleep patterns (Archibald, 2005)
USN/NPS Large Warship	2007	Compare NSWW to actual work/rest patterns (Haynes, 2007)
USN/NPS Large Warship	2008	Compare NSWW to actual work/rest patterns (USN cruisers) (Mason, 2009)
USN/NPS Large Warship	2009	Compare NSWW to actual work/rest patterns (USN frigate) (Green, 2009)
USN/NPS/Other Small Warship	2011	Evaluate SAFTE model for maritime workplace application (Brown, 2012)
USN/NPS/ONR Small Warship	2013	Effects of sleep deprivation on performance (Davey, 2013)
RCN/DRDC Submarine	2007	Assessment of CF submarine watch schedule variants (Paul et al., 2010; Paul et al., 2008)
RCN/DRDC Large Warship	2011	Assessment of CF frigate watch schedule variants (Paul et al., 2012)

The second group of studies extends the research by adding a deliberate comparison of watch schedules and/or comprehensive work/sleep information (for example, not just how much people are sleeping, but also how much they are working and what type of work they are doing). The latter group, of more formally designed evaluative studies, is focused on making a choice between two or more options, or increasing knowledge beyond a description of actual sleep and/or performance, in order to look for relationships that explain sleep quantity/quality and the follow-on effect on performance. As shown in Table 12, every study in this second group looked at a comparison between alternative watch schedules, although this is not the only type of evaluation that could occur.

Table 12. Studies comparing watch schedules.

Group/Type	Year	Study Focus
USN/NPS/other DOD Submarine	2003	Evaluate rotating (contemporary) and compressed (alternative) schedules (Duplessis et al., 2007; Osborn, 2004)
USN/NPS Large Warship	2012	Compare 6/6 and 3/9 watch routines (Young, 2013)
Swedish National Road and Transport Research Institute (VTI)	2007	Compare 6/6 and 4/8 watch routines (Lutzhoft et al., 2007)
Finnish Institute of Occupational Health/Others Civilian Ships	approx 2007	Compare 6/6 and 4/8 watch routines for bridge officers (Härmä et al., 2008)
Multinational (Sweden/UK/others) Project Horizon Civilian Ship (Simulators)	2011	Effects of sleepiness on performance under different watch patterns, predominately 6/6 and 4/8 (Project Horizon, 2012; van Leeuwen et al., 2013)

There is overlap between the two groups; that is, some studies in the more descriptive group provided comparisons between different departments onboard, or groups of people. It would appear that earlier studies conducted by a country/group or with regard to a particular platform type, collected data for baseline purposes or to create or validate modeling tools, while later studies investigated alternative watch schedules to identify areas for potential improvement. More of these evaluative studies are expected to appear in the literature in the future. The 2013 USS BENFOLD and 2014 USS NIMITZ studies conducted by the USN and NPS (N. Shattuck, personal communication, November 26, 2014) fall into this category. All of the civilian studies that were reviewed are in the second group, albeit with quite a lot of overlap. This is especially true for Project Horizon where the data collected has been used to produce a fatigue management modeling tool called MARTHA (Project Horizon, 2012).

3. Study Variables and Methods

The studies reviewed in this thesis used similar variables and methods of analysis. Specific aspects of interest are reported here; noting that this is by no means an inclusive listing.

a. Age and Gender

Age and gender were not recorded or analyzed in all studies; none of the results suggest that either is vitally important for naval workload studies which may be why they were not considered in detail. However, age and gender details are very quick and easy to collect and should be noted to ensure completeness of data sets and to allow later analysis if desired. Gender was not found to be significant with respect to sleep or performance in the selected studies; however, this finding could be due to the small numbers of female compared to male participants in the male-dominated maritime and military environments. Sawyer (2004), in a NPS thesis involving personnel onboard an aircraft carrier, recommended that further studies attempt to achieve a more equitable gender balance. This inequity is not surprising since the sample of females is representative of the population. Militaries are not usually gender balanced. Age was found to be of interest in some cases, but results were not consistent across studies, and age-related analysis was not always completed. Implying generalization of results with respect to age is not viable without further research. Age may be a proxy for a range of other factors, such as rank, time in service and time at sea, but this finding is expected only if there is direct and consistent correlation with age. This relationship would be less true for the RAN than the USN, since the RAN has more latitude with respect to age on entry, and does not have an “up-or-out” policy. Where specified, average age, and age range, of participants for civilian studies was higher than seen in military studies, as is shown in Table 13.

Table 13. Average age of participants in specified studies (standard deviation in parentheses where available).

Platform	Average (Mean) Age
Aircraft Carrier (Nguyen, 2002)	25.6 years
Large Warship (Mason, 2009)	29 (7.2) years
Large Warship (Miller et al., 1998)	29 years, range 22-40 years
Large warship (Paul et al., 2012)	32.9 (7.7) years, range 21-48 years
Small Warship (Brown, 2012)	35.8 (5.92) years
Submarine (Paul et al., 2008)	38.6 (7.6) years, range 26-54 years
Civilian ship (Lutzhoft et al., 2007)	41.5 (9.9) years

b. Participants

It is uncommon for substantial monetary or other incentives to be provided for participation in sleep studies such as those compared in this thesis. This is particularly true in naval studies; most participation relies upon the investigator to explain to potential volunteers the importance of their participation to the overall results, thus eliciting their cooperation. Study participation takes time out of already busy days for the individual volunteering for these studies. However, since participants within an industry are generally aware of the increased safety risk and performance degradation caused by fatigue, they are often pleased to be able to assist in furthering research that is expected to lead to improved conditions and/or outcomes.

It is important to ensure that fatigue risk management studies do not induce additional and unnecessary fatigue. In the recommendations for implementation of a fatigue risk management system (FRMS), the FAA specifies the use of a “measurement methodology that will be sufficient to demonstrate that operations under the FRMS do not induce additional fatigue relative to operations under the prevailing prescriptive rules” (United States Department of Transportation, 2013). Any study of this nature must also expect that if conditions change or tasking varies, volunteers may withdraw. The studies reviewed here all relied upon volunteers, and usually did not involve a large proportion of the vessel’s population. Deliberate efforts were made to avoid self-selection bias and achieve a representative mix of personnel from across departments, ranks, gender, etc., but initial and/or final numbers of participants were not always large enough for valid statistical comparisons to be made. For example, the large warship study by Green (2009) experienced a 50 percent dropout rate in participation, calling into question whether the final dataset was representative of the population. This problem may be able to be avoided if a much larger sample is recruited for a study, although resources may not support this decision.

c. Department and Watch Schedule

Studies that compared departments found statistically significant differences in sleep quantity and quality, probably related to different watch routines employed by different departments. This finding was also of interest due to the possibility that the culture of individual departments varies and could account for some of the differences observed (Haynes, 2007; Mason, 2009). One submarine study suggested that minimum sleep requirements should be calculated by department to allow for the different risk factors associated with different tasks in each department (Chapman, 2001). Departmental data did not have large sample sizes (ranging from 2 to 15 participants); departmental divisions were not always the same. Some departments might be combined for a study, but not in the same way each time, and departments varied by ship type and country. Watch schedules also varied by department, so departments were difficult to compare. Consequently, generalizations were not available from this review. In the future, departmental data should be collected for all studies, noting that additional questions should be included to ensure that a person is actually working within their assigned department (for example, misclassification can occur for personnel working as a food service attendant (Green, 2009)), and specifying what the exact watch routine is for the work area.

Officers tended to be grouped as their own department, rather than being included with a work area. It does make sense to analyze data by rank or rank grouping, and senior personnel (both officers and enlisted) were found to receive less sleep than junior personnel in a number of studies (Archibald, 2005; Green, 2009; Mason, 2009). Inclusion of officers within work areas may also be relevant, especially if the cultural aspects of a department are being investigated.

