NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

ASSOCIATION BETWEEN DRIVER-REPORTED SLEEP AND PREDICTED LEVELS OF EFFECTIVENESS BASED ON THE FATIGUE AVOIDANCE SCHEDULING TOOL

by

Aaron E. Heisinger

June 2009

Thesis Advisor: Nita Lewis Miller
Second Reader: Matthew G. Boensel

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 June 2009	3. REPORT TYPE AND DATES COVERED Master's Thesis	
4. TITLE AND SUBTITLE Association Between Driver-Reported Sleep and Predicted Levels of Effectiveness Based on the Fatigue Avoidance Scheduling Tool		**5. FUNDING NUMBERS**	
6. AUTHOR(S) Christopher D. Bertram			
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.

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

13. ABSTRACT (maximum 200 words)

    Current military operations require a high state of operational readiness. Service members and civilian workers are tasked with performing in a near non-stop environment without proper rest and recuperation. Unit and individual effectiveness depend upon initiative, judgment, courage, and motivation, which are all enhanced by the ability to think clearly and logically – attributes that are degraded by fatigue. This thesis seeks to determine the extent to which fatigue plays a part in human factors related to large truck mishaps. This study is conducted using the Large Truck Crash Causation Study data base and assesses drivers' predicted level of effectiveness employing the Sleep, Activity, Fatigue, and Task Effectiveness Model as instantiated in the Fatigue Avoidance Scheduling Tool (FAST). The entire population of truck crashes is categorized into two groups, those with human factors causes and those with non-human factors causes. A comparison of the two groups shows a statistically significant difference between the two groups in reported sleep and predicted levels of effectiveness. This result shows that fatigue is more prevalent and is potentially an important contributing factor to human factors related mishaps. Heightened levels of fatigue diminish situational awareness, judgment, and decision-making capabilities and can result in serious, sometimes even deadly consequences. It is recommended that fatigue avoidance strategies such as FAST be implemented in training and operational planning. Such strategies can assist in the development of more efficient and potentially safer sleep-work schedules.

14. SUBJECT TERMS Large truck crash, long-haul truck drivers, heavy truck operations, Large Truck Crash Causation Study, effects of fatigue on driver performance, long-haul operations, commercial transportation carriers, Federal Motor Carrier Safety Administration, National Highway Traffic Safety Administration			15. NUMBER OF PAGES 81
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT UU

NSN 7540-01-280-5500

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

THIS PAGE INTENTIONALLY LEFT BLANK

ASSOCIATION BETWEEN DRIVER-REPORTED SLEEP AND PREDICTED LEVELS OF EFFECTIVENESS BASED ON THE FATIGUE AVOIDANCE SCHEDULING TOOL

Aaron E. Heisinger
Captain, United States Marine Corps
B.A., Linfield College, 2000

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

**NAVAL POSTGRADUATE SCHOOL
June 2009**

Author: Aaron E. Heisinger

Approved by: Associate Professor Nita Lewis Miller
                    Thesis Advisor

                    Senior Lecturer Matthew G. Boensel
                    Second Reader

                    Robert F. Dell
                    Chairman, Department of Operations Research

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

Current military operations require a high state of operational readiness. Service members and civilian workers are tasked with performing in a near non-stop environment without proper rest and recuperation. Unit and individual effectiveness depend upon initiative, judgment, courage, and motivation, which are all enhanced by the ability to think clearly and logically – attributes that are degraded by fatigue. This thesis seeks to determine the extent to which fatigue plays a part in human factors related to large truck mishaps. This study is conducted using the Large Truck Crash Causation Study data base and assesses drivers' predicted level of effectiveness employing the Sleep, Activity, Fatigue, and Task Effectiveness Model as instantiated in the Fatigue Avoidance Scheduling Tool (FAST). The entire population of truck crashes is categorized into two groups, those with human factors causes and those with non-human factors causes. A comparison of the two groups shows a statistically significant difference between the two groups in reported sleep and predicted levels of effectiveness. This result shows that fatigue is more prevalent and is potentially an important contributing factor to human factors related mishaps. Heightened levels of fatigue diminish situational awareness, judgment, and decision-making capabilities and can result in serious, sometimes even deadly consequences. It is recommended that fatigue avoidance strategies such as FAST be implemented in training and operational planning. Such strategies can assist in the development of more efficient and potentially safer sleep-work schedules.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

EXECUTIVE SUMMARY

Current military operations require a high state of operational readiness. Service members and civilian workers are tasked with performing in a near non-stop environment without proper rest and recuperation. Unit and individual effectiveness depend upon initiative, judgment, courage, and motivation, which are all enhanced by the ability to think clearly and logically – attributes that are degraded by fatigue.

Fatigue is a major concern in managing today's ongoing military operations. Marines, Sailors, Soldiers and Airmen are experiencing more frequent and longer deployments to austere locations. Fatigue is so commonplace that it is often unrecognized by our troops. Mission accomplishment is foremost in minds of our armed services that fatigue is either ignored or accepted as the price of high operational tempo. In any case, fatigue is often unreported. However, heightened levels of fatigue diminish situational awareness, judgment, and decision-making capabilities and can result in serious, perhaps deadly consequences.

This thesis studies the possible effects of fatigue on operator performance in a dynamic environment. This thesis uses the Large Truck Crash Causation Study (LTCCS), which is a collection of detailed accounts of over 1,000 large truck crashes in the U.S. from 2001 to 2003. Sleep schedules derived from driver estimates of sleep following crash events are used to develop sleep-work schedules. These schedules are used to calculate predicted levels of effectiveness from the Fatigue Avoidance Scheduling Tool (FAST), which uses the Sleep, Activity Fatigue and Task Effectiveness (SAFTE) Model.

There are 545 cases that contained sufficient sleep histories to establish valid sleep-work schedules. These cases are divided into two causal groups (human factors (HF) and non-human factors (NONHF) based on a survey containing short, detailed descriptions of causal factors that lead to the accident. The survey is administered to six subject matter experts seeking their opinion on whether or not the cause of the accident is HF related.

This thesis finds that fatigue is much more prevalent in HF-related mishaps. The FAST results revealed that approximately 16% of the entire HF population was operating below 90% predicted level of effectiveness at the time of the accident. This compares to roughly 6% of the NONHF population. Further, operators in almost 3% of the HF cases fell below 77% level of effectiveness while not a single case of the NONHF group fell below 83% effectiveness. The significance of low effectiveness is that risk of accidents increases with increasing fatigue.

This result shows that fatigue is a potential contributing factor to human factors related mishaps. Heightened levels of fatigue diminish situational awareness, judgment, and decision-making capabilities and can result in serious, sometimes even deadly consequences. It is recommended that fatigue avoidance strategies such as FAST be implemented in training and operational planning. Such strategies can assist in the development of more efficient and potentially safer sleep-work schedules.

ACKNOWLEDGMENTS

I would like to acknowledge Dr. Nita Miller and Senior Lecturer Matthew G. Boensel for their continuous support, encouragement, guidance, and perspective with assisting me in attaining this achievement.

I would also like to thank Dr James C. Miller for his time, assistance and expertise in this field of study, especially FAST; Rob O'Donnell for keeping me up-to-date with the latest version of FAST as well as assisting with any software issues; LtCol Sergio Posadas for helping me with my writing technique and enforcing the "Light, Layered and Linked" concept; and Prof Sam Buttrey for, on his own time and breaking from his busy schedule, providing key insight and assistance on the statistical analysis of the data.

To Finnley and Jacob, the best daughter and son a dad could ask for. Thank you for your hugs and smiles of encouragement while I was away at school. I look forward to this summer and spending more time with both of you.

I would also like to thank my wife, Jessica. Without her support and steadfast devotion to our family, this thesis effort would not have been possible. Her unconditional support throughout this curriculum and thesis process made this endeavor possible.

THIS PAGE INTENTONALLY LEFT BLANK

I. INTRODUCTION

A. EFFECTS OF FATIGUE ON OPERATIONAL PERFORMANCE

The effects of fatigue continue to be a major concern in today's military operations. With current U.S. military obligations spanning the globe, an increasing number of military personnel and civilian workers are asked to operate at higher levels of performance without receiving adequate sleep (Drummond, 2007). Our service members are working longer hours in near continuous combat operations around the clock with little opportunity to rest properly (Davenport, 2007). "Fatigue is so prevalent and such a part of our culture that we scarcely see or recognize it. It's the big gray elephant we muscle out of the cockpit when we fly, step around when we enter the bridge and push aside when we peer into the periscope" (Davenport, 2007).

The world situation is one of continual and complex conflict. Military operations around the world require a high state of operational readiness. An important part of this readiness is maintaining fully operational forces home and abroad. These challenges make well-developed sleep-work-rest schedules crucial. A study conducted by the British Medical Association (August 2000) on service members found sleep deprivation to have a negative effect on mood, motivation, attention, alertness, short-term memory, ability to accomplish routine tasks, job performance and physical ability. Effects of sleep deprivation may vary from person to person and are dependent on several factors (e.g. age, gender, physical and mental health, and the operational environment) (British Medical Association, 2000).

Sleep deprivation is known to impair some aspects of working memory and the ability to multi-task or solve complex problems. Sleep deprivation can also have a more serious effect; it tends to reduce an individual's sensitivity to risk (Drummond, 2007). This risk is especially important to the commanders leading combat forces (Drummond, 2007).

Fatigue is also an important issue in long-haul truck operations. In comparison to ongoing military operations, long-haul truck drivers experience similar, irregular working

conditions such as driving for extended periods of time and working inconsistent schedules. Both groups are tasked to perform in near continuous, around the clock operations while receiving inadequate rest. Working assignments may require them to deploy or be on the road for months at a time. Working extended hours can cause the individual to eat, sleep and perform at irregular times.

Long-haul drivers regularly drive extended hours and face deadlines while conforming to demanding schedules (Johnson, 2009). Drivers who do not receive adequate sleep or stay awake for long periods of time are subjected to sleep deprivation. Sleep deprivation is known to impair judgment, slow reaction times, and decrease the ability to assess one's own level of fatigue (Moore et al., 2009; Mara, 1998). This degradation in driver performance can lead to serious consequences or even death.

This thesis provides insight on the effects of fatigue on operator performance based on the data for long-haul truck driver accidents. This thesis uses the Large Truck Crash Causation Study (LTCCS) dataset to develop sleep-work schedules that are read into the Fatigue Avoidance Scheduling Tool (FAST) to obtain predicted levels of effectiveness. The purpose is to identify possible effects of fatigue on human factors related accidents.

B.   THESIS ORGANIZATION

This thesis is organized into five chapters. The first chapter provides an introduction to the purpose of the thesis. Chapter II contains the literature review that covers sleep, sleep deprivation, fatigue, the LTCCS, the SAFTE Model, and the FAST. Chapter III provides the methodology of the study, including how the data for the study is selected and used. Chapter IV presents the results from the analysis. Lastly, Chapter V contains the discussion of the results, conclusions from the thesis, and recommendations for follow-on research.

II. LITERATURE REVIEW

A. INTRODUCTION

To understand the effects of fatigue on human factors related accidents, key terms must be defined: sleep and its effects on fatigue; the purpose, methodology and functionality of the LTCCS; and the SAFTE Model and FAST. This chapter begins by focusing on sleep, sleep deprivation and effects of fatigue on human performance. Next, the LTCCS is described to fully detail the reason for its study and its usefulness in conducting this thesis. Lastly, the SAFTE Model and FAST are discussed to provide the reader with a general understanding of how FAST uses SAFTE to produce predicted levels of effectiveness.

B. SLEEP

Sleep is defined as "a natural, regularly recurring condition of rest for the body and mind" by Webster's New World College Dictionary (2009). Even though we spend almost one-third of our lives asleep, the precise functions of sleep eludes researchers (Miller et al., 2007). Sleep is a physiological requirement that must be met for sustaining a healthy lifestyle while balancing behavior (mood), improving memory and concentration, and maintaining cognitive performance levels (Moore et al., 2009). The average adult requires roughly eight hours of sleep per night to offset sleep debt and achieve optimal performance (Miller et al., 2007).

Miller et al. (2007) describe two categories of sleep experienced by the human brain: rapid eye movement (REM) and non-rapid eye movement (NREM). Each type of sleep serves a different purpose and is broken down by observable changes in behavior. NREM sleep is divided into four gradually deeper sleep stages and REM sleep has a single sleep period. According to Miller et al. (2007), adequate amounts of both REM and NREM sleep are needed for optimal human performance.

Sleep debt is the termed used for the accumulation of lost sleep. Sleep debt may be caused from acute sleep deprivation, due to an extended period of wakefulness; shift

work changes; or travel crossing time zones. Acute sleep debt can be observed when conducting continuous operations or performing for extended periods of time. Another form of sleep debt, and one common among service members and truck drivers, is chronic sleep debt. Chronic sleep debt is when an individual goes multiple nights of receiving inadequate amounts of sleep (i.e., less than eight hours per night) and is typically seen during sustained operations. No matter if sleep debt is attributed to acute or chronic sleep deprivation, sleep debt will continue to accrue and at some point must be paid back. The only form of "re-payment" is adequate amounts of quality sleep.

C.   SLEEP DEPRIVATION

Sleep deprivation can be dangerous not only to the individual but also dangerous to others. Sleep deprivation adversely affects cognitive performance and decision-making abilities. Specific tasks, such as driving or flying, require certain levels of technical proficiency which present serious risks or safety hazards if the person is sleep deprived. Colten and Altevogt (2006) have stated that there are an estimated 110,000 injuries and 5,000 fatalities per year in vehicular accidents involving large or commercial trucks. In their publication, driver fatigue (including sleep deprivation) is determined to be the causal factor in 57% of the accidents that lead to the driver's death (Colten & Altevogt, 2006).

