PHD

Evaluating Cognitive Screening as a Possible Solution to Reducing Accidents and Improving Workplace Productivity through Early Preventive Detection of Fatigue-Impairment in the Construction Industry

Powell, Ronald

Award date:
2012

Awarding institution:
University of Bath

Evaluating Cognitive Screening as a Possible Solution to Reducing Accidents and Improving Workplace Productivity through Early Preventive Detection of Fatigue-Impairment in the Construction Industry

Ronald I Powell

A thesis submitted for the degree of Doctor of Philosophy

University of Bath

Department of Architecture