Watch schedule information is difficult to collect and record accurately, particularly when terms such as “Straight 8s”, “Five and Dime” or “Dogged 4s” may not be commonly understood by all participants and readers. In addition to actual time on watch, a watch schedule needs to specify when personnel have an opportunity to sleep. It is important to know if people have other duties when they are off watch and/or are able to sleep during the day. If there is a complex system, a change in watch schedule, or a routine watch amendment (such as on some civilian vessels where the Master stands one watch per day, allowing the usual watch standers an “unscheduled” break from watch), this deviation must be noted. Diagrams may be effective in reporting these types of watch schedules, such as the four-section weekly rotating watch system shown in Figure 7.

Watch before shifting

Time	Work	Sleep			
0100	1		2	3	
0200	1		2	3	4
0300	1			3	4
0400	2			3	4
0500	2			3	4
0600	2				4
0700	3				4
0800	3				4
0900	3				
1000	4				
1100	4				
1200	4				
1300	1				
1400	1				
1500	1				
1600	2				
1700	2	1			
1800	2	1			
1900	3	1			
2000	3	1	2		
2100	3	1	2		
2200	4	1	2		
2300	4	1	2	3	
2400	4		2	3	

Saturday before shifting

Time	Work	Sleep			
0100	1		2	3	
0200	1		2	3	4
0300	1			3	4
0400	2			3	4
0500	2			3	4
0600	2				4
0700	3				4
0800	3				4
0900	3				
1000	4				
1100	4				
1200	4				
1300	1				
1400	1				
1500	1				
1600	1				
1700	2				
1800	2				
1900	2				
2000	2	1			
2100	3	1			
2200	3	1			
2300	3	1	2		
2400	3	1	2		

Sunday shifting of watch

Time	Work	Sleep			
0100	4	1	2		
0200	4	1	2	3	
0300	4		2	3	
0400	1		2	3	
0500	1		2	3	
0600	1			3	
0700	2			3	
0800	2			3	
0900	2				
1000	3				
1100	3				
1200	3				
1300	4				
1400	4				
1500	4				
1600	1				
1700	1				4
1800	1				4
1900	2				4
2000	2	1			4
2100	2	1			4
2200	3	1			4
2300	3	1	2		4
2400	3	1	2		

Figure 7. An example of a 3/9 watch shifted forward on Sunday with no consecutive watches (from Roberts, 2012, p. 68).

d. Other Variables

Other variables were noted when they were the focus of a study, such as the light, motion, and noise data collected by Archibald (2005). Caffeine and tobacco/nicotine use information was sometimes collected but seldom reported, usually because it was not the major focus of the study. In the civilian ship simulator study, it was noted that participants were restricted to no more than four cups of coffee per day (Project Horizon, 2012).

e. Data Analysis and Validation

Variables measured in these studies ran the gamut: they were objective and subjective, observed and self-reported, with continuous and non-continuous data. Parametric statistical analysis was used for normally distributed continuous data; non-parametric analysis was used for analysis of Likert-like sleepiness or mood scale results. Some studies utilized multivariate regression analysis, with a mixed effects model most often specified; sleep quantity was the usual dependent variable.

Techniques used to validate data were not widely reported, although actigraphy downloads were commonly checked against sleep-wake logs, and then individually “cleaned” to ensure the highest possible level of accuracy. Onboard actigraphic data download and cross-checking, perhaps daily, was suggested as a means of increasing accuracy, to avoid the potential for difficulties encountered when attempting to query results with individuals after the study is complete.

Different sleepiness scales were used across the studies reviewed, with each research group tending to stay with one particular scale; for example, the USN/NPS studies use the Epworth Sleepiness Scale while Scandinavian-based groups tend to use the locally-developed Karolinska Sleepiness Scale. Tools for deriving predicted effectiveness also varied, again with a local or language of origin effect being visible. FAST was used in all of the RAN, USN and CF studies but in none of the civilian studies. The ability to input intended sleep times into

FAST and calculate predicted effectiveness means that potential risks can be avoided. This idea was replicated across the studies reviewed, with the RAN NMD and heat map diagrams playing a similar role in risk avoidance (Grech et al., 2014). Data from Project Horizon was used to produce a civilian ship-oriented tool called MARTHA which can be used for alertness or performance prediction (Project Horizon, 2012). IMPRINT was not used in any of the studies reviewed, but may be appropriate for later studies. It would also be interesting to compare predictions of FAST, NMD, and MARTHA for the same data sets to determine how their results compare with one another.

f. Baseline Performance

Establishing an individual's baseline performance when they are not fatigued is seldom possible, particularly in military field studies. While there is no guarantee that having baseline data would substantially change the final results, it would be useful to verify that idea. Many of the studies reviewed suggested that further research should include baseline data collection. FAST assumes a three-day preconditioning period to reduce errors induced by an incorrect historical baseline. The default value for this preconditioning period is eight hours of quality sleep per day. Some studies that utilized FAST reported that the first three days of data were not used in the analysis, but it was not clear in every case that preconditioning had occurred. In the real world of naval operations, three consecutive days of eight hours of quality sleep each night is rare, assuming that it occurs will have a significant impact on the study results. Another assumption in FAST is that the individual is awake for one hour before and after watch. While this may be true on average, it would be prudent to validate this assumption using actual data or surveys. The experience of the author suggests that the length of time before and after watch differs depending on the specific watch and the number of days a vessel has been at sea.

Even with accurate baseline data, the location, tasking or previous activity of participants could have an impact on results. Paul (2012) specifically notes that participants in the CF submarine study commenced the study in a fully rested state due to a one-month defect rectification period, but that participants in the CF large warship study had been at sea for three weeks prior to the helicopter embarkation of the study team. Green (2009) suggests that workload studies should take place in warlike conditions; and Miller et al. (1998) noted that the USCG vessels studied were not subject to a high operational tempo during the data collection period.

Performance improvement due to the learning effect was raised as an issue by Brown (2012), with respect to the switching test used as a performance indicator. She also noted that the PVT does not have a learning effect, thereby avoiding this issue. Some suggestions for remedying this problem include choosing a test that does not have a learning effect, baselining data, or administering a test for seven days prior to the study as suggested by Davey (2013) with regard to the switching test.

A more complicated problem was reported by Paul et al. (2008, p. 3), demonstrating that it is difficult to predict with certainty what the second and third order effects of an action may be, and suggesting that researchers perhaps should not share too much data with participants during the study:

> Probably the most compelling reason that the PVT data are of questionable utility is that a significant number of subjects were competing for the fast reaction time of the day, every day and this resulted in a shift in the area of the speed-accuracy trade-off curve at which these subjects were choosing to perform. Essentially, for good reaction time data, the subjects should respond as quickly as they can without making mistakes in which case the tolerable error rate is about 2%. However, in their quest for speed, accuracy was sacrificed and many of the subjects had as many errors as correct responses making their data unusable.

g. Readability

An important component of any research is the ability to communicate the method and results in a manner that makes it easy for the intended audience to understand, but with enough detail for a more informed reader to be able to: a) make an assessment regarding generalization of results, b) use the method to replicate results, and c) provide sufficient background detail to enable collection of comparable data. Several of the studies reviewed here have characteristics that make them reader friendly. Two are provided here as examples that ensure maximum attention and understanding by a variety of readers. Firstly, use of a concise summary paragraph regarding methodology employed, with sufficient detail regarding statistical analysis to enable replication, similar to that provided in the CF studies, for example:

> Such ‘interval data’ is not normally distributed and is therefore analysed via non-parametric statistics. The Kruskal-Wallis analysis was used to assess group differences, and the Friedman Analysis of Variance (ANOVA) to test repeated measures across days. The Wilcoxon test was also used to assess matched pairs of cells. … A split-plot ANOVA with 3 between factors (i.e. 3 different watch system variants) and 12 repeated measures (i.e. 12 days at sea) was used for analysis of the VAS data. (Paul et al., 2008, pp. 4–5)

The second example is to provide a summary of measures used and how they were collected, as shown in Figure 8 from the USCG cutter study.