Partial sleep deprivation (PSD) can occur when a person is deprived of a single stage of sleep. Miller et al. (2007) stated that if given the opportunity to sleep after PSD, a person will return to the stage of sleep that is missed. Sleep that is lost can then be recovered. Total sleep deprivation (TSD) arises when an individual is kept awake continuously. When provided the chance to sleep following TSD, the individual will recover by entering into the deep stages of sleep.

Chronic sleep deprivation often occurs in those who work long hours (e.g., military personnel, emergency responders or doctors, truck drivers) or those who suffer from sleep disorders (e.g., sleep apnea, insomnia) that interfere with sleep on a regular basis (Miller et al., 2007). Some signs of sleep deprivation are: difficulty waking up in

the morning; fluctuations in behavior/temperament; diminished cognitive performance and decision-making capability; and falling asleep while on the job (Saisan et al., 2008).

D.   FATIGUE

Humans have an internal biological need for sleep with a circadian clock that prompts them to sleep at night with periods of activity during the day (Rosekind et al., 1996). Long hours of wakefulness, shift work, change in work schedules, jet lag, sleep deprivation and circadian desynchronization can significantly degrade productivity (Miller, 2008). These factors or occurrences may lead to fatigue, decreased vigilance, poor decision-making, and other performance effects that diminish operational effectiveness.

Early warning signs of fatigue include lethargy, apathy, moodiness, reduced vigilance, fixation, slower reaction times, and poor decision making (Dinges, 1995). The average person requires eight hours of sleep per night; however, individual sleep requirements will vary. Some individuals are less affected by sleep deprivation than others. In general, it is assumed that a loss of two hours of sleep in a single night can significantly reduce effectiveness the next day (Drake, 2008). Sufficient amounts of quality sleep are required each night to ensure optimal performance.

Fatigue often sets in as a result from an unfulfilled need for sleep. As fatigue advances, the brain may cause uncontrolled shifts from attentiveness to sleep to meet physiological needs (Drake, 2008). Drake defines microsleeps as involuntary sleep lapses that may last for a couple of seconds up to a few minutes. The more fatigued a person is, the longer in duration and more frequent these microsleeps become (Drake, 2008). These microsleep episodes, also called mental lapses, can have dangerous consequences "as the brain has switched to sleep mode and is not processing stimuli" (Davenport, 2005). Davenport continues by asserting that this may be particularly dangerous, as the individual does not even realize that these microsleeps are taking place (Davenport, 2005).

Fatigue and sleepiness may not be apparent in troops engaged in rigorous activity due to noise levels, excitement, physical activity, caffeine or nicotine usage, thirst and hunger, and/or operational commitments. Sleep-deprived troops may not recognize fatigue or sleepiness during pre-mission drills or rehearsal because of built-up anticipation, excitement or anxiety. However, once activity subsides, and tropps are awaiting transportation or standing by for follow-on orders, weariness and other fatigue-related symptoms are often exhibited.

Long-haul truck drivers are conditioned to their surroundings and lengthy time spent on the road. They are accustomed to long hours of wakefulness, and extended periods of time performing the same task, while receiving little time to rest properly. The droning of the engine, the monotonous task of driving mile after mile, and frequent use of caffeine or nicotine are just a few aspects of long-haul truck operations.

Currently, long-haul truck drivers are driving longer hours, logging more miles and taking shorter rest stops to increase profits (Johnson, 2009). The U.S. Department of Transportation ruled in 2005 that truck drivers who transport goods or property may not drive more than eleven hours after ten consecutive hours off (Federal Motor Carrier Safety Administration, 2009). Many large trucks are equipped with a sleeper or berthing area which allows the driver to pull over and rest when they feel the need to sleep. Mitler et al (1997) conducted a sleep study of fatigue effects on long-haul truck drivers, investigating 80 drivers. Their study examined the routes and hours driven- per day and compared four driving schedules (two in the U.S. involving five 10-hour trips and two in Canada involving four 13-hour trips). The study showed empirically that the drivers averaged 5.18 hours in bed and a mere 4.78 hours of actual sleep over a five-day period (Mitler et al., 1997). In their conclusion, they present their findings that long-haul truck drivers received less sleep than is required to maintain proper alertness while driving.

E.   LARGE TRUCK CRASH CAUSATION STUDY

The Motor Carrier Safety Improvement Act of 1999 required the Federal Motor Carrier Safety Administration (FMCSA) to conduct a study to determine the causes of, and contributing factors to, crashes involving commercial motor vehicles (Hedlund and

Blower, 2006). FMSCA and the National Highway Traffic Safety Administration (NHTSA) wanted to understand the causes behind serious accidents involving large trucks (trucks that have a gross weight over 10,000 pounds). The resulting report includes an extensive database of large truck crashes and focuses on crash prevention analysis. The purpose of the project is to raise awareness of large truck crash causes.

From April 2001 to December 2003, 120,000 large truck mishaps were reported across the nation. A representative sample is selected for inclusion in the LTCCS using 24 data collection sites in 17 states by researchers from NHTSA's National Automotive Sampling System (NASS) and state truck inspectors. Data is gathered on 1,070 crashes involving 2,284 vehicles. Developers of the study made certain that each accident analyzed in the study had at least one large truck and that the crash resulted with at least one party being injured. Each accident report is investigated by experienced crash and truck inspectors - both responsible for recording post-crash inspections. Each case came under several reviews where additional information and assessments are coded at a central location. The data collected contained thorough crash investigation reports, truck inspection records, photographs of the crash scene and crash diagrams, driver history and sleep records, and interviews from drivers involved in the crash. All of these aimed at creating or uncovering all the details of the crash.

1. LTCCS Methodology

The LTCCS focused on the events taking place prior to the crash. The decision about what type of data to collect is driven by the desire to show the wide range of factors that are linked to large truck accidents. Detailed information about vehicle description; physical condition and experience of the drivers; truck carrier operation, truck components, and load; and road and environmental conditions are gathered so that the role of each variable could be reviewed (Hedlund & Blower, 2006).

The approach taken in the analysis and in data collection centered on the idea that traffic accidents are probabilistic incidents. It is important that the analysts and truck inspectors obtain a good description of the pre-crash events that lead to a mishap. Specific LTCCS variables are essential to the methodology and understanding of the

study. LTCCS created a method of coding pre-crash events such as critical event, critical reason, crash type and pre-crash decision-making. LTCCS created a critical event and critical reason variable to assist with the assessment or determination of pre-crash causation that led to the crash.

Establishing critical events is the first step. Follow-on data are built off this critical event. Hedlund and Blower (2006) defined *critical events* as the event that immediately precipitated the crash. Only one critical event is determined per documented crash. The critical event is the event that places the vehicles on a path that makes the accident inevitable. C*ritical reason* is termed as the immediate failure that led to that critical event (Hedlund & Blower, 2006). The critical reason is why the critical event occurred. Hedland and Blower (2006) divided the critical reasons into three categories: driver, vehicle, or environment. Potential critical reasons included driver condition and decision making, roadway and environmental conditions, and vehicle failure(s).

Critical reasons are used to provide information on why the critical event occurred. Critical reason is not intended to find the "cause" of the accident but rather as one aspect of the information about how the mishap occurred. A critical reason is the basis why the crash is unavoidable or destined to occur.

2.   Data Analysis

There are 1,070 crashes recorded involving a combination of large trucks and other vehicles — 2,284 vehicles in all. This study provides more data about truck crashes than can be located anywhere else, as it includes different types of motor carriers, mechanical condition of the trucks and vehicles, detailed information about driver condition and decision making, and recent sleep schedules of drivers. Because of the large data set, several types of analyses from descriptive statistics to conditional probability calculations can be performed.

LTCCS data may be used to evaluate conditional probabilities to measure the risks involved in crashes for drivers and/or vehicles. As an example, driving at night and

using prescription drugs did not necessarily cause accidents; however, each individual factor is run through a system to show increased risk. LTCCS provides details on the outcome in each crash and by properly implementing the right analyses; one can test for involvement of driving at night, the use of prescription drugs, or the interaction of the two.

3.   LTCCS Limitations

The study is not appropriate for evaluating risk factors across all subsets of traffic accidents (e.g., crashes occurring on rural highways, inner city streets). The overall risk or probability of a crash on particular roadways cannot be ascertained from the data set alone, even if the type of crash on certain roads is identifiable.

Insufficient or missing data limited the size of the dataset for follow-on analysis and possibly presented bias (Hedland & Blower, 2006). The authors of the study did not have the chance to review the data for accuracy or completeness when submission of the final LTCCS file occurred. Variables that are directly observed by LTCCS personnel are considered accurate and complete (e.g., vehicle and environmental data). However, variables that are offered second-hand or derived from interviews may be interpreted as incomplete or biased by some users of the LTCCS dataset.

4.   Functionality of the LTCCS

The LTCCS provides a general-purpose data file developed for problem identification. The dataset is intended to be used to assist in estimating the number of truck accidents involving a specific factor and the association of that factor to the crash (Hedlund & Blower, 2006). The study collected over 1,000 data variables, describing all aspects of the crash (i.e., drivers, vehicles, and environment), and their assessments are considered to be thorough and complete. The large and detailed dataset allows for assessment of certain isolated factors (e.g., demographics, driving experience, and sleep history) from immediate events of the crash. The robustness of the LTCCS offers potential consideration and evaluation for implementing several types of crash countermeasures.

The LTCCS dataset may be useful in supporting studies of causation (Hedlund, 2003). As stated previously, the purpose of the research is to collect and preserve objective and detailed information about pre-crash events and parties involved. The data base can also be used for identifying and evaluating the significance of an issue and then comparing the issue among others. In the end, the data may assist in explaining physical and behavioral incidents involved in crashes that may be better understood in order to develop and test interventions to reduce such occurrences.

5.   **Follow-on Studies on Large Truck Accidents**

The LTCCS dataset contains the largest and most detailed account of large truck collisions to date. Several studies use the data provided in the LTCCS. The database can be used to study crash risk using a wide range of statistical analysis. Using the LTCCS dataset provides endless possibilities on how to approach crash risk, causation analysis, or various methods of countermeasures.

Paquette (2007) conducted a pilot study using the LTCCS database. His study estimated performance levels in the drivers that are documented in the LTCCS. The goal of his pilot study is to estimate the contribution of driver fatigue to the causes of large truck crashes (Paquette, 2007). He implemented a case-control study design in an attempt to evaluate the relationship between human factors mishaps and driver fatigue. Paquette concluded from evaluating a small subset of the data that driver fatigue played a significant role in large truck crashes. Although this has been the only study undertaken that looked directly at reported sleep history, his study did not use all of the data that is readily available from the LTCCS datasets.

Knipling (2004) from Virginia Tech Transportation Institute reported on crash risk and its variance among commercial truck drivers. Using the LTCCS data set, he found the "critical reason" for the crash is attributed to the other driver or vehicle 70% of the time while only 30% of the time to the truck driver. Reviewing another study examining "fault" in a 1994-97 North Carolina police-reported truck crash, Knipling concluded that truck drivers are assigned "fault" in 48% of the accidents versus 40% for the other vehicle (Knipling, 2004).

Krishnaswami and Blower (2003) conducted a study looking at the possibility of improving crash injury outcomes of large truck occupants by using suitable crash protection systems. They identified four objectives in conducting their research: 1) performance of a survey of the state of the art in truck occupant protection systems; 2) collection of truck crash data from US road system and analyze; 3) development of truck crash simulation models and occupant injury models; 4) and quantitative analysis potential benefits of various occupant protection countermeasures (Krishnaswami & Blower, 2003).

The publicly available crash data is reviewed to identify the key factors associated with truck driver injury (Krishnaswami & Blower, 2003). Previously, little research has been performed on crashworthiness of trucks or injury devices for drivers in trucking accidents. The crash data does not provide the much-needed descriptive injury information of the drivers involved in the accident. At the time of their study, data was still being collected for the LTCCS and was not yet available.

Although the LTCCS data set was not utilized for their study, it is apparent that once it was compiled, the detailed accounts of reported truck accidents would supply much-needed information for in-depth research on future countermeasures and causal factors.

Johnson (2008) investigated human factors issues related to the possible use of lane departure warning systems (LDWS) to reduce side collision and run-off-road crashes. Lane departures are classified as either intentional (e.g., to overtake slower vehicles or obstacle avoidance) or unintentional (driver fatigue, distraction or inattention) (Johnson, 2008). The LTCCS database is used to determine the different types of accidents, such as roadway departures and inattention, which could be affected by using LDWS. The dataset from the LTCCS, coupled with safety records from eight large commercial trucking fleets, provides data for his study.

Johnson concluded that the frequency of lane departure and run-off-road accidents is low but the consequences of these incidents are relatively high. Also, he found that the relative frequency of lane departure mishaps varied between carrier companies. This

indicated that the implementation of LDWS or a variation of LDWS is dependent on the individual company's operational experience and familiarity with the system.

Another study of heavy truck accidents, conducted by Kharrazi and Thomson (2008), used the LTCCS dataset to determine if common maneuvers that cause a loss of vehicular control can be identified. Truck accidents were analyzed in relation to the type of accident, loss of control, critical maneuver, vehicle combination and different types of road characteristics (Kharrazi & Thomson, 2008). According to their findings, large trucks are involved in 12% of all fatalities reported in 2004 while accounting for only 7% of total miles driven (2008). Approximately 20% of large truck accidents are caused by loss of control, which is considered to be drastically reduced by using an Electronic Stability Control, active steering or further integration of braking and steering (Kharrazi & Thomson, 2008).

Their findings were summarized as follows: 1) loss of control is associated with 19% of trucks involved in accidents; 2) vehicle roll-over is more common than yaw instability (55% of trucks lost control and rolled-over, 31% incurred yaw instability and 14% experienced both); 3) negotiating a curve is the key critical maneuver that led to a loss of control (59%), next is avoidance maneuver (11%) and road edge recovery (11%); 4) preventing yaw instability could potentially lead to about a 20% reduction in rollover mishaps (Kharrazi & Thomson, 2008).