3.12 Summary of Measures

Work Demand
- Work-rest schedule (Daily Log)
- Cumulative sleep debt (Daily Log, Questionnaire)
- Number of sleep periods per day (Daily Log)
- Task descriptions of metabolic demand (Daily Log, Questionnaire)
- Description of ship motion (ships' logs)
- Perception of motion discomfort (Daily Log)
- Engine room temperatures (ships' logs)

Effort
- Perceptions of physical and mental workload (Daily Log)
- General motivation (Questionnaire)

Performance (overt)
- Operational task incidents (observation)
- Computer-based performance tests (simultaneity, vigilance, pattern recognition, code substitution, tapping speed)

Fatigue (covert)
- **Circadian Effects**
  - Body temperature rhythm (Thermoscan)
  - Task performance rhythms (simultaneity, vigilance, pattern recognition, code substitution, tapping speed)
  - Sleepiness rhythm (SVASS)
- **Acute**
  - Stanford Sleepiness Scale (Daily Log)
- **Cumulative**
  - Perception of sleepiness (SVASS)
  - Combined sleep debt and recovery sleep pattern (Daily Log, Questionnaire)
  - General level of arousal (FIT 2000)
  - Computer-based performance task data (simultaneity, vigilance, pattern recognition, code substitution, tapping speed)

Figure 8. Summary of measures (from Miller et al., 1998, p. 16).

4. Study Conclusions

Not every study that was reviewed reached a definitive conclusion or proved an expected result, usually as a result of missing data or limitations on the sample size (e.g., Davey, 2013; Green, 2009; Lutzhoft et al., 2007). Many studies provided information or suggestions appropriate to list as a gap or best practice; these items are discussed in a later section. This section details the average amount of sleep obtained by study participants, and uses this information to illustrate how conclusions and results can be misinterpreted.

When available, the average amount of sleep per 24-hour period for all participants in each study is provided in Table 14, together with the NSWW daily sleep allowance, and the amounts of daily sleep recommended by experts and obtained by Americans as reported by Roberts (2012).

Table 14. Average amount of sleep per 24-hour period for specified studies/groups (highest to lowest).

Platform	Average Sleep per 24-hours
NSWW allowance	8 hours
Sleep experts recommend	7.5 to 8 hours
Large Warship (Haynes, 2007)	7.3 hours
Large Warship (Grech et al., 2014)	7.2 hours
Large Warship (2) (Mason, 2009)	7.15 hours
Submarine (Osborn, 2004)	7.1 hours
Large Warship (1) (Mason, 2009)	6.93 hours
American national average	6.9 hours
Large Warship (Green, 2009)	6.72 hours
Large Warship (Young, 2013)	6.3 hours
Aircraft Carrier (Nguyen, 2002)	6.17 hours

While none of the values reaches the eight hours planned for in the NSWW, the range of 6 to 7.5 hours may seem fairly reasonable, more than half of the platforms have participants getting more average sleep per day than the American average provided by Roberts (2012). This conclusion would, however, be erroneous since overall averages can be somewhat misleading and should not be taken as applicable to all people or situations. Consider how the following additional information that is available in Table 15 expands our understanding of the results of the studies.

Table 15. Average amount of sleep per 24-hour period for specified studies, by watch schedule or department (note that "Operations" department is not the same in the RAN and USN).

Platform & Average Sleep	**Average Sleep per 24-hours by watch or department**
NSWW allowance 8 hours	8 hours
Sleep experts 7.5 to 8 hours	7.5 to 8 hours
Large Warship (Grech et al., 2014) 7.2 hours	Operations and engineering received less sleep at sea than supply and electrical departments during all study stages
Large Warship (combined) (Mason, 2009)	Operations 7.32 hours Supply 6.93 hours Combat 6.69 hours Engineering 6.33 hours
Submarine (Osborn, 2004)	7.1 hours (existing schedule) 6.6 hours (modified experimental) 6.3 hours (experimental)
American national average	6.9 hours
Large Warship (Green, 2009) 6.72 hours	Supply 8.29 hours Combat 6.62 hours Nav/Admin 6.46 hours Operations 6.36 hours Engineering 5.82 hours
Large Warship (Young, 2013) 6.3 hours	3/9 watch 6.53 hours 6/6 watch 5.52 hours
Aircraft Carrier (Nguyen, 2002) 6.17 hours	Below deck workers 7.37 hours Above deck workers 4.74 hours
Project Horizon (Project Horizon, 2012)	4/8 team two 7 hours 4/8 team one 6 hours 6/6 watch "markedly less," in two parts

The two tables demonstrate the dangers associated with the generalization of results, and incomplete understanding of the background of a study. The sleep of participants in Green's large warship study seemed marginally acceptable in Table 14, with an average sleep amount of 6.72 hours per day. However, when we look at Table 15 we have much more information to draw upon. First, we need to examine the supply department more closely to find out what they are doing that allows them to get more sleep (8.29 hours on average). Perhaps more importantly we need to find out why the engineering department is only getting an average of 5.82 hours sleep per day, and then rectify that issue.

Departments are combined in different ways in different studies; the results for a department on one ship are not the same as for the same department on another ship and may not be the same for the same department at a different time. Even more concerning is that we know is that it is not just sleep quantity we need to worry about, but equally important is sleep quality. There is no allowance made in these studies for factors affecting sleep quality or for split sleep (more than one sleep episode per day). Unless sleep is measured actigraphically and adjusted for sleep quality, the actual sleep may be even less than initially reported. The situation involving split sleep can be particularly misleading since the sum of two sleep episodes is not considered to be somewhat less than rather than equivalent to their total (Lutzhoft et al., 2007).

5. Crew Endurance Risk Factors

The crew endurance risk factors used in this thesis were described in detail in Chapter III, the broad groups are: sleep quantity, sleep quality, work schedule, work type, individual and lifestyle, organization and culture, and environment. All of the studies considered looked at sleep quantity in detail, and most considered sleep quality in some manner, but the other crew endurance risk factors investigated were not considered by all the studies reviewed, and the results of the analysis of most of these risk factors were not reported. Work schedules were commonly considered, but type of work was not. For example, a sailor could have been doing manually demanding physical labor for 10 hours or monitoring a visual display. Both are fatiguing but they cannot be assumed to be identical in their end result. Individual and lifestyle factor data collected only included caffeine, tobacco, and alcohol use, and this data was not reported in any detail. Organizational and cultural factors were mentioned when there was a potential change management issue; for example, announcements over the intercom and loudspeaker may have been postponed until after 1000 to allow the night watch to sleep late; meeting times were moved to earlier in the day to better accommodate sleep schedules. Environmental factors were mentioned only when they were a topic of interest in the study, such as when motion, noise or temperature variables were investigated.

Even though reporting of results regarding crew endurance risk factors was limited, the results that were provided suggest that some generalizability of results is appropriate even though studies are not always completely generalizable. People are people, many of the crew endurance risk factors used here apply equally to people in any platform. For example, a 6/6 two-section system results in decreased sleep quantity and quality (due to fragmented sleep) when compared to other two watch systems; the platform makes little difference.

All of the crew endurance risk factors investigated here are likely to be of interest in future research, particularly sleep quality and work schedule which are influenced by organization and culture. For sleep quality, issues of sleep fragmentation, whether related to watch schedule or not, and organizational changes that may allow sleep to be of better quality are topics of interest. Having individual rooms and padding doors, as mentioned by Watt (2009) as a means of zealously guarding aircrew sleep quality may not be possible onboard ships, but there are likely to be other potential improvements that could improve the quality of the sleeping compartments. The bio-mathematical performance effectiveness models used have had some validation in the maritime environment (Brown, 2012) but further investigation as to their accuracy in alternative conditions is appropriate. The extension of research into lighting, noise and similar factors may result in relatively simple solutions for increased sleep quality and thereby enhance performance.

B. THE EVOLUTION OF SLEEP STUDIES INTO CREW ENDURANCE MANAGEMENT PROGRAMS

The study features discussed in the previous sections demonstrate that more recent studies investigate sleep quantity, and then use it as a means of calculating cognitive effectiveness, and/or also measure performance. The study of crew endurance is maturing, moving from a focus on what is occurring; to measuring or predicting performance. The next major milestones are to determine how to achieve the highest possible cognitive effectiveness, decrease risk, increase safety, and make sleep hygiene an individual responsibility enabled by their organization. Tools such as the USCG CEMS (Comperatore et al., 2005) and Roberts (2012) advice to command on watch schedules are examples of the information now available to assist personnel. The maturing of this process can be compared with the ADDIE model for training development, the OODA loop, or many other scientific method or research methodology theories. For example, similar approaches are found in the *Technical Cooperation Program's Guide for Understanding and Implementing Defense Experimentation* (GUIDEx) (Bowley et al., 2006) or Bardach's (2011) eightfold path to effective problem solving. A simple problem-solving model for any problem/issue/question is a modified ADDIE model[3] consisting of four steps :

1. Assess the situation.
2. Develop a potential solution/s.
3. Implement a solution.
4. Evaluate the results (and then return to the beginning).

The study of crew endurance has moved through this cycle multiple times, solving many problems and coming up with many more questions. If each group represented in the studies reviewed is considered over time, the evolution from assess (i.e. measure sleep) to develop (come up with an alternative watch

[3] The ADDIE model is considered sourceless (Molenda, 2003), therefore this modification shares that trait.

schedule) and implement/evaluate (observe and come up with another question to assess) is apparent.

The USCG as an organization has the most mature implementation of crew endurance for the groups studied, with assessment beginning in the 1980s leading to the development of the first version of CEMS in 2003. Updates (CEMS v2), and translation to other similar industry areas (towing vessels) were accomplished in 2005. For USCG ships, anecdotal evidence indicates that sleep hygiene is understood by all individuals, and fatigue management is supported at the organizational level as "business as usual".

The USN may have done more at-sea studies than any other navy, and the USN/NPS collaboration is a very useful way to both produce reports and to educate personnel from across the world regarding fatigue. Davey (2013, p. 9) states that "Efforts to correct the issue of fatigue within the Navy were at a standstill until the early 2000s when Dr. Nita Shattuck, in conjunction with NPS thesis students, began to renegotiate the Navy's "sleep when you're dead" culture." It is apparent that fatigue management has been important to many personnel in the USN over a long period of time but challenges to organizational level communication endure despite the best efforts of numerous personnel. A 3 November 2014 link shared on the NPS Facebook page mentioned that research by NPS Associate Professor Nita Shattuck into non-standard watch schedules was supported by the Commander, Submarine Forces, Vice Admiral Michael Connor (Naval Postgraduate School, 2014). It quickly garnered the comment shown in Figure 9:[4]

[4] While not necessary given that it is in the public domain, the author did attempt to contact the commenter, listed on Facebook as a retired USN Captain, for permission to reproduce this comment. The then-Commander Cutler Dawson mentioned is assumed to be VADM Dawson (retired), who served in the USN for 34 years and in 2014 was President and CEO of the Navy Federal Credit Union (Roberts, 2014).

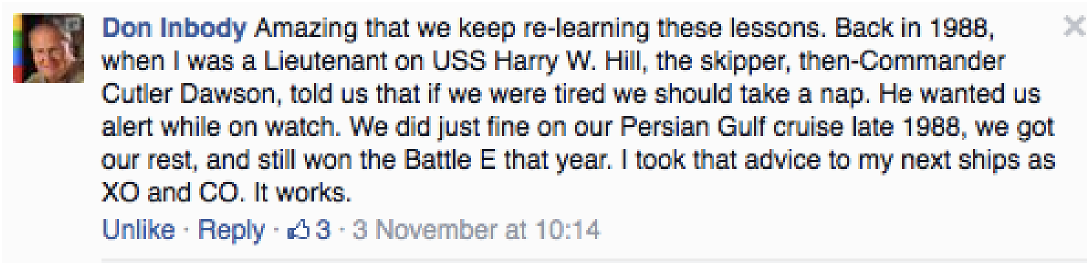

Figure 9. NPS Facebook post comment (from Inbody, 2014).

Gaining traction to achieve organizational awareness and implementation of a FRMS is an ongoing process, especially in an organization as large as the U.S. DOD. The movement from descriptive studies assessing sleep in submarines in 2001 (Blassingame, 2001; Gamboa, 2002), to operational studies all over the world on multiple platforms of all types, continuing through to the current evaluative testing of alternative watch schedules, and planned production of further fatigue related educational programs are indications that the USN is making a great deal of headway.

The RAN has also been formally investigating sleep onboard platforms since at least 2001 (Chapman, 2001), albeit at a much slower pace than the much larger USN. In 2011, the RAN completed what may be the largest underway actigraphic study in the world (Grech et al., 2014). The project built on the experience of the USN by consulting with NPS faculty in the development and implementation of the project. High level support via the New Generation Navy (NGN) program (NGN Project 13—*People-focused work practices*), the dedication of an entire chapter to fatigue in the 2014 version of the Navy Safety Management System publication, a domestic safety award for the efforts aboard HMAS WARRAMUNGA during and after the 2011 study, and the implementation of the NMD at sea are all contributing to a broader understanding of the importance of fatigue. The next steps could be to confirm applicability of previous results to other platforms through detailed analysis of data collected aboard a RAN patrol boat in 2013, another example of fruitful collaboration with the NPS. In addition, ongoing submarine studies are planned for 2015–2016; and work could continue toward a standalone FRMS or crew endurance publication for the RAN. Ultimately personnel in ships should be able to operate under the same type of crew rest principles that naval aircrew already adhere to, monitored regularly using tools such as the NMD to ensure potential problems with schedules are identified and rectified before they occur.

Watch schedule research in the RCN was furthered due to an unfortunate circumstance involving a fatal fire onboard a RCN submarine in 2005. The board of inquiry modeled the fatigue of personnel onboard which led to field studies that showed even higher levels of fatigue on RCN submarines than had been predicted by the model. The International Submarine Watch Schedule Symposium hosted by Canada in 2009 (attended by the RAN, USN, and others) discussed watch schedule alternatives. As a result, the CF implemented a revised two-section watch system across their submarine fleet, improving modeled cognitive effectiveness by 30 percent compared to the original system (M. Paul, personal communication, October 24, 2014). The CF have also been proactive with work on the extension of submarine watch schedule research to surface vessels. In 2010, they published a CF fatigue risk management guide, which although aviation-driven is still applicable to the maritime environment (Cheung et al., 2010).

Civilian shipping studies were outside the scope of this thesis and have not been comprehensively reviewed here. However, a few comments about civilian shipping studies may be appropriate. The 2010 ‘Manila Amendments’ to International Maritime Organization (IMO) Standards of Training, Certification and Watchkeeping (STCW) allow for a 91-hour workweek, or up to 98 hours per week for two weeks in exceptional circumstances (Project Horizon, 2012). This change has no doubt predicated some research. An alarming statistic generated by Project Horizon found that 40 percent of participants fell asleep on watch (Project Horizon, 2012). This finding should prompt even more action about the issue of fatigue and work schedules in the civilian shipping industry.

The results of civilian shipping studies may not be directly applicable to the naval environment: it is unlikely that 40 percent of navy watch standers are falling asleep, especially given the larger numbers of personnel on watch at any time and the fact that they may have additional tasking compared to their civilian counterparts. However, it is highly likely that some navy watch standers are falling asleep on watch and that many are performing at sub-optimal levels, every step needs to be taken to prevent this from occurring.

C. GAPS AND BEST PRACTICES

The descriptive and evaluative studies reviewed in this thesis had many differences and a number of commonalities. They detailed what was successful during their study and what was not, either specifically or generally. The specific advice and the more intangible messages gleaned from looking carefully at the whole group of studies can be combined to compile a list of practices that should be encouraged, or avoided. In addition, suggestions can be made as to what additional research might be needed to increase the body of knowledge regarding fatigue and crew endurance. Gaps identified from the comparison of the selected studies are grouped by subject and provided in Table 16. Discussion of the issue is included *in situ,* if not previously covered.

Table 16. Gaps identified from comparison of studies.