These studies either have been supported by or have addressed the requirement for the data that is only available in the LTCCS. The elements of data that is collected in the LTCCS are available in several forms. Some of the different reports include general description and report of the accidents, diagrams of the mishap scene, photographs of the vehicles involved, interviews from persons involved in the accident as well as non-motorists, driver's assessment of the accident (including the events that led up to the mishap and driver behavior), and information gathered from the large truck carrier companies regarding driver history and safety records (McKnight, 2004).

Investigative analysis develops causal inferences through the collection and analysis of the facts that are reported. The validity of these inferences is dependent on

the amount and accuracy of the data (McKnight, 2004). Some of the accidents contain missing or incomplete information while others are gathered from personal interviews. McKnight goes on to state that the confidence in the assumptions as to human factors causes will differ with the amount of and validity of relevant information (2004).

F. SLEEP, ACTIVITY, FATIGUE, AND TASK EFFECTIVENESS (SAFTE) MODEL

The SAFTE Model has been under development by Dr. Steven R. Hursh for more than a decade. The model has been developed for the Department of Defense as a means of identifying performance problems and developing a favorable operational planning schedule based on hypothetical work-rest-sleep schedules (Hursh et al., 2004). SAFTE is able to produce predicted levels of effectiveness of an individual given a past, present and future work-rest-sleep assignment.

The SAFTE Model is based on sleep and circadian rhythm research collected over a 20-year period (Hursh et al., 2003). SAFTE is designed to take into account a wide range of scheduling and sleep conditions over any given time period and produce valid performance predictions (Hursh et al., 2003). The model incorporates quantitative information about circadian rhythms, cognitive performance (recovery rates associated with sleep and decreasing rates associated with wakefulness), and cognitive performance related to sleep inertia. These elements produce a three-process model of human cognitive effectiveness. The latest version of SAFTE predicts of performance under a wide range of scheduling conditions.

The model is homeostatic and modifies its predictions of future effectiveness based on current sleep history. A circadian process influences both performance and sleep management. Sleep management depends on hours of wakefulness, hours of sleep achieved, current sleep debt, circadian process, and sleep quality. Performance or cognitive effectiveness is dependent on circadian process as well as the balance between quality sleep and sleep debt, and sleep inertia (Hursh and Eddy, 2005). Managers or unit leaders can use the fatigue management system to predict the onset of fatigue, reduce

operator error due to fatigue, and to improve operator safety and effectiveness. A general design of the SAFTE Model is shown in Figure 1.

The model begins with the 'SLEEP RESERVOIR' box. The lines within the block signify levels of sleep within the reservoir. The lowest point, or trough, shows that the reservoir is completely empty, while at its peak, the reservoir is at capacity. The sleep reservoir is replenished during sleep periods and exhausted during hours of wakefulness. The rate at which the reservoir is filled depends on the quality of sleep and sleep intensity. The sleep intensity node is modeled as a function of the time of day (circadian process) and current level of the sleep reservoir. Sleep quality is determined by various external factors such as performance requirements and real-world demands (Hursh et al., 2004). Performance use is the output of the above model.
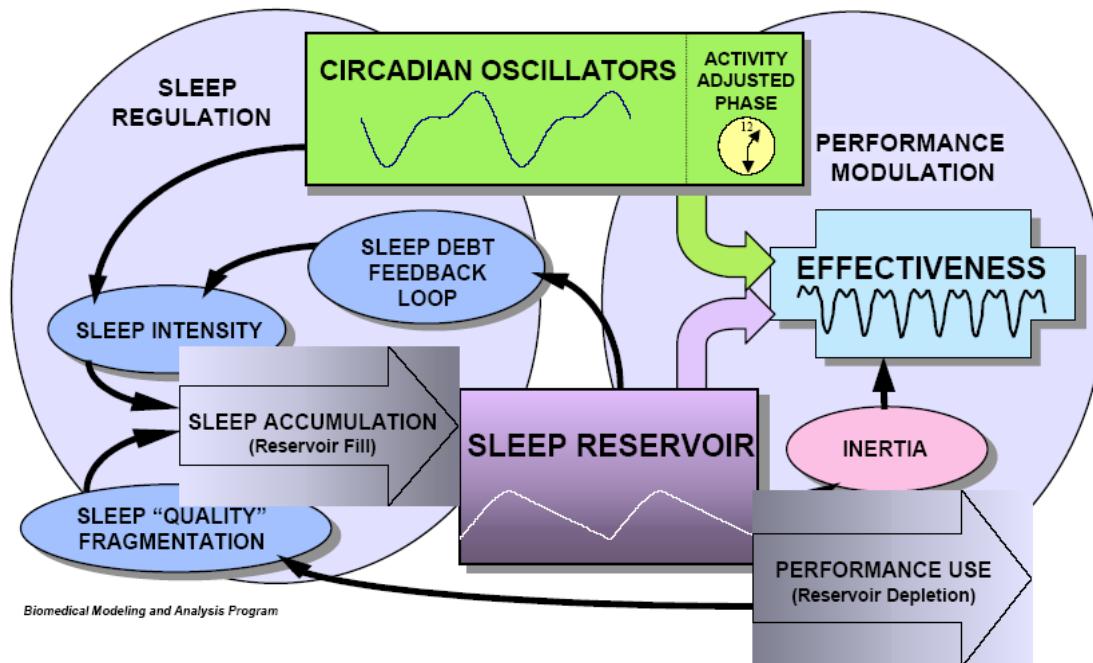

Figure 1.   SAFTE Model (From: Hursh, 2003)

14

Current models of sleep and performance have limitations and the SAFTE Model is no exception. Hursh et al. (2004) list two features that are not provided by the SAFTE Model. SAFTE does not:

- Offer an estimation of group or unit variation on an average performance prediction

- Integrate individual characteristics such as age, time of day, or sleep requirements for full performance.

These individual attributes or limitations may not be important if the purpose of the model is to predict the mean group or unit performance or to develop a basic work-rest-sleep schedule that will be used by the group or unit. In these instances, an ordinal prediction is adequate in determining which alternative schedule should be used or whether the average performance at a future time is expected to fall below acceptable levels (Hursh et al., 2004).

The performance of the model is dependent upon the data used to establish a sleep history prior to the time of the prediction of effectiveness. Also, fatigue models must assume some performance level or measurement as a standard to make a prediction. The SAFTE Model uses two sets of parameters to predict an individual's alertness and vigilance or psychomotor vigilance test (cognitive throughput) (Hursh et al., 2004).

All models are either accepted or rejected by their ability to make informative predictions to the user(s). Perhaps the greatest challenge to overcome in fatigue modeling is how to minimize the difference between the controlled environment and real-world operations (Hursh et al., 2004).

1. Validation of the SAFTE Model

Validation means a model must be a predictor of performance effectiveness. FAST, based on the SAFTE Model, takes in a realistic representation of the circadian process. Again, the circadian process manages the amount of sleep as a function of time of day and also takes into consideration the effects of sleep inertia. To validate the model, testing provided empirically derived data with remarkable predictive accuracy

(Hursh et al., 2003). Based on its predictive accuracy and performance SAFTE was selected from several competing models. In 2002, SAFTE was independently compared to six models from around the world and judged to have the least error (Fatigue and Performance Workshop, 2002). DoD has implemented SAFTE as the model of choice for predicting levels of effectiveness based on fatigue related impairment (Hursh et al., 2004).

G.   FATIGUE AVOIDANCE SCHEDULING TOOL (FAST)

FAST is an instantiation of the SAFTE Model. FAST uses the SAFTE Model to estimate a person's predicted effectiveness. The main purpose of FAST is to create a user-friendly computerized tool for mission or operational planners/schedulers. Based on time of day and amount of sleep an individual receives prior to and during the prescribed time period, FAST provides the user with a predicted level of performance effectiveness (Hursh et al., 2003). FAST is useful for assessing work and sleep schedules to determine if there will be any foreseeable problems with the schedule. Figure 2 features the graphic user interface (GUI).

FAST can also be used by the scheduler to insert a nap or rest period to boost predicted performance effectiveness or to calculate a predicted level of effectiveness if sleep duration is increased. The ability to plan for rest provides the planners with the capability to achieve higher levels of performance given limited opportunity to rest or combating interrupted sleep. In the end, more efficient work-rest-sleep schedules can be developed based on average or predicted effectiveness scores for upcoming or planned operations.

The software is user friendly and has the familiarity of Microsoft Windows software. FAST provides the user with a large number of settings, from display options to setup features. FAST is dependent upon user input to provide performance effectiveness levels. The user or scheduler can choose a time period as short as six hours to one as long as thirty days. FAST creates a graphical display depicting performance effectiveness over the set time period.

The green (upper) portion of the chart shows the performance level of an individual between 90–100% and represents the safe environment. The yellow sector, or caution area, indicates an individual's predicted effectiveness level between 65–90%. Lastly, the red section, or danger zone, indicates an individual's predicted effectiveness level below 65%. The green-yellow-red color scheme is a familiar that is readily interpretable.

Dates are presented near the top of the chart (in the white region) and the corresponding 24-hour time period is at the bottom of the chart. The blue and red bars at the bottom of the display indicate sleep and work scheduled throughout the day and night (Blue = Sleep and Red = Work). The scale along the left side of the chart shows the level of performance effectiveness (0 – 100%). The user has the ability to select one of four displays to overlay the chart and that option will be located on the far right side of the graph. The displays available include blood alcohol content (BAC) equivalence, lapse index, sleep reservoir, or acrophase (the time at which the peak of the circadian rhythm occurs).

One useful feature is the dashboard. The dashboard is an option that the user can either enable or disable. The dashboard provides the user with a "snapshot" of predicted performance effectiveness at any given time. The dashboard provides date/time information, level of effectiveness at any time of day, and five different fatigue indicators. The dashboard divides the data into two categories: performance and fatigue factors. The performance category displays five sub-categories showing physical aspects of fatigue (e.g., mean cognitive performance, effectiveness, lapse index, reaction time, and sleep reservoir). The fatigue factors category contains five sub-categories as well and includes sleep (past 24 hours), chronic sleep debt, hours of wakefulness, time of day, and out of phase. These features assist the user in analyzing and/or determining the decrease in predicted performance levels and possible countermeasures to employ.
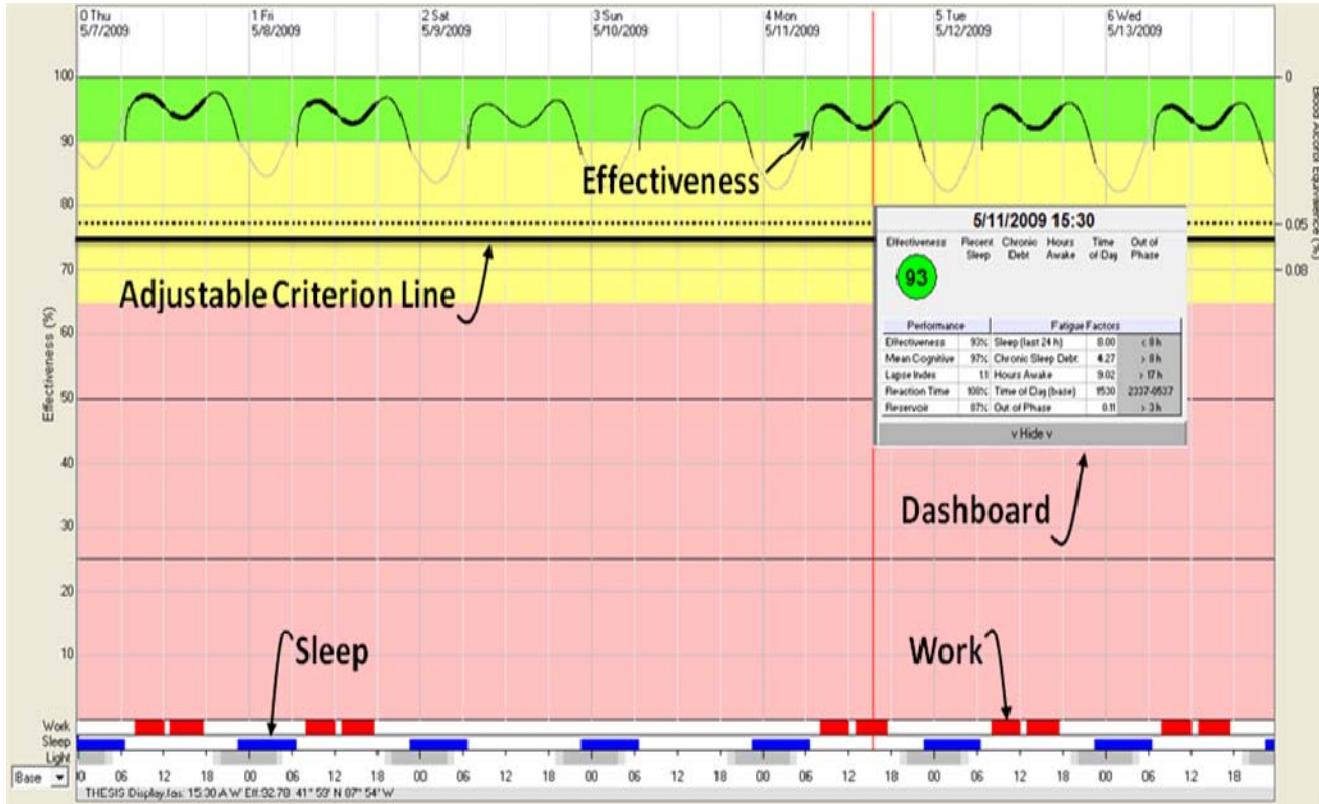
Figure 2.   FAST GUI Display

FAST is also useful for conducting retrospective analyses on fatigue-related incidents or mishaps (Hursh et al., 2003). Information on sleep habits, schedules, and sleep quality of the individual can be collected and entered into FAST. This information provides insight as to the level of effectiveness of the operator at the time the incident took place. With the combination of a documented sleep and event record, an analysis of the mutual effects of time of day and sleep history may be conducted.