Gap	Discussion
Environment	Extreme or unusual conditions (such as bad weather) have not been extensively studied.
Sleep quality	Further investigation of sleep quality is necessary; this is a large topic, and some study has already been done or is currently underway.
Individual factors	Crew endurance risk factors such as the effect of pharmacological interventions and other influences such as caffeine, tobacco, alcohol, family stress, and so on have not been considered in great detail.
Individual tolerance factors	Individuals are affected in different ways by their environment; for example, some get seasick, while others may not. There are individual tolerances for sleep deprivation. Little research on this topic was evident, when it did occur it was due to exceptionally

Gap	Discussion
	large individual differences (Paul et al., 2008, p. 6), which may be occurring in other studies without being recognized, thereby biasing results.
Concurrent studies	Most studies were done in isolation. Conducting the same study at the same time on a variety of platforms would allow for cross-comparison by holding many extraneous variables constant.
Manpower policies	There is a potential gap between what studies are investigating and what assumptions manpower requirements organisations use. Increased communication between researchers and manpower personnel may be appropriate. A number of studies have found that the 3/9 watch system is preferred over others, such as the 4/8 or 5/10. This is hardly surprising given that this is comparing a four-section watch system (the 3/9) to a three-section watch system (4/8 or 5/10); logically the four-section watch system is likely to be preferred. For surface ships, current manpower requirements are largely based on a three-section watch system. If there are enough people for a four-section watch system to be used then there could be an issue with other work not being completed, or an oversupply of personnel to an area. There is little value in recommending a four-section watch system to improve crew endurance if manpower requirements policies dictate a three-section watch system basis for allocation of personnel.
Promulgation of study evaluations	There is evidence of improvements resulting from several of these studies. For example, HMAS WARRAMUNGA implemented substantial changes to routines (Turner, 2012), and ongoing USS NIMITZ work predicts performance improvement due to changing of watch schedules. Notwithstanding, little evidence is available to show that evaluation is occurring and that FRMS are working. Even less information is available to demonstrate organizational level learning. The Watt study (2009) demonstrated successful scheduling, but the methodology does not appear to be used widely. The 2010 CF submarine study recommended an 8-4-4-8 routine rather than the then current 6-6-6-6 routine, with a predicted 23 percent overall increase in mean cognitive effectiveness (Paul et al., 2010), but results of the recommended at-sea trial are not publically available. The information was obtained via personal correspondence. Changes to routines, meal times, etc., are not always simple and may even be considered to be “more trouble than they are worth,” but if evidence from another study showed a 23 percent increase in cognitive effectiveness, it is difficult to argue about the value of implementing the plan in similar situations.

Gap	Discussion
Evaluation of software and tools	Ongoing evaluation of fatigue management tools is needed. FAST has been validated with actual versus predicted effectiveness data (Brown, 2012). Some issues related to ship movement are known, but can be accommodated. FAST, like most tools, is evolving; the versions used in the studies compared in this thesis ranged from beta through to version 3. Ongoing evaluation would be useful to ensure validity.
Comparison of software and tools	Comparison of different fatigue management tools, such as FAST, NMD and MARTHA, by calculating effectiveness predictions for a single data set, would increase knowledge about the generalizability of conclusions across research efforts. Similarly, the use of IMPRINT for further comparison and validation could be investigated.
Education	It is time to take further steps in the education process. CEMS, and the U.S. Navy Safety Center videos are a good start, but somewhat concerning is anecdotal evidence that fatigue prevention measures do not survive past a particular command team. Eventually those commanders will have higher level command jobs, and education will occur. Having education more rapidly available, at the organizational level, would be beneficial.

The use of best practices is also important, not just during studies but also afterward when results are promulgated and others are able to benefit from a completed study by using relevant results in their own unit. A personal example is included here to demonstrate a situation where widespread promulgation of best practices may have improved crew endurance. Many studies, completed over many years, have found that rotating watches forward rather than backward, or not rotating the watch at sea at all, is preferable. However, the authors experience and recent watch comparison studies still refer to ships having a three-section, dogged, backwardly rotating watch system, such as the one shown in Figure 10, which is in common use in the RAN.

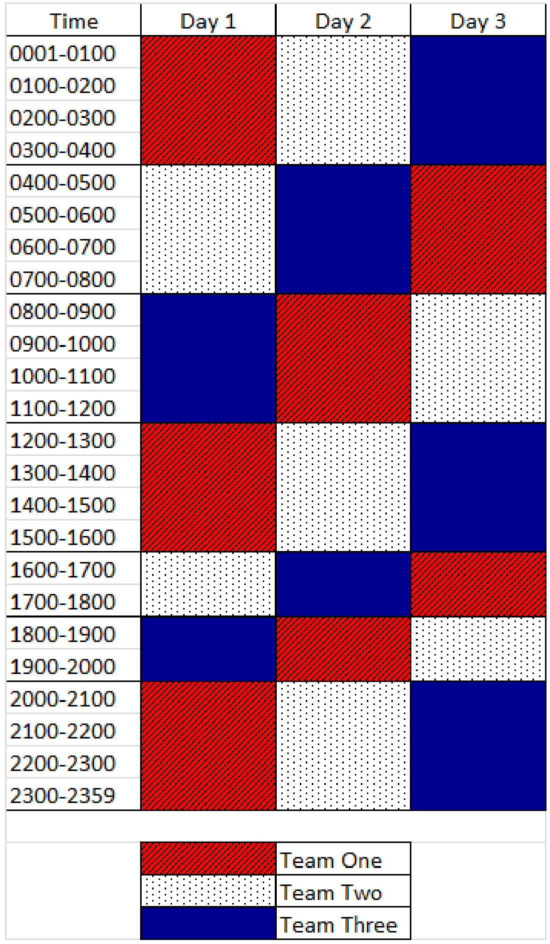

Figure 10. Three-section watch system, dogged, rotating.

As a patrol boat navigator in the late 1990s it did not occur to the author to schedule personnel in a forward rotating system, and I do not recall it being suggested. Previous to that time, when at the whim of other schedulers, I was always placed in a backwardly rotating system. These decisions were probably made because rotating forward is not as simple or intuitive as directing someone to come in earlier for the next watch rotation. Standing fixed watches (i.e. not rotating watches at all) is usually unpopular with those keeping watch in the middle of the night. But, given a diagram[5] and told that using that plan rather than rotating backward will substantially increase the minimum predicted effectiveness for personnel and decrease the percentage of time that personnel are below 70

[5] See, for example, Figure 7, p. 44.

percent predicted effectiveness (Roberts, 2012, appendix B), I would have used it. Unfortunately, I did not have this information; but by promulgating gaps and best practices, awareness of better ways to mitigate fatigue and manage crew endurance will increase.

Table 17. Best practices for naval workload studies.

Practice	Discussion
Sample size	When designing studies, use as large a sample as possible; strive to accommodate all volunteers, and seek more. A larger proportion of involvement by the ships company means that withdrawals are less likely to bias results and that sufficient data will be available to ensure that results are representative of the population. If equipment limitations are the only factor precluding additional participation, consider borrowing additional equipment.
Baseline data	Collect baseline data, for example, baseline sleep data, or three days worth of data for FAST preconditioning. A good example of this is evident in the 2007 underway submarine watch evaluation (Duplessis et al., 2007).
General data	Include general information in studies. Collect more demographic information (age, gender) than you think you may need. The opportunity is not available later, and the details could be useful to you or others for reasons that were not evident at the time of the study. Include departmental and employment data even if it may not be specifically required. If there is a question where you already think you know the answer (for example, what hours do you work?), provide a list of options and an “other” box to reduce the workload for participants.
Sleep disorders	Ask about sleep disorders. Some individuals may have, or develop, sleep disorders that impact their performance. Regular health screening should uncover these conditions, but questions can also be asked of study participants to ascertain if a sleep disorder is likely so that observations that could skew results are not included in the data set.
Objective measures	Use actigraphy rather than self-reported sleep, and download and verify actigraphic data against activity logs during the study.