1.   Limitations of FAST

As with all fatigue models and/or tools, there are limitations to FAST. First, FAST requires schedules (past, present, or future) to be input by the user. FAST will look at the entire time period and allow the user to define each 15-minute interval of time as either sleeping, working, or awake. Depending on the time period and number of people the user wants to evaluate, this can take several minutes to several hours.

FAST also allows the user to decide on the quality of sleep an individual receives (excellent, good, fair, or poor). Each sleep quality has a direct impact on the predicted performance effectiveness level. Possessing the capability to accurately depict sleep further adds to the quality of prediction by FAST because it captures as much realism as possible. For example, being underway or on deployment, there are several factors that can affect the amount or quality of sleep a person receives; loud noises, change of time zones, rough seas, and inadequate sleeping quarters can all have different affects on different people.

The second limitation is that FAST does not take into account any method or type of countermeasure. The use of caffeine, dextroamphetamine (Dexedrine), or sleeping pills (to assist in sleeping) are often used and readily available to mitigate fatigue; however, it is not a factor or variable available for use in FAST.

2.   Using FAST at the Operational Level

DoD believes FAST to be the most accurate and operationally practical fatigue model currently available to our armed services (Hursh et al., 2004). FAST allows for quick and easy entry of work-rest-sleep schedules to provide predicted levels of effectiveness. FAST also displays a graphical representation of performance over a projected period of time. FAST presents a truthful representation of expected performance that can be used to develop more effective and efficient work or training schedules.

FAST may also be used as an analysis tool of mishaps that are deemed to be caused by fatigue. Using FAST in this way, prior sleep-work-rest schedules are entered and a forecasted level of effectiveness is displayed based on the time period entered. The individual's level of effectiveness can then be assessed at the exact time of the mishap. Hursh et al combine this functionality of FAST with other findings of an investigation to conduct further analysis of the mixed effects of time of day and recent sleep history as contributing factors to safety-related events (2004).

THIS PAGE INTENTIONALLY LEFT BLANK

III.   METHODOLOGY

A.   INTRODUCTION

The purpose of the methodology chapter is to illustrate the steps and procedures applied to obtain, use and analyze the data contained in LTCCS datasets. The LTCCS contains a large amount of descriptive data that can be used for analysis to identify individual crash causation factors (Report to Congress on the Study of the LTCCS, 2006). The database is available for public use and offers colleges, universities, and individuals the capability to review the datasets to gain a better understanding of truck crash factors (Report to Congress on the Study of the LTCCS, 2006). The database is available in various formats such as Excel, SAS, or text files and can be downloaded from the following Web site: http://ai.fmcsa.dot.gov/ltccs/default.asp.

This chapter starts by describing the purpose for conducting large truck crash research and the methods used to obtain crash data. Next, the participants who are examined in this thesis are described and a demographic breakdown of these individuals is provided. A descriptive outline on the selection of the data (cases) considered in this thesis is presented. Once the data is selected, the method of their classification is offered. The method behind the classification gives the logic of the categorization of the critical reason variables pertaining to human factors. This chapter concludes with a short explanation of the study design implemented for this thesis and the statistical analysis that is conducted.

B.   METHOD

The LTCCS is a collaborative project between FMSCA and NHTSA.   The primary objective of the study was to conduct investigative and statistical analysis on large truck crashes around the nation. Prior to the LTCCS, no other motor vehicle crash database in the U.S. contained sufficient or detailed information on large truck crashes. Crash researchers, along with state truck inspectors, deployed to every feasible crash site as soon as possible after the accident occurred. An investigative approach develops

causal factors through the collection and analysis of the facts as to why the crash transpired (McKnight, 2004). Inferences as to the causes of large truck crashes are derived mainly from information discovered at the crash scene through the use of interviews and observations. Statistical analysis usually involves the comparison of the characteristics of people, things, or conditions involved in the crashes with control samples from the population at large that are similar to the cases except for the particular characteristic under examination (McKnight, 2004).

Based on experience from previous studies, investigative analysis is most successful in identifying the immediate contributors to crashes, whereas statistical analyses are important to identifying causal contributors (Hedlund, 2004).

Exhaustive efforts ensured reported crashes involved at least one large truck (Gross Vehicle Weight (GVW) >10,000 lbs.) (Hedlund et al., 2003). For each accident reported, data is collected on up to 1,000 factors and is entered into the study. Some of the datasets included an assessment of the driver's actions and behavior leading up to and during the mishap, condition of the truck and other vehicle(s) post accident, and road and environmental conditions (Craft, 2007).

The LTTCS database contains 1,070 crashes and 2,284 vehicle records. The LTCCS gathered detailed information regarding personal injuries of drivers and occupants, vehicular damage, driver assessment (including background investigations on drivers driving records), sleep history, and environmental and road conditions at the time of the accident than any other crash study in the U.S. The LTCCS database lends itself to a wide variety of study and analysis as to crash risk, crash causation, and crash countermeasures.

C. PARTICIPANTS

Participants are individuals involved in an accident and included in the LTCCS. There are a combined total of 3,014 drivers and occupants observed in the Occupants dataset within LTCCS database. However, only the truck drivers and drivers of the other

vehicles are examined for this analysis. This reduced the total number of observations to 2,258. Case identification numbers are then matched to their corresponding records in the DriverSleep dataset.

After matching case identifications between the two datasets, records that did not contain sufficient sleep history or have missing data are removed from further consideration. Any records that indicated a shift change are subsequently removed as well. The decision to remove all cases with incomplete or missing data, as well as the drivers enduring a shift change, is made to ensure an accurate and justifiable sleep history could be generated. After these cases are removed based on the previous criteria, there are 1,368 records still existing out of the 2,258 drivers reported (approximately 61% of the data still possible for use).

1. Human Factors, Non-human Factors and Other

Critical reason variables, defined as the immediate failure that led to that critical event, are established by the LTCCS (Craft, 2007). Critical reasons are comprised of driver condition and decision-making, roadway and environmental conditions, as well as vehicle failure(s). Driver critical reasons are designed to allow the user to label these reasons into four main categories:

- Non-performance (e.g., driver fell asleep, driver is overcome by heart attack or seizure)

- Recognition (driver inattention, distraction, poor situational awareness)

- Decision (driving too fast for given road/traffic conditions, following too closely)

- Performance (driver panic, poor directional control, overcompensation)

These critical reasons are used in the categorization of human-factors or nonhuman factors related accidents.

Critical reasons are described using brief descriptions detailing the wide range of variables. These coded variables, with their descriptions, are developed to assist with assessing causal factors that lead to the accident. The CrashAssessment Excel file

contained all the critical reason codes, with heading ACRReason, used in this analysis. A survey is created based on the CrashAssessment data to assist with categorizing critical reasons into three groups. These categories are human factors, nonhuman factors and other. The survey form is located in Appendix A.

The survey was distributed via email to six subject matter experts in a Human Systems Integration area of expertise. It was requested that they provide their professional assessment on categorizing each coded description (variable) that is presented in the survey. The participants of the survey are provided with only an Excel document containing the variables and short description of the critical reason codes. No other background information regarding the mishap, driver(s) or vehicle(s) is made available.

Prior to the survey being sent out, it was decided that there had to be at least 66.67% (4 out of 6) in agreement to successfully categorize the critical reason code. If a code did not receive the required 66.67% it was removed from further consideration. The results of the surveys are found in Appendix A. After collecting the surveys and categorizing all of the code descriptions (HF, NONHF or "Other") the related codes could be arranged and distributed to the appropriate grouping. The remaining HF, NONHF, and "Other" codes are examined and then reviewed in the CrashAssessment dataset.

The 1,368 cases that are remaining after assessing sleep history records and possible shift changes are examined strictly on their ACRReason codes. There is a total of 24 HF codes, 15 NONHF codes, and 4 "Other" codes. After paring down the data to the matching 1,368 cases in the CrashAssessment dataset, it is possible to sort the data based on ACRReason codes using the data filter function. Once each individual code is reviewed for each of the three categories, there is a combined total of 545 cases remaining (496 HF records and 49 NONHF records). All of the "Other" coded records are removed from further consideration. This left 545 cases out of a total 3,014 cases (or 2,258 drivers only) to be analyzed in this thesis.

D.   FORMATTING THE DATA

Additional Excel spreadsheets are developed by extracting case identification numbers (CaseID), last sleep start and last sleep end dates and times from the DriverSleep dataset, and the crash date and time from the Crash dataset.   The spreadsheet that is generated is necessary to develop a comma-delimited text file. The resulting text files are formatted using a C program. The code that is written is used to convert text files into the format that is necessary for input into FAST. Appendix G displays a copy of the code that is written and Appendix H illustrates proper format of the .txt file.

1.   Running the Data

Once this process is complete for all 545 cases, the data is manually imported, one by one, into FAST. FAST produces predicted levels of effectiveness based on reported work and sleep scheduels. FAST is also capable of presenting a graphical representation of levels of effectiveness over the period of time depicted in the work-sleep schedules. Predicted levels of effectiveness are generated from FAST for both HF and NONHF groups. These scores are recorded in Excel for follow-on statistical analysis.

2.   Study Design and Statistical Analysis

A case-control study design is implemented to verify if a statistically significant difference between the HF and NONHF populations. The goal is to observe if the data from the HF group produced different results from the NONHF group. If the results from the two groups are similar, it is a conclusion that fatigue does not act as a factor on large truck crashes. A K-S test is performed using S-Plus statistical software.

THIS PAGE INTENTIONALLY LEFT BLANK

IV. RESULTS

A. DEMOGRAPHICS

This study examines reported sleep patterns of drivers involved in large truck crashes. There are 545 cases; 496 are classified as human factors-related (HF) and 49 as non-human factors-related (NONHF). A demographic analysis of the drivers is provided in Figure 3. This detailed breakdown illustrates that there is a similar demographic make-up between HF and NONHF populations. This also shows that there is an equal representation of males and females, ethnic backgrounds, and all drivers over the age of 16. In HF and NONHF populations, males make up 85% and 82% of the drivers respectively and the median of the age of each population is 39. The standard deviation (SD) for the HF group is 14.0 while that of the NONHF group is 13.82.

Demographic Breakdown (HF and NONHF)

Gender

	HF	% of total population	NONHF	% of total population
Male	422	85.08%	40	81.63%
Female	74	14.92%	9	18.37%

RACE

	HF	% of total population	NONHF	% of total population
White	333	67.14%	28	57.14%
Black	71	14.31%	12	24.49%
Hispanic	69	13.91%	8	16.33%
American Indian	3	0.60%	0	0.00%
Asian	10	2.02%	1	2.04%
Other	5	1.01%	0	0.00%
Unknown	5	1.01%	0	0.00%

Age

	HF	NONHF
Average Age	39.33	39.20
Oldest	90.00	81.00
Youngest	16.00	18.00
Standard Deviation	14.00	13.82

Figure 3.   Demographic Breakdown (HF and NONHF)

B.   FAST RESULTS

The data is collected over the time period from April 2001 to December 2003 and contains a great deal of driver and crash information. The thesis focuses on the reported sleep and driving patterns in the 24-hour time period immediately prior to the mishap. Using driver accounts of their previous sleep period, which is their last reported sleep (sleep start and sleep end times), we established sleep-work schedules for implementation by FAST. FAST uses a preconditioning function that assumes each individual receives eight hours of excellent sleep for the three days prior to the first recorded (observed) day. As drivers may not have received this assumed eight hours of sleep and are not well rested prior to the accident, results would have been skewed if adjustments to preconditioning is not considered.   Figure 8 illustrates the HF and NONHF predicted levels of effectiveness. FAST assumes the individual received three consecutive days of excellent sleep. The distributions, shown in Figure 4, of HF and NONHF FAST scores are similar and are negatively skewed.
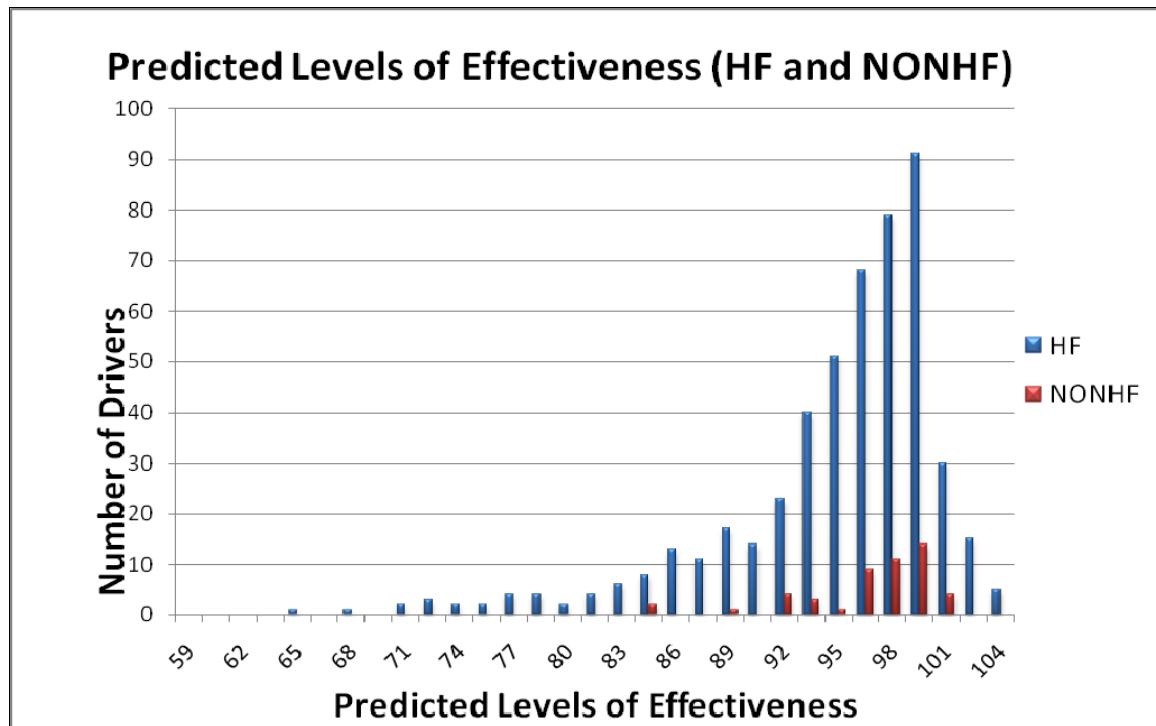
Figure 4.   FAST Results Producing Predicted Levels of Effectiveness for HF and NONHF Populations
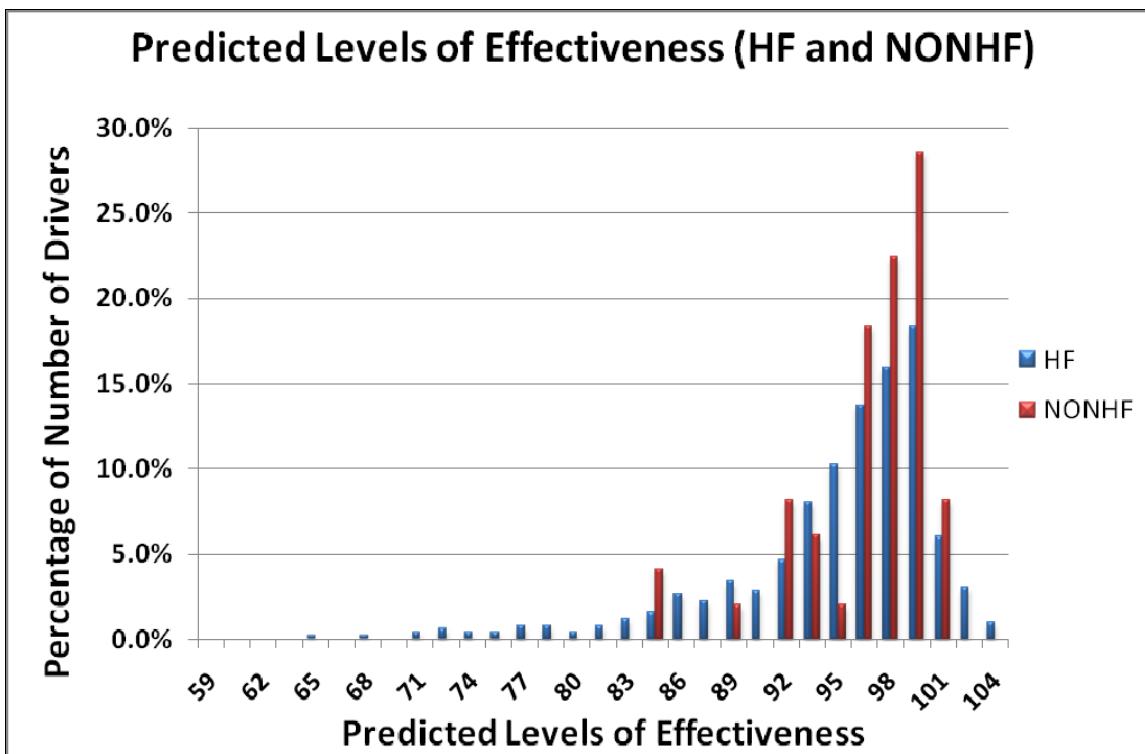
Figure 5.   FAST Results Producing Predicted Levels of Effectiveness for HF and
                NONHF Populations (Percentage Breakdown)

The mean level of predicted effectiveness for HF population is 94.184 (SD = 6.19). The mean level of predicted effectiveness of the NONHF population is 96.182 (SD = 3.76). The observed statistic is – 1.998 (mean (HF) – mean (NONHF)).

Box plots of Predicted Levels of Effectiveness
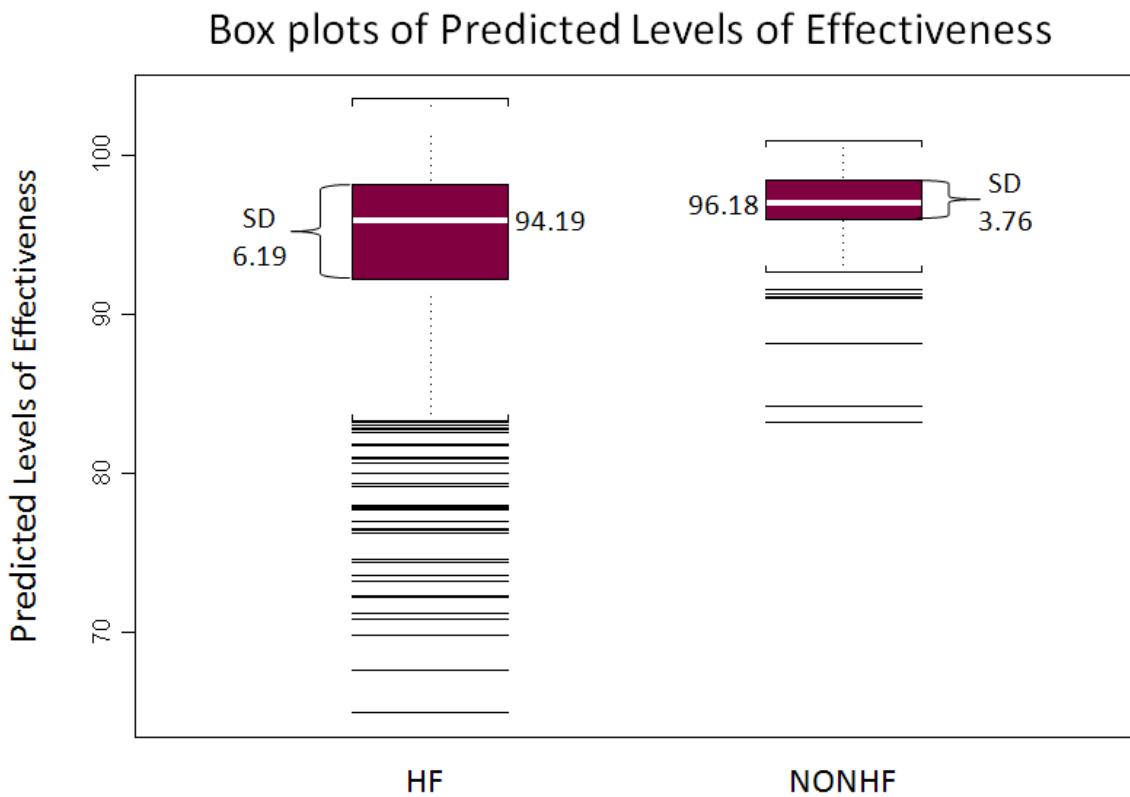
Figure 6.   Box plots of Predicted Levels of Effectiveness

C.   STATISTICAL RESULTS

Significance tests reveal whether an observed effect, such as a difference between two means, could reasonably occur by chance when selecting a random sample. If not, then there is evidence that the effect observed in the sample reflects an effect that is present in the population.

The null hypothesis, H_o, states or describes some aspect of the statistical behavior of a set of data. For this study, the null hypothesis is that there is no difference in the predicted levels of effectiveness between the HF and NONHF populations.   The alternative hypothesis is that there is a difference. This study uses a one-tail (or one-sided alternative) test of significance.

Comparing two samples raises a few arguments about using a permutation test versus traditional formula-based tests, such as the t-test. Hypotheses for the t-test are presented in terms of two population means:

$$H_o: \mu_{HF} - \mu_{NONHF} = 0$$
$$H_a: \mu_{HF} - \mu_{NONHF} < 0.$$

The null hypothesis states that average predicted levels of effectiveness are the same for both groups. The one-sided alternative proposes the HF group has a lower predicted level of effectiveness than that of the NONHF population.

Permutation tests are statistical significance tests in which a reference distribution is obtained by calculating all possible values of the test statistic under re-arrangements of the labels on the observed data (Good, 2000). Plainly stated, the way by which treatments are assigned to subjects in an experimental design is paralleled in the analysis of that design. Likewise, a permutation test involves the shuffling of observed data to determine how "unusual" an observed outcome is. If the labels or identifiers are transferable under the null hypothesis, then the resulting tests provide exact significance levels (Gordon, 2000).

Good (2007) lists several advantages to using a permutation test:

- Permutation tests can be performed for any statistic (so the approach permits the user to choose the test statistic best suited for the task at hand).
- Permutation tests can be used for analyzing unbalanced designs (i.e., mixtures of categorical, ordinal, and metric data).
- Permutation tests are flexible and robust when dealing with missing data and violations of assumptions.
- Permutation tests may reduce costs of experiments and surveys through sample size reduction.

For this thesis, there are 496 cases labeled as HF and 49 labeled as NONHF. For a permutation test, all observations of both HF and NONHF are combined. The data is then re-distributed into groups of the same sizes as the original HF and NONHF

observations. Using S-Plus, it is possible to perform such a test. This permutation test requires a short algorithm to re-sample the observed data. Appendix I illustrates the code used to perform the test.

After running the program several times, S-Plus calculates p-values ranging from 0.003 to 0.015. These p-values suggest that the observed value of -1.998 is very unlikely to have been seen in a world in which the HF and NONHF labels were assigned at random.

While a permutation test is the primary statistical test used for this thesis, there are several other tests that are used in this study. All of the following tests are conducted using S-Plus statistical software package: Pooled-Variance Two-Sample t-test; Kolmogorov-Smirnov (K-S) Goodness-of-Fit Test; Wilcoxon Rank-Sum Test.

Performing a t-test produced a $t(543) = -2.22$ with p-value = 0.0269. The null hypothesis is rejected in favor of the alternative hypothesis: difference in means does not equal 0. The Wilcoxon rank-sum test yielded $Z = -2.1049$ and a p-value = 0.0353. Again, the null hypothesis is rejected in favor of the alternative hypothesis: μ is not equal to 0.

Performing a K-S test assumes the null hypothesis to be that the data follow a specified distribution whereas the alternative hypothesis is that the data does not follow the specified distribution. A two-sample K-S test gives a $ks = 0.2571$ and a p-value = 0.0032. This K-S test determines that the two datasets differ significantly. Again, reject the null hypothesis and accept the alternative hypothesis: the cumulative distribution function (cdf) of the HF population does not equal the cdf of the NONHF population for at least one sample point.

V. CONCLUSION AND RECOMMENDATIONS

A. CONCLUSIONS

Instability around the world places a great demand upon today's military forces. With a wide range of military operations, such as humanitarian assistance and disaster relief, non-combatant evacuation operations, patrols with host nation forces, and green water operations in the littorals, our U.S. service members must be properly trained and ready to deploy at a moment's notice. Efficient training, operating, and deployment plans must be developed with due consideration for sleep-work schedules.

This is the first study using large sets of data on driver sleep history extracted from LTCCS. As developed driver sleep-work schedules are input into FAST, predicted levels of effectiveness of the drivers at the precise time of the accident are determined. This thesis shows that at a level of effectiveness of 90.00, only 81.7% of all HF drivers have a predicted level of effectiveness equal to or higher while 93.9% of all NONHF drivers have a predicted level of effectiveness equal to or higher than 90.00. Only 3% of the entire observations, HF and NONHF, are below a predicted level of effectiveness of 77%. Hursh et al. (2006) show that there is visible degradation in human performance when FAST scores drop to levels of effectiveness of 77%. The percentage of drivers who are shown to be below 90.00 in the HF group is 18.3% while there is only 6.1% in the NONHF population.

There is a significant difference between mean predicted levels of effectiveness of the drivers in the HF and NONHF mishaps. However, there are limitations to this thesis and data collection process. One limitation stems from the LTCCS database itself and the validity of reported sleep history. All sleep histories are used in good faith and it is accepted that the reported sleep is the actual amount of sleep the drivers received in the 24 hours prior to the mishap. Another limitation is the exclusion of all of the drivers who underwent a shift change. Excluding these drivers whose sleep estimates indicated that they were experiencing a shift change was a conservative approach but may have affected the results. Had these drivers' sleep estimates been used these observations would have

produced considerably lower levels of effectiveness, further confirming the association of sleep patterns and accidents as well as underestimating the effects of fatigue on shift work.

B.   RECOMMENDATIONS

Effects of fatigue on operator performance are evident and fatigue has been found to be a causal factor in HF related accidents. FAST is a useful tool that can assist in the development of sleep-work-rest schedules. FAST can aid in the reduction of fatigue related accidents or as an investigative tool to provide insight to why the mishap might have occurred. It is recommended that FAST be introduced or integrated into military operations departments to assist in the planning phases or development of training and operating schedules. This would allow the schedule writers to input schedules into FAST to gain insight as to possible ORM issues dealing with fatigue. It would provide commanders with developing and assessing operational requirements based on levels of effectiveness of his troops.

It is also recommended that the LTCCS dataset be revisited with an attempt to develop all of the sleep histories on record. Close examination of the sleep histories that were omitted or those observations that were deleted could allow investigators to generate valid and justifiable sleep histories. Follow-on studies are needed to gain a better understanding of the role of fatigue on operator performance and the association of fatigue in large truck accidents.

LIST OF REFERENCES

Bootstrap Methods and Permutations Tests. Retrieved June 8, 2009, from
http://bcs.whfreeman.com/ips5e/content/cat_080/pdf/moore14.pdf

British Medical Association. (2000). Retrieved May 15, 2009, from *Implications for health and safety of junior doctor's working arrangements:* http://www.bma.org.uk/ni/healthcare_policy/working_hours_conditions/implicationsforhealthandsafetyofjuniordoctorsworkingarrangements.jsp?page=3

Colten, Harvey R. and Altevogt Bruce M. (2006). Sleep Disorders and Sleep Deprivation: An Unmet Public Health Problem. The National Academies Press Washington, DC.