Practice	Discussion
Use of previous research results	Start with a "best guess" based on research from a variety of sources, military and civilian, and potentially outside the maritime industry. If one study shows that a particular two-section watch system is much better than another one, then it is very likely that similar results will be achieved elsewhere. There is no need to reinvent the wheel. For example, Roberts' (2012) guide for commanders has information regarding many watch schedules, including day-time naval work requirements and the corresponding FAST predictions. This resource would provide a good place to start when choosing an alternative watch schedule to evaluate.
Use of chronohygiene principles	The fundamentals of shift work adhering to the chronohygiene principles listed by Duplessis et al. (2007) (some taken from as far back as 1992, and no doubt earlier) have not changed. Use them. 1. Adherence to 24-hour work-rest schedules 2. Adoption of a long, protected period of uninterrupted sleep 3. Minimization of sleep fragmentation 4. Maximizing time off between shifts 5. Limiting work shifts to eight hours per 24-hour period 6. Adopting a fixed vice rotating shift 7. Minimizing shift turnover frequency
Program sleep	Program sleep as well as—or instead of—programming work, and use the most effective watch schedule available. Examining predicted sleep is how FAST and other models predict effectiveness, and the theory can be used in reverse to ensure adequate sleep. An option used by an F-15 squadron with great success in an unpredictable operational environment was programming sleep time rather than work time, and zealously guarding sleep quality (Watt, 2009). It is not easy to achieve this for a number of reasons. First, it goes against cultural norms to encourage rest in the military. Next, protecting sleep is a complex issue, especially in shared berthing facilities. Finally, trainees must become proficient during each watch time, not just one, so rotation of watches is attractive to trainers and schedulers.
FRMS	Mandate the use of organizational FRMS. If you do not have one; get one or use someone else's.
Reporting results	Provide results in plain English, but explain exactly what you did. Sleep and crew endurance studies need to be read and understood by high-level decision makers who are not statisticians, or modeling experts. It would be helpful if titles for papers were carefully selected to maximize the chance of researchers finding them online with distribution made as broad as possible, including online access (for example, author and/or organization upload).

Practice	Discussion
Watch schedule comparisons	Comparison of watch schedules should only occur when the same number of hours per day is being worked in each.
Solutions versus treatments	Look for causes rather than treating symptoms.

V. CONCLUSIONS AND RECOMMENDATIONS

In a 2008 fatigue research meta-analysis, Allen, Wardsworth, and Smith (2008) stated “Fatigue has been noticeably under-researched in the maritime domain compared to other transport sectors.” The author would assert, happily, that this statement is now less true, but there is still some way to go. It is reasonable that workload studies are not the only task or the highest priority for any ship, naval or civilian. Operational tasking has priority; time is money in terms of delivery of goods transported by sea, and most vessels are already busy with the tasks that they have been assigned. Generally, study teams have to take what they can get and make the best of it in terms of programming/availability of platforms, the number of volunteer participants, how many drop-outs a study experiences prior to completion, the quality of self-reported data, and so on. What has been achieved with respect to workload studies and education regarding the risks of fatigue is a testament to the dedication and tenaciousness of the many individuals and groups who recognize the importance of crew endurance and the role of fatigue management, and are educating others about these issues.

A. CONCLUSIONS

All those involved in what has been done thus far in naval crew endurance management should feel rightly proud. However there is still work to be done. Major gaps in the research identified in this review are the assessment of crew endurance risk factors, particularly sleep quality, onboard ships, the formal design and conduct of evaluative research rather than descriptive efforts, and organizational level fatigue management policy and education. Fatigue risk management systems are still evolving but have not matured to a level guaranteeing sailors routinely receive adequate, quality, sleep at sea.

Based on the review of studies in this thesis three main ideas are summarized as recommendations for future crew endurance studies. These ideas are not revolutionary and their applicability extends well beyond crew endurance. A pictorial representation is provided in Figure 11.

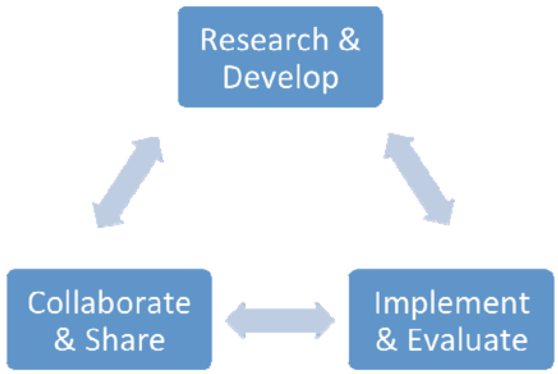

Figure 11. A recommended approach for future crew endurance studies.

1. Research and Develop

Look at what else has been done. Think critically about the limitations and assumptions from other studies. The results are not likely to be completely generalizable, but they are a better place to start than with nothing at all. If a 6-6 two-section watch system is not the best option in a Canadian submarine, it probably is not the best option in any other conventionally powered submarine either (or nuclear submarine, or surface ship). Come up with potential solutions; the more suggestions there are, the more likely it is that one of them will work. Having said that, do not get stuck here. Do not develop forever and do not keep replicating work that has already been done. Move on from descriptive study, measuring the symptoms, to more evaluative study that identifies and fixes the cause of an issue or challenge.

2. Implement and Evaluate

It is not enough to merely identify a potential problem and suggest a possible solution; researchers must also implement the suggested solution and evaluate the results to see if it works. It is then possible to take note of what does not work, and go back to research and development and see if it can be fixed.

3. Collaborate and Share

Collaboration increases a pool of information, enhances knowledge, and reduces replication. At the most basic level, crew endurance is about safety and individual performance; it does not need to be a closely held secret. Certainly there are particular aspects of studies and specific results regarding performance of military personnel that are not appropriate to share, but this should not mean that the distribution of a study is always limited.

Collaboration is not always the cheapest or easiest option. Often a collaborative team-based approach takes longer to implement than an individual one. But it is equally true that the results of collaborative efforts are generally superior, and internet-based communication has significantly reduced the potential costs of collaborating. It is important to collaborate early since this can avoid unnecessary duplication of activity.

B. RECOMMENDATIONS

This thesis does not presume to have examined all areas of crew endurance, or to have transformed the author into a subject matter expert capable of formulating organizational-level study plans for the next decade. It has shown that the RAN and the USN have invested heavily in workload studies and are evolving toward a situation where adequate, quality sleep at sea is both a possibility and a requirement. Collaboration between the RAN and the USN is excellent, and care should be taken to ensure that it does not diminish, to the detriment of both parties.

The RAN is much smaller than the USN, and there is a similar mismatch in the quantum of investment and resources available to research, based on the relative size of the population, economy, industry etc. The RAN has completed fewer workload studies than the USN, and the results are not as widely distributed. However, high study participation rates, the greater flexibility in cultural change etc. that may be possible in a smaller group, and the existence of some unique platforms and schedules could allow Australia to take the lead in some fatigue research efforts. Australia and the RAN need to be astute in leveraging opportunities and focusing expertise to maximize ongoing cooperative research in the areas of sleep, fatigue mitigation, and crew endurance.

It is recommended that the RAN pursue the following opportunities in order to further ongoing, mutually beneficial cooperative study with the USN, and other appropriate groups:

1. Provide the USN with regular updates on NMD implementation and compare results with FAST analysis (ongoing).

2. Formalize arrangements for mutually beneficial sharing of equipment, personnel expertise, and study results (1–2 years).

3. Complete detailed analysis of the RAN patrol boat data collected in 2013 and confirm applicability of results to other small warship platforms, such as LCS (1–2 years).

4. Develop a shared database of data sets to enable cross-validation and comparison of tools, especially NMD and FAST (1–3 years).

5. Investigate watch systems common in the RAN but not widely used in the USN, specifically three-section watch systems such as the one previously shown in Figure 10, p. 61 (3–5 years).

6. Complete RAN submarine studies planned for 2015–2016, in discussion with USN/NPS, and perhaps the RCN, to enable comparable analysis of results with similar USN/RCN studies (4–7 years).