Craft, Ralph. (2007). Analysis Brief: LTCCS Summary Federal Motor Carrier Safety Administration Office of Research and Analysis Publication No. FMCSA-RRA-07-017 Retrieved June 14, 2009, from: http://www.fmcsa.dot.gov/facts-research/research-technology/analysis/FMCSA-RRA-07-017.htm

Davenport, Nicholas. (2007). Fatigue – A Root Cause. *Approach Magazine*. p. 4-5.

Drake, Jack. (2008). Mitigating Flight Crew Fatigue. *Rotor Magazine.* p. 8-10.

Dinges, David F. (1995) Performance Effects of Fatigue. *Fatigue symposium proceedings/National Transportation Safety Board and NASA Ames Research Center*.

Drummond, Sean P. (2007). Effects of Total Sleep Deprivation and Recovery Sleep on Cognitive Performance and Brain Function. Accession Number: ADA474587.

Federal Motor Carrier Safety Administration Interstate Truck Driver's Guide to Hours or Service. Retrieved May 31, 2009, from http://www.fmcsa.dot.gov/rules-regulations/truck/driver/hos/fmcsa-guide-to-hos.pdf

Good, Philip. (2000). Permutation Tests: A Practical Guide to Resampling Methods for Testing Hypothesis. 2nd edition. Springer, New York.

Hedlund, James. *Statistical Analyses of Large Truck Crash Causation Study Data* (Committee for Review of the Federal Motor Carrier Safety Administration's Truck Crash Causation Study, Letter Report of September 4, 2003, pp. 16-49). Retrieved May 22, 2009, from: http://gulliver.trb.org/publications/reports/tccs_sept_2003.pdf.

Hedlund, James and Blower, Daniel. (2006). Large Truck Crash Causation Study (LTCCS) Analysis Series: Using LTCCS Data for Statistical Analyses of Crash Risk. Retrieved June 14, 2009, from http://www.fmcsa.dot.gov/facts-research/research-technology/analysis/ltccs.htm

Hursh, S.R., Raslear, T.G., Kaye, A.S., Fanzone, J.F. (2006). Validation and Calibration of a Fatigue Assessment Tool for Railroad Work Schedules, Summary Report. Department of Transpiration, summary report, October 2006. DOT/FRA/ORD-06/21. Retrieved June 14, 2009, from: http://www.trb.org/news/blurb_detail.asp?id=7059

Hursh, S.R., Eddy, D., and Elsmore, T. (June 2004). The Fatigue Avoidance Scheduling Tool: Modeling to Minimize the Effects of Fatigue on Cognitive Performance. Document Number 2004-01-2151. Retrieved June 14, 2009, from: http://www.sae.org/technical/papers/2004-01-2151

Hursh, S. R., Redmond, D. P., Johnson, M.L., Thorne, D.R., Belenky, G., Balkin, T.J., Storm, W.F., Miller, J.C., Eddy, D.R. (2004). Fatigue Models for Applied Research in Warfighting. Aviation, Space, and Environmental Medicine. Vol 00, No. 0, Month 2004. Retrieved June 14, 2009, from: http://mljohnson.pharm.virginia.edu/pdfs/301.pdf

Hursh, S.R., Redmond, D.P., Johnson, M.L., Thorne, D.R., Belenky, G., Balkin, T.J., Miller, J.C., Eddy, D.R. and Storm, W.F. (2003). The DOD Sleep, Activity, Fatigue and Task Effectiveness Model. Proceedings of the 12th Conference on Behavior Representation in Modeling and Simulation (BRIMS 2003), 03-BRIMS-001, 14-15 May 2003, Mesa, Arizona.

Johnson, Gordon. (2009). Driver Fatigue. Retrieved May 31, 2009, from http://semi-accident.com/fatigue.html

Johnson, Steven L. (2008). Human Factors Study of Driver Assistance Systems to Reduce land Departures and Side Collision Accidents. Department of Industrial Engineering. University of Arkansas.

Kharrazi, Sogol and Thomson, Robert. (2008). Study of heavy truck accidents with focus on maneuvers causing loss of control. Department of Applied Mathematics. Chalmers University of Technology. Retrieved June 14, 2009, from: http://inderscience.metapress.com/media/egpptmqqxrj7dyvnyd9l/contributions/7/4/8/0/7480533t4t72nk73.pdf

Knipling, Ronald R. (2004). *Evidence and Dimensions of Commercial Driver Differential Crash Risk.* Proceedings of the Third International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design.

Krishnaswami, Vasanth and Blower, Daniel. (2003). Feasibility of Heavy Truck Occupant Protection Measures. Transportation Research Institute. University of Michigan. Retrieved June 14, 2009, from: http://deepblue.lib.umich.edu/bitstream/2027.42/1544/2/97646.0001.001.pdf

Mara, J.R. (1998). Drowsy Driving and Automobile Crashes. Sponsored by the National Center on Sleep Disorders Research / National Highway Traffic Safety Administration. Retrieved June 14, 2009, from: http://www.nhtsa.dot.gov/people/injury/drowsy_driving1/drowsy.html#NCSDR/NHTSA

McCallum, Marvin, Sanquist Thomas, Mitler M., and Krueger, G. (2003). Commercial transportation operator fatigue management reference. Transaction Agreement No. DTRS56-01-T-003. U.S. Department of Transportation Research and Special Programs Administration.

McKnight, James. (2004). *Investigative Analysis of Large Truck Accident Causation* (Committee for Review of the Federal Motor Carrier Safety Administration's Truck Crash Causation Study, Letter Report of September 4, 2003, pp. 50-66), Retrieved June 5, 2009, from http://gulliver.trb.org/publications/reports/tccs_sept_2003.pdf.

Miller, James C. and Eddy, Douglas R. (2008). Operational Risk Management of Fatigue Effects II.

Miller, N.L., Matsangas, P., and Shattuck L.G. (2007). "Fatigue and its Effect on Performance in Military Environments." In P.A. Hancock and J.L. Szalma (Eds.), *Performance Under Stress*. Ashgate Publications.

Mitler, M.M., Miller, J.C., Lipsitz, J.J., Walsh, J.K., and Wylie, C.D. (1997). "The Sleep of Long-Haul Truck Drivers." The New England Journal of Medicine. Volume 337:755-762, Number 11.

Moore, R.T., Kaprielian, R., and Auerbach, J. (2009). "Asleep at the Wheel" Report of the Special Commission on Drowsy Driving. Retrieved May 30, 2009, from http://science-education.nih.gov/supplements/nih3/sleep/guide/info-sleep.htm

Paquette, Brian. (2007). Estimating fatigue in drivers involved in large truck crashes, analysis of the LTCCS dataset. University of South Florida.

Rosekind, M.R., Gander, P.H., Gregory, K.B., Smith, R.M., Miller, D.L., Oyung, R., Webbon, L.L., Johnson J.M. (1996). Managing Fatigue in operational Settings 1: Physiological Considerations and Countermeasures.

Saisan, J., de Benedictis, J., Barston, S. and Segal, R. (2008)
Sleeping Well; What You Need To Know: Sleep Requirements, Needs, Cycles
and Stages. Retrieved May 30, 2009, from
http://www.helpguide.org/life/sleeping.htm

U.S. Department of Transportation. (2006). "Report to Congress on the Large Truck
Crash Causation Study." Federal Motor Carrier Safety Administration. MC-
R/MC-RRA.

# APPENDIX A.   SURVEY OF CRITICAL REASONS CATEGORIZATION AND RESULTS

ACRReason

HF	NON-HF	OTHER	N	PERCENT	CODE	CAUSE
X			N	X.X	X	Driver experienced microsleep
6			43	1.9	111	Internal distraction
6			28	1.2	112	External distraction
6			138	6	113	Inadequate surveillance (e.g., failed to look, looked but did not)
6			3	0.1	118	Other recognition error (Specify):
6			69	3	120	Too fast for conditions to be able to respond to unexpected
6			2	0.1	121	Too slow for traffic stream
6			36	1.6	122	Misjudgment of gap or other's speed
6			34	1.5	123	Following too closely to respond to unexpected actions
6			20	0.9	138	Other decision error (Specify):
6			3	0.1	139	Unknown decision error
5	1		59	2.6	100	Sleep (acutally asleep)
5	1		53	2.3	119	Unknown recognition error
5	1		22	1	124	False assumption of other road user's actions
5	1		44	1.9	125	Illegal maneuver
5	1		6	0.3	127	Inadequate evasive action (e.g. braking only, not braking)
5	1		23	1	128	Aggressive driving behavior
5	1		63	2.8	140	Too fast for curve/turn
5	1		1	0	141	Panic/Freezing
5	1		40	1.8	142	Over compensation
5	1		38	1.7	143	Poor directional control (e.g.failing to control vehicle with skill)
5		1	59	2.6	110	Inattention (i.e. daydreaming)
5		1	2	0.1	149	Unknown performance error
4	2		40	1.8	101	Heart attack or other physical impairment of the ability to act
4	2		1	0	500	Signs/signals missing
3	1	2	98	4.3	199	Type of driver error unknown
3	3		1	0	505	Road design - roadway geometry (e.g. ramp curvature)
3	3		1	0	507	Road design - other

HF	NON-HF	OTHER	N	PERCENT	CODE	CAUSE
2	3	1	17	0.7	203	Cargo shifted
2	3	1	1	0	504	View obstructed by other vehicles
2	4		1	0	522	Fog
2	4		2	0.1	523	Windgust
2	4		7	0.3	530	Glare
1	5		16	0.7	200	Tires/wheels failed
1	5		8	0.4	201	Brakes failed
1	5		1	0	202	Steering failed
1	5		1	0	204	Trailer attachment failed
1	5		2	0.1	205	Suspension failed
1	5		1	0	208	Car body, doors, and/or hood failed
1	5		20	0.9	509	Slick roads (low friction road surface due to ice, loose debris, anything)
1	5		3	0.1	541	Transmission/engine failure
1	4	1	2	0.1	298	Other vehicle failure (Specify):
1	4	1	1	0	299	Unknown vehicle failures
	5	1	1	0	518	Other highway - related condition (Specify):
	5		18	0.8	540	Degraded braking capability
	2	4	17	0.7	999	Unknown reason for critical event
		6	1,229	53.8	0	Critical event not coded to this vehicle
		6	3	0.1	108	Other critical non-performance (Specify):
		6	6	0.3	109	Unknown critical non-performance

Results from CriticalReasons

Number of Questions

HF - NONHF - OTHER

Frequency
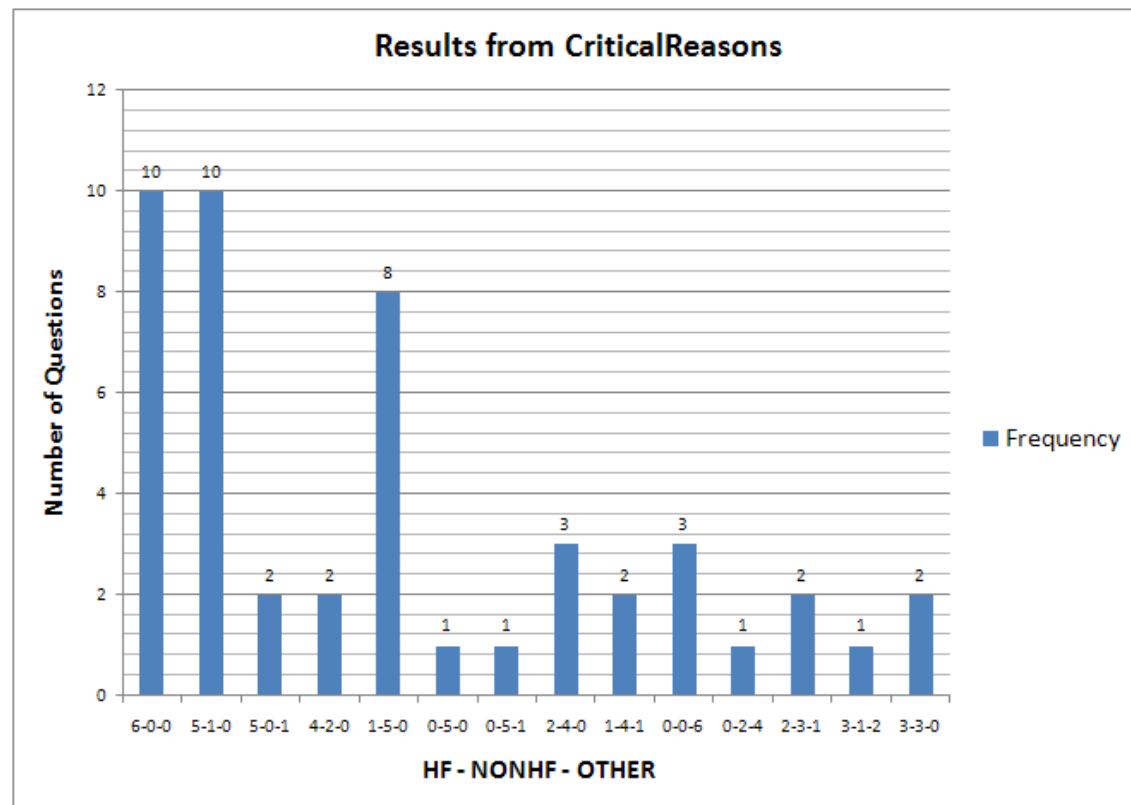
THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX B. DISTRIBUTION OF DRIVERS BY GENDER AND BY GROUP (HF AND NONHF)


Distribution of Drivers by Gender
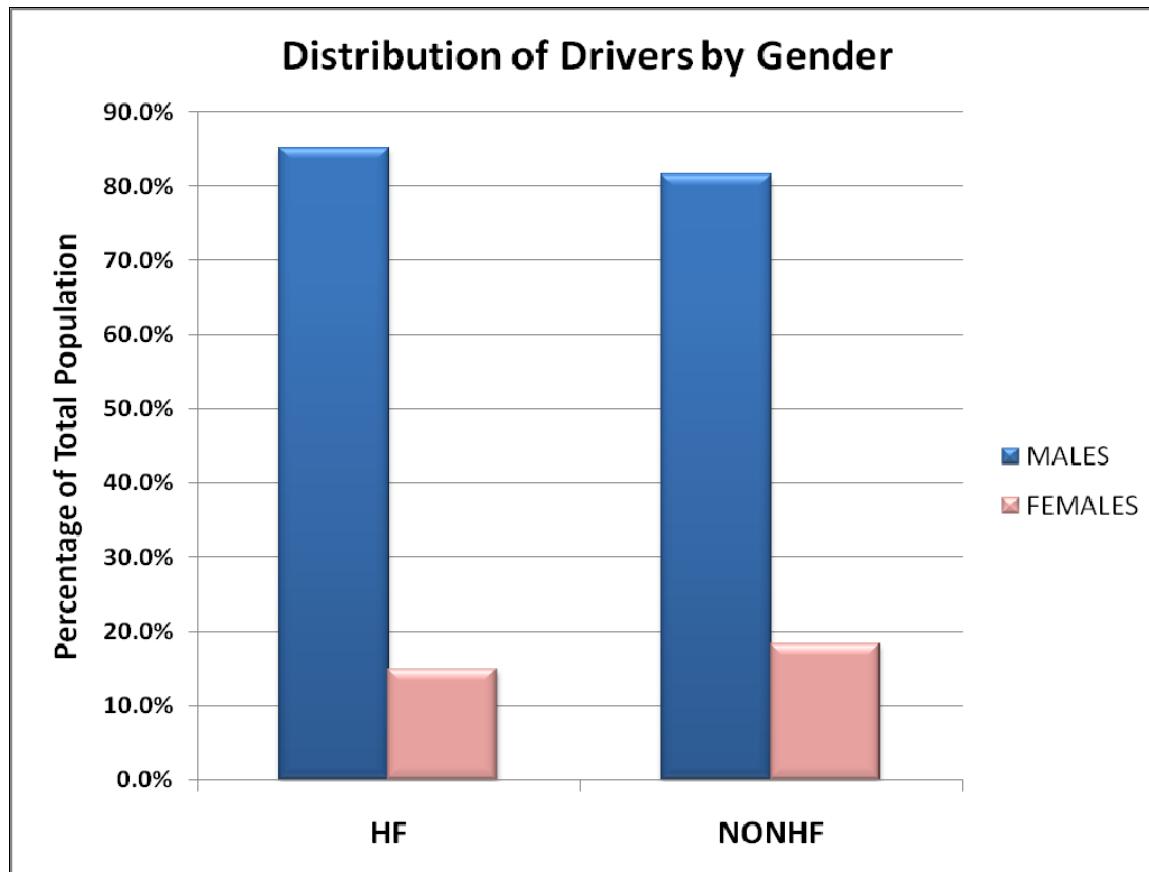
THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX C.   BREAKDOWN OF DRIVERS BY ETHNICITY FOR HF AND NONHF POPULATTIONS
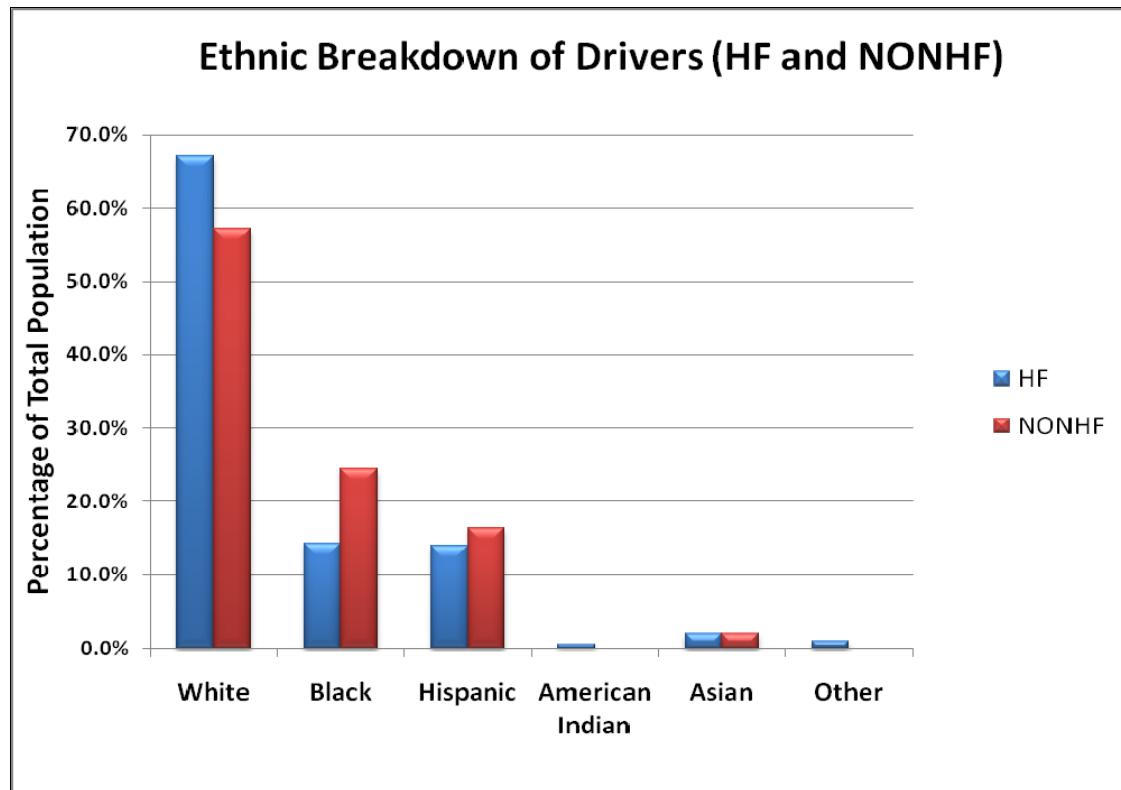
Ethnic Breakdown of Drivers (HF and NONHF)

THIS PAGE INTENTIONALLY LEFT BLANK

Driver Ages (HF and NONHF)

Driver Ages (HF and NONHF)
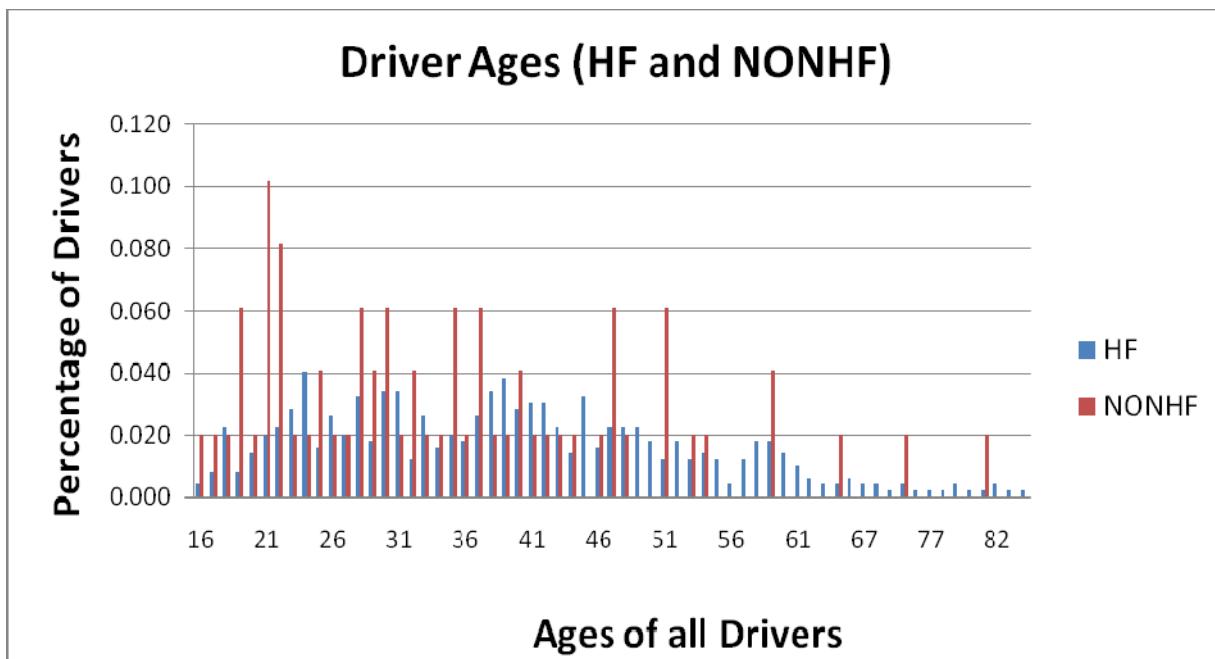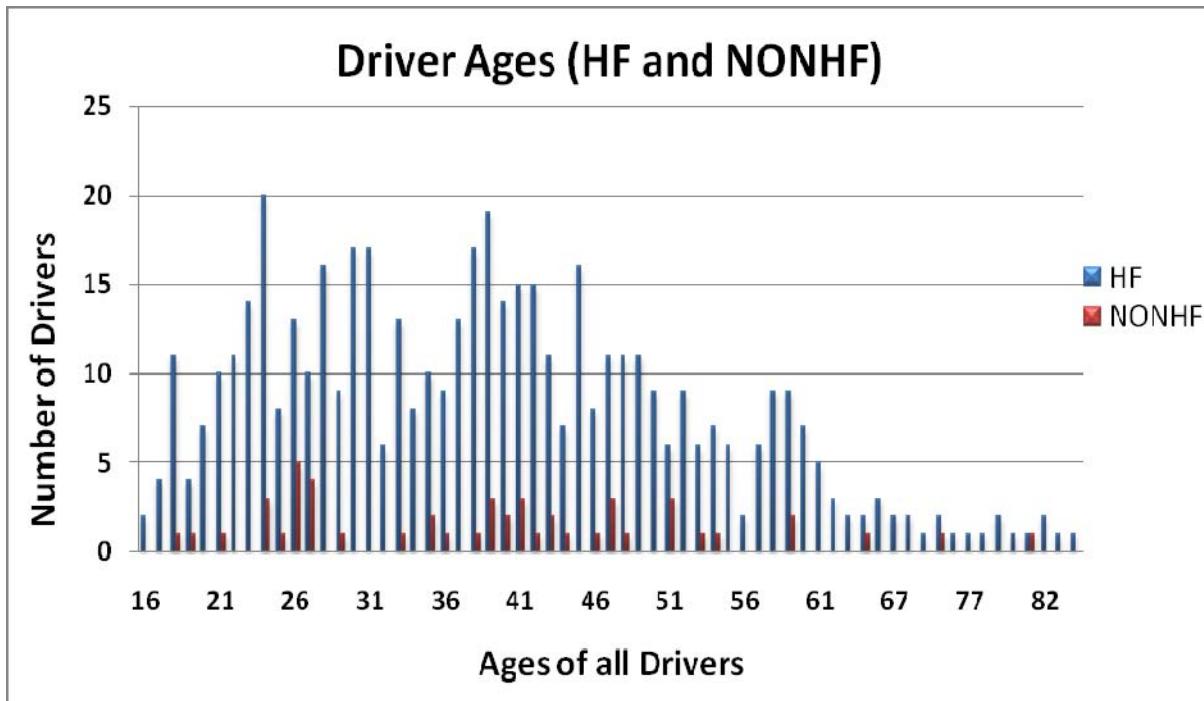
THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX E.   BOX PLOTS OF ALL DRIVERS BY AGE AND GROUP (HF AND NONHF)

Box plots of Ages of All Drivers (HF and NONHF)
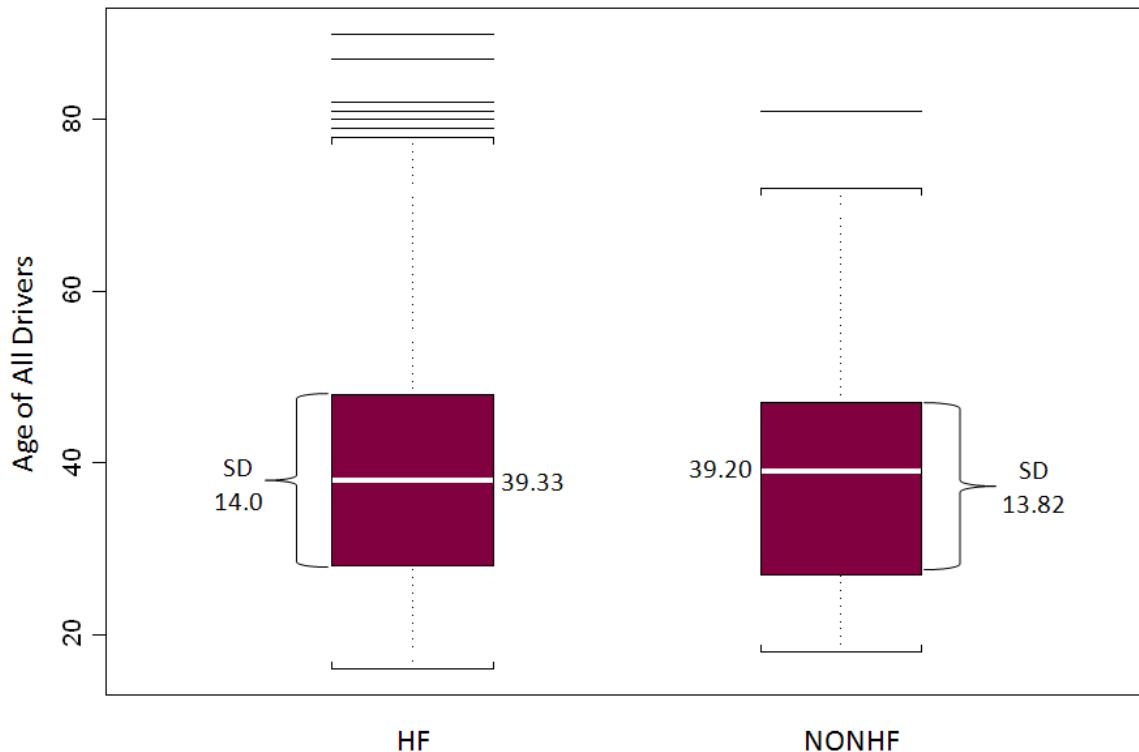
THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX F. HISTOGRAMS OF HF AND NONHF PREDICTED LEVELS OF EFFECTIVENESS

Predicted Level of Effectivness
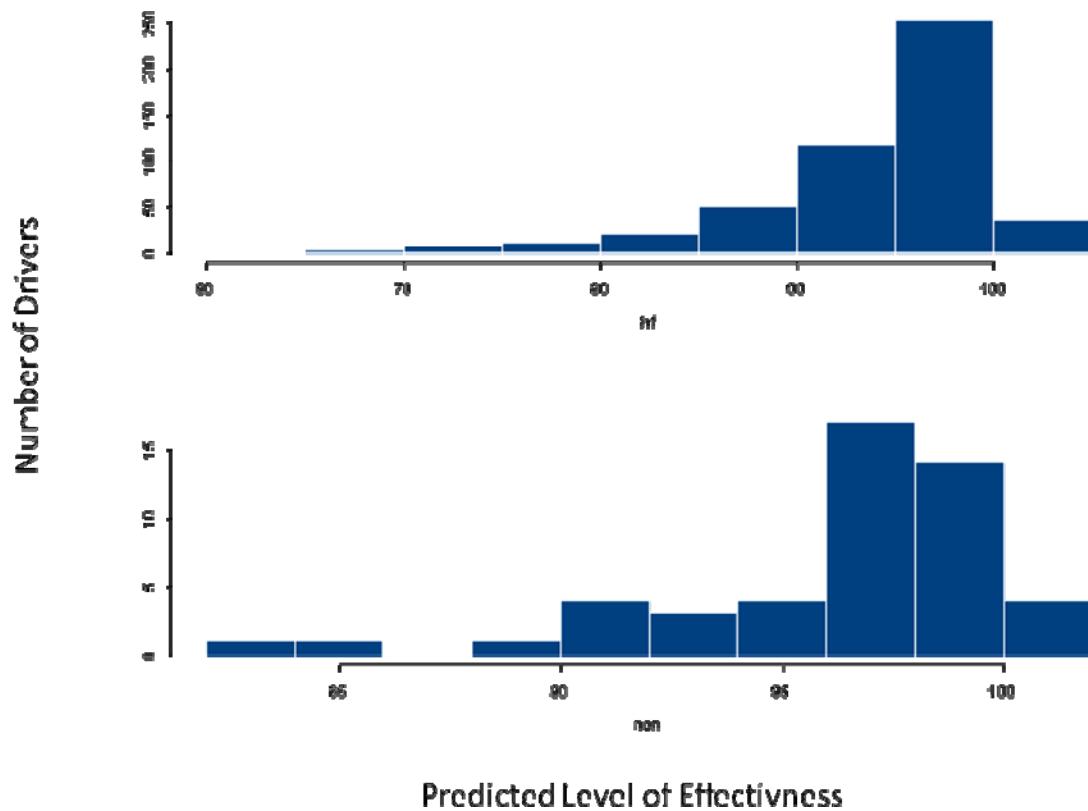
THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX G. C++ CODE FOR CONVERSION OF EXCEL FILES INTO FAST READABLE TEXT FILES

```
#Author J.C. Miller
#include <stdio.h>
#include <string.h>
#include <stdlib.h>.5
FILE *input_file;
char input_name[64];
char buffer[128], subj_no[16], start_date[16], start[16], stop[16], crash[16];
int i, n;
void Make_txt ( char subj[16], char day[16], char begin[16], char end[16], char work[16]);
int main()
{
printf ( "\nFASTinput app\nDr. James C. Miller\n\n" );
/* open custom text file made from custom Excel file */
strcpy ( input_name, "/Users/jamesandjoymiller/Desktop/HF.txt" );
input_file = fopen ( input_name, "r");
if ( input_file == NULL )
{
puts( "*** Can't find raw data file ***");
exit(1);
}
/* count the number of lines in the input file */
while ( !feof ( input_file ))
{
fgets ( buffer, 128, input_file );
n++;
}
printf ( "lines = %d\n", n );
rewind ( input_file );
/* get data from the input file */
for ( i = 0; i < n-1; i++ )
{
fgets ( buffer, 128, input_file );
strcpy (subj_no, strtok ( buffer, "," ));
strcpy (start_date, strtok ( NULL, "," ));
strcpy (start, strtok ( NULL, "," ));
strcpy (stop, strtok ( NULL, "," ));
strcpy (stop, strtok ( NULL, "," ));
strcpy (crash, strtok ( NULL, "," ));
strcpy (crash, strtok ( NULL, "," ));
/* debug print to console */
```

```
printf ( "%d: subj_no = %s, start = %s, stop = %s, crash = %s -- ", i+1, subj_no, start,
stop, crash );
Make_txt ( subj_no, start_date, start, stop, crash ); /* call function to make single text file
for FAST input
*/
}
/* end main */
fclose ( input_file );
exit(0);
}
Make_txt ( char subj[16], char day[16], char begin[16], char end[16], char work[16] )
/* make single text file for FAST input */
{
FILE *txt_file;
char txt_name[64];
char temp[8], short_name[8];
int check, begin_h, begin_m, end_h, end_m, work_h, work_m;
int sleep_start, sleep_end, work_mark;
int awake17=17*60, add_sleep_start, add_sleep_end;
int early_h, early_m;
int i, len=8;
/* convert string times to hour and minute integers */
strcpy ( temp, strtok ( begin, ":\n" )); /* start time of reported sleep */
check = atoi ( temp );
begin_h = check;
strcpy ( temp, strtok ( NULL, ":\n" ));
begin_m = atoi ( temp );
strcpy ( temp, strtok ( end, ":\n" )); /* end time of reported sleep */
end_h = atoi ( temp );
strcpy ( temp, strtok ( NULL, ":\n" ));
end_m = atoi ( temp );
strcpy ( temp, strtok ( work, ":\n" )); /* time of accident -- for 1-minute work period */
work_h = atoi ( temp );
strcpy ( temp, strtok ( NULL, ":\n" ));
work_m = atoi ( temp );
/* correct for over-midnight sleep period */
if ( end_h < begin_h )
{
end_h += 24;
work_h += 24;
}
/* check and correct for sleep start before 07:00 due to FAST default run-in sleep */
if ( check < 7 )
{
begin_h += 24;
```

```c
end_h += 24;
work_h += 24;
}
sleep_start = ( begin_h * 60 ) + begin_m;
sleep_end = ( end_h * 60 ) + end_m;
work_mark = ( work_h * 60 ) + work_m;
/* if needed, add an early sleep period to prevent waking period longer than 17 hours */
add_sleep_start = 0;
add_sleep_end = 0;
if ( begin_h > 23 )
{
add_sleep_end = sleep_start - awake17;
add_sleep_start = add_sleep_end - (( begin_h - 23 ) * 60 );
if ( add_sleep_start < 420 ) add_sleep_start = 420;
}
/* debug output to console */
printf ( "%d:%d to %d:%d, then %d:%d -- ", begin_h, begin_m, end_h, end_m, work_h,
work_m );
printf ( "%d to %d, then %d -- ", sleep_start, sleep_end, work_mark );
printf ( "add = %d to %d\n", add_sleep_start, add_sleep_end );
/* make txt file for FAST with 1s, 0s and a w */
strcpy ( txt_name, "/Users/jamesandjoymiller/Desktop/HF_txt_files/" );
strcat ( txt_name, subj ); strcat ( txt_name, ".txt" );
txt_file = fopen ( txt_name, "w");
if ( txt_file == NULL )
{
puts( "*** Can't open txt file ***");
exit(1);
}
/* print header and 1s, 0s and a w to text file for FAST input */
fprintf ( txt_file, "%s, 07:00:00, 60\n", day );
fprintf ( txt_file, "subject no. %s\n", subj );
for ( i = 0; i < add_sleep_start; i++ ) fprintf ( txt_file, "0\n" );
if ( add_sleep_start ) fprintf ( txt_file, "*added sleep period\n" );
for ( i = 0; i < add_sleep_end - add_sleep_start; i++ ) fprintf ( txt_file, "1\n" );
for ( i = 0; i < sleep_start - add_sleep_end; i++ ) fprintf ( txt_file, "0\n" );
fprintf ( txt_file, "*reported sleep period\n" );
for ( i = 0; i < sleep_end - sleep_start; i++ ) fprintf ( txt_file, "1\n" );
for ( i = 0; i < work_mark - sleep_end; i++ ) fprintf ( txt_file, "0\n" );
fprintf ( txt_file, "*1-min work period for crash\n" );
fprintf ( txt_file, "w\n" );
for ( i = 0; i < 60; i++ ) fprintf ( txt_file, "0\n" );
/* end Make_txt */
fclose ( txt_file );
return;}
```

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX H.   SAMPLE TEXT FILE TO INPUT INTO FAST

Example with header:
3/21/06, 07:00:00, 60 *date, start time, epoch length (*denotes comment)
Subject no. 329006526-2
*three minutes of wakefulness
0
0
0
*three minutes of sleep
1
1
1
*three minutes of work (not case sensitive w = W or vice verse)
W
W
W
*three more minutes of wakefulness without work
0
0
0
…
…
…

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX I.   ALGORITHM USED FOR PERMUTATION TEST CONDUCTED IN S-PLUS

```
function(n = 1000, d1 = HF, d2 = NONHF)
{
        len1 <- length(d1)
        len2 <- length(d2)
        data <- c(HF, NONHF)
        the.sum <- sum(data)
        result <- numeric(n)
        for(i in 1:n) {
                new.HF <- sample(data)[1:len1]
                new.sum <- sum(new.HF)
                result[i] <- (new.sum/len1) - ((the.sum - new.sum)/len2)
        }
        return(result)
}
```

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX J.   SAMPLE INPUT FOR PERMUTATION TEST USING S-PLUS

```
> HF <- scan("clipboard")  ##Vector of HF predicted levels of effectiveness from Excel
> NONHF <- scan("clipboard")  ##Vector of NONHF predicted levels of effectiveness
> length(HF)  ##Length of the Vector
[1] 496
> length(NONHF)  ##Length of the Vector
[1] 49
> mean(HF) - mean(NONHF)  ##Difference between the mean levels of effectiveness
[1] -1.998049
> sam.sim.out <- sam.sim()  ##Algorithm from Appendix H
> summary(sam.sim.out)
Min.         1st Qu.       Median        Mean          3rd Qu.       Max.
-2.3667      -0.6223       -0.0040       0.0033        0.5663        3.0136
> sum(sam.sim.out <= -1.998)  ##Number of values below -1.998
[1] 8
> 8/1000
[1] 0.008  ##p-value
```

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX K. DESCRIPTIVE STATISTICS AND TWO-SAMPLE F-TEST & T-TEST (HF AND NONHF) USING EXCEL

HF		NONHF	
Mean	94.18399194	Mean	96.18204082
Standard Error	0.277910176	Standard Error	0.537421089
Median	95.885	Median	96.93
Mode	96.87	Mode	93.46
Standard Deviation	6.189353501	Standard Deviation	3.761947623
Sample Variance	38.30809676	Sample Variance	14.15224991
Kurtosis	3.553667587	Kurtosis	3.55474114
Skewness	-1.740874042	Skewness	-1.787824094
Range	38.54	Range	17.69
Minimum	64.98	Minimum	83.24
Maximum	103.52	Maximum	100.93
Sum	46715.26	Sum	4712.92
Count	496	Count	49
Largest(1)	103.52	Largest(1)	100.93
Smallest(1)	64.98	Smallest(1)	83.24
Confidence Level(95.0%)	0.546028996	Confidence Level(95.0%)	1.080557502

63

F-Test Two-Sample for Variances

	HF	*NONHF*
Mean	94.18399194	96.18204082
Variance	38.30809676	14.15224991
Observations	496	49
df	495	48
F	2.706855588	
P(F<=f) one-tail	2.9671E-05	
F Critical one-tail	1.468750422	

t-Test: Two-Sample Assuming Unequal Variances

	HF	*NONHF*
Mean	94.18399194	96.18204082
Variance	38.30809676	14.15224991
Observations	496	49
Hypothesized Mean Differer	0	
df	77	
t Stat	-3.302422577	
P(T<=t) one-tail	0.000727664	
t Critical one-tail	1.664884538	
P(T<=t) two-tail	0.001455328	
t Critical two-tail	1.991254363	

64

INITIAL DISTRIBUTION LIST

1.   Defense Technical Information Center
        Ft. Belvoir, Virginia

2.   Dudley Knox Library
        Naval Postgraduate School
        Monterey, California

3.   Commanding General, Training and Education Command
        MCCDC, Code C46
        Quantico, Virginia

4.   Director, Marine Corps Research Center
        MCCDC, Code C40RC
        Quantico, Virginia

5.   Marine Corps Tactical Systems Support Activity (Attn: Operations Officer)
        Camp Pendleton, California

6.   Director, Operations Analysis Division
        Code C19, MCCDC
        Quantico, Virginia

7.   Marine Corps Representative
        Naval Postgraduate School
        Monterey, California

8.   CAPT Nicholas Davenport
        Naval Safety Center
        Norfolk, Virginia