7. Work toward a standalone RAN FRMS or crew endurance policy, including stricter application of fatigue policy measures for naval sea-going personnel (5–10 years).

LIST OF REFERENCES

Åkerstedt, T., Anund, A., Axelsson, J., & Kecklund, G. (2014). Subjective sleepiness is a sensitive indicator of insufficient sleep and impaired waking function. *Journal of Sleep Research, 23*(3), pp. 240–252.

Åkerstedt, T., Axelsson, J., Lekander, M., Orsini, N., & Kecklund, G. (2013). The daily variation in sleepiness and its relation to the preceding sleep episode—a prospective study across 42 days of normal living. *Journal of Sleep Research, 22*(3), pp. 258–265. doi:10.1111/jsr.12014

Allen, P., Wardsworth, E., & Smith, A. (2008). Seafarers' fatigue: A review of the recent literature. *International Maritime Health, 59*(1–4), pp. 81–92.

Ancoli-Israel, S., Cole, R., Alessi, C., Chambers, M., Moorcroft, W., & Pollak, C. (2003). The role of actigraphy in the study of sleep and circadian rhythms. American Academy of Sleep Medicine review paper. *Sleep, 26*(3), pp. 342–392.

Archibald, K. (2005). *Effects of noise, temperature, humidity, motion and light on the sleep patterns of the crew of HSV-2 SWIFT.* Master's thesis, Naval Postgraduate School, Monterey, CA.

Avers, K., & Johnson, W. B. (2011). A review of Federal Aviation Administration fatigue research: Transitioning scientific results to the aviation industry. *Aviation Psychology and Applied Human Factors, 1*(2), p. 87.

Banks, J. O., Avers, K. E., Nesthus, T. E., & Hauck, E. L. (2012). A comparative study of international flight attendant fatigue regulations and collective bargaining agreements. *Journal of Air Transport Management, 19*, pp. 21–24.

Bardach, E. (2011). *Practical guide for policy analysis: The eightfold path to more effective problem solving* (4th ed.). Washington, DC: CQ Press.

Basner, M., Mollicone, D., & Dinges, D. F. (2011). Validity and sensitivity of a brief psychomotor vigilance test (PVT-B) to total and partial sleep deprivation. *Acta Astronautica, 69*(11), pp. 949–959.

Blassingame, S. R. (2001). *Analysis of self-reported sleep patterns in a sample of US Navy Submariners using nonparametric statistics*. Master's thesis, Naval Postgraduate School, Monterey, CA.

Bowley, D., Comeau, P., Edwards, R., Hiniker, P. J., Howes, G., Kass, R. A., . . . Vaughan, J. (2006). *Guide for understanding and implementing defense experimentation (GUIDEx)* (version 1.1 ed.). Ottawa: The Technical Cooperation Program (TTCP).

Brachet, T., David, G., & Duseja, R. (2012). The effect of shift structure on performance: The role of fatigue for paramedics. *American Economic Journal: Applied Economics,* 4, (April 2012), pp. 219–246.

Brown, S. A. (2012). *Maritime platform sleep and performance study: Evaluating the SAFTE model for maritime workplace application.* Master's thesis, Naval Postgraduate School, Monterey, CA.

Browning, B. W. (2007). Leadership in desperate times: An analysis of endurance: Shackleton's incredible voyage through the lens of leadership theory. *Advances in Developing Human Resources, 9*(2), pp. 183–198.

Caldwell, J. A., Mallis, M. M., Caldwell, J. L., Paul, M. A., Miller, J. C., & Neri, D. F. (2009). Fatigue countermeasures in aviation. *Aviation, Space, and Environmental Medicine, 80*(1), pp. 29–59.

Chapman, S. (2001). *The management of stress and fatigue amongst Royal Australian Navy submariners: A strategic, operational and financial imperative*. Canberra A.C.T.: Australian Defence Headquarters, Department of Defence.

Cheung, B., Vatanian, O., Hofer, K., & Bouak, F. (2010). *General recommendations on fatigue risk management for the Canadian Forces.* (Technical Report No. TR 2010-056). Toronto: Defence Research and Development Canada.

Commonwealth of Australia. (2014). *Australian book of reference 6303-Navy safety systems manual* (5th ed.). Canberra.

Comperatore, C. A., Rivera, P. K., & Carvalhais, A. B. (2005). *U.S. Coast Guard guide for managing crew endurance risk factors - version 2.1.* (Final Report No. CG-D-13-01). Groton, CT: U.S. Coast Guard Research and Development Center.

Cordle, C. J., & Shattuck, N. L. (2013). A sea change in standing watch. *U.S. Naval Institute Proceedings* (239 January 2013), pp. 35–39.

Davey, J. D. (2013). *Effects of sleep deprivation on U.S. Navy watchstander performance onboard the independence class littoral combat ship (LCS-2).* Master's thesis, Naval Postgraduate School, Monterey, CA.

Duplessis, C., Miller, J., Crepeau, L., Osborn, C., & Dyche, J. (2007). Submarine watch schedules: Underway evaluation of rotating (contemporary) and compressed (alternative) schedules. *The Journal of the Undersea & Hyperbaric Medical Society Inc., 34*(1), p. 21.

Folkard, S., Lombardi, D. A., & Tucker, P. T. (2005). Shiftwork: Safety, sleepiness and sleep. *Industrial Health, 43*(1), pp. 20–23.

Folkard, S., & Tucker, P. (2003). Shift work, safety and productivity. *Occupational Medicine (Oxford, England), 53*(2), pp. 95–101.

Ftouni, S., Sletten, T. L., Howard, M., Anderson, C., Lenné, M. G., Lockley, S. W., & Rajaratnam, S. M. (2013). Objective and subjective measures of sleepiness, and their associations with on-road driving events in shift workers. *Journal of Sleep Research, 22*(1), pp. 58–69.

Gallicchio, L., & Kalesan, B. (2009). Sleep duration and mortality: A systematic review and meta-analysis. *Journal of Sleep Research, 18*, pp. 148–158.

Gamboa, D. S. (2002). An analysis of the effects of environment and career longevity on the sleep patterns of enlisted U.S. Navy submariners. Master's thesis, Naval Postgraduate School, Monterey, CA.

Gander, P. H. (1986). International aircrew sleep and wakefulness after multiple time zone flights: A cooperative study. In R. Curtis Graeber (Ed.), *Crew factors in flight operations: IV. Sleep and wakefulness in international aircrews.* NASA Technical Memorandum 88231. Ames Research Center, Moffett Field, California.

Gander, P. H., Mulrine, H. M., Berg, M. J., Smith, A. A. T., Signal, T. L., Wu, L. J., & Belenky, G. (2014). Effects of sleep/wake history and circadian phase on proposed pilot fatigue safety performance indicators. *Journal of Sleep Research.* doi:10.1111/jsr.12197.

Grech, M. R., Roberts, W., Hamilton, S., Turner, M., & Cleary, A. (2014). *Crew endurance at sea: An analysis of sleep, work-hours and fatigue across a deployment period (HMAS WARRAMUNGA).* Unpublished manuscript.

Green, K. Y. (2009). *A comparative analysis between the Navy Standard Workweek and the actual work/rest patterns of sailors aboard U.S. Navy frigates.* Master's thesis, Naval Postgraduate School, Monterey, CA.

Härmä, M., Partinen, M., Repo, R., Sorsa, M., & Siivonen, P. (2008). Effects of 6/6 and 4/8 watch systems on sleepiness among bridge officers. *Chronobiology International, 25*(2–3), pp. 413–423.

Haynes, L. E. (2007). *A comparison between the Navy Standard Workweek and the actual work and rest patterns of U.S. Navy sailors.* Master's thesis, Naval Postgraduate School, Monterey, CA.

Huey, B. M., & Wickens, C. D. (Eds.). (1993). *Workload transition: Implications for individual and team performance.* Washington, DC: National Academies Press.

Inbody, D. (2014, November 3). [Facebook comment]. Retrieved from https://www.facebook.com/NPSmonterey?fref=ts

Lauderdale, D. S., Knutson, K. L., Yan, L. L., Liu, K., & Rathouz, P. J. (2008). Self-reported and measured sleep duration: How similar are they? *Epidemiology (Cambridge, Mass.), 19*(6), pp. 838–845. doi:10.1097/EDE.0b013e318187a7b0.

Lazzaretti, P. C. (2008). *HSI in the USN frigate community operational readiness and safety as a function of manning levels.* Master's thesis, Naval Postgraduate School, Monterey, CA.

Lutzhoft, M., Thorslund, B., Kircher, A., & Gillberg, M. (2007). *Fatigue at sea.* (No. VTI Rapport 568A). Sweden: VTI.

Mason, D. R. (2009). *A comparative analysis between the Navy Standard Workweek and the work/rest patterns of sailors aboard U.S. Navy cruisers*. Master's thesis, Naval Postgraduate School, Monterey, CA.

McLean, A., Grech, M. R., & Elischer, P. (2009). *Collins class submarines: A study examining the issues of sleep, fatigue and performance.* (No. DSTO-TR-2341 (Limited Distribution)). Melbourne, Australia: Maritime Platforms Division, Defence Science and Technology Organisation.

MedicineNet.com. (2014). Definition of fatigue. Retrieved from http://www.medterms.com/script/main/art.asp?articlekey=9879

Miller, C., & Strohl, K. P. (2011). Sleep deprivation: Legislative and regulatory issues. *Sleep and Safety.* p. 156.

Miller, J. C., Smith, M. L., & McCauley, M. E. (1998). *Crew fatigue and performance on U.S. Coast Guard cutters.* (No. CG-D-10-99 2). Washington, DC: U.S. Coast Guard.

Miller, N. L., Crowson, Jr., J. J., & Narkevicius, J. M. (2003). Human Characteristics and Measures in Systems Design. In Booher, H. R. (Ed.), Handbook of human *systems integration* (Vol. 23) (Chapter 19, pp. 699–742). New York: John Wiley & Sons.

Miller, N. L., & Nguyen, J. (2003). Working the nightshift on the USS JOHN C. STENNIS: Implications for enhancing warfighter effectiveness. *Proceedings of the Aerospace Medical Association,* May.

Miller, N. L., Shattuck, L. G., & Matsangas, P. (2007). *Fatigue in military operational environments*. Monterey, CA: Naval Postgraduate School.

Molenda, Michael. (2003). In search of the elusive ADDIE model. *Performance improvement,* 42(5), May-June, pp. 34–37.

Monk, T. H., Buysse, D. J., Billy, B. D., Fletcher, M. E., & Kennedy, K. S. (2013). Polysomnographic sleep and circadian temperature rhythms as a function of prior shift work exposure in retired seniors. *Healthy Aging & Clinical Care in the Elderly,* 2013(5), pp. 9–19. doi:10.4137/HACCE.S11528.

Naval Postgraduate School (2014, November 3). [Facebook post]. Retrieved from https://www.facebook.com/NPSmonterey?fref=ts

Nguyen, J. L. (2002). *The effects of reversing sleep-wake cycles on sleep and fatigue on the crew of USS JOHN C. STENNIS.* Master's thesis, Naval Postgraduate School, Monterey, CA.

O'Connor, P., Buttrey, S. E., O'Dea, A., & Kennedy, Q. (2011). *An assessment of the relationship between safety climate and mishap risk in U.S. Naval aviation*. NPS-OR-11-004. Monterey, CA: Naval Postgraduate School.

Osborn, C. (2004). *An analysis of the effectiveness of a new watchstanding schedule for U.S. submariners.* Master's thesis, Naval Postgraduate School, Monterey, CA.

Oxford Dictionaries. (2014a). Endurance. Retrieved from http://www.oxforddictionaries.com/us/definition/english/endurance

Oxford Dictionaries. (2014b). Sleep. Retrieved from http://www.oxforddictionaries.com/definition/english/sleep

Paul, M. A., Hursh, S. R., & Miller, J. C. (2010). *Alternative submarine watch schedules: Recommendations for a new CF watch schedule.* (No. TR 2010-001). Toronto: Defence Research and Development Canada.

Paul, M. A., Ebisuzaki, D., McHarg, J., Hursh, S. R., & Miller, J. C. (2012). *An assessment of some watch schedule variants used in CDN patrol frigates: OP Nanook 2011.* (No. TR 2012-078). Toronto: Defence Research and Development Canada.

Paul, M. A., Gray, G. W., Nesthus, T. E., & Miller, J. C. (2008). *An assessment of the CF submarine watch schedule variants for impact on modeled crew performance.* (No. TR 2008-007). Toronto: Defence Research and Development Canada.

Project Horizon. (2012). *Project horizon - A wake-up call: Research into the effects of sleepiness on the cognitive performance of maritime watchkeepers under different watch patterns, using ships' bridge, engine and liquid cargo handling simulators.* (No. 8022 NAU014 Horizon Report GB.indd).

Pulsar Informatics & FAA. (2014). FAA fatigue risk assessment tool. Retrieved from http://faafatiguerisk.pulsarinformatics.com

Roberts, D. A. (2012). *Analysis of alternative watch schedules for shipboard operations: A guide for commanders.* Master's thesis, Naval Postgraduate School, Monterey, CA.

Roberts, C. R. (2014, May 7). An interview with Cutler Dawson. *The News Tribune*. Retrieved from http://www.thenewstribune.com/2014/05/07/3184613/an-interview-with-cutler-dawson.html

Sawyer, T. L. (2004). *The effects of reversing sleep-wake cycles on mood states, sleep, and fatigue on the crew of the USS JOHN C. STENNIS.* Master's thesis, Naval Postgraduate School, Monterey, CA.

Tryon, W. W. (2004). Issues of validity in actigraphic sleep assessment. *Sleep. 27*(1), pp. 158–165.

Turner, M. J. (2012, June 12). *Performance Optimisation and the use of FMD/MONICAR onboard HMAS WARRAMUNGA*. HMAS WARRAMUNGA Administrative Instructions, Commanding Officers Directive 27/12.

Tvaryanas, A. P., & Miller, N. L. (2010). *Human systems integration domain trade-offs in optimized manning - The task effectiveness scheduling tool*. Monterey, CA: Naval Postgraduate School.

United States Coast Guard. (2005). *Implementing the crew endurance management system on towing vessels*. Washington, DC: U.S. Department of Homeland Security.

United States Department of Transportation. (2013). In Federal Aviation Administration (Ed.), *Advisory circular: Fatigue risk management systems for aviation safety* (AC 120-103A ed.).

United States Department of Transportation. (2011). *Flightcrew member duty and rest requirements.* Retrieved from http://www.faa.gov/regulations_policies /rulemaking/recently_published/media/2120-AJ58-FinalRule.pdf

United States Navy. (2009). *NATOPS general flight and operating instructions OPNAV INSTRUCTION 3710.7U*. Retrieved from http://doni.daps.dla.mil/Directives/03000%20Naval%20Operations%20and %20Readiness/03-700%20Flight%20and%20Air%20Space%20Support% 20Services/3710.7U.pdf

van Leeuwen, W. M., Kircher, A., Dahlgren, A., Lützhöft, M., Barnett, M., Kecklund, G., & Åkerstedt, T. (2013). Sleep, sleepiness, and neurobehavioral performance while on watch in a simulated 4 hours on/8 hours off maritime watch system. *Chronobiology International*, *30*(9), pp. 1108–1115.

Watt, C. G. (2009). *Aircrew fatigue management.* Master's thesis, Air War College, Montgomery, AL.

Yokeley, M. T. (2012). Effects of sleep deprivation on U.S. Navy surface ship watchstander performance using alternative watch schedules. Master's thesis, Naval Postgraduate School, Monterey, CA.

Young, R. L. (2013). A comparison of sleep and performance of sailors on an operationally deployed US Navy warship. Master's thesis, Naval Postgraduate School, Monterey, CA.

THIS PAGE INTENTIONALLY LEFT BLANK

INITAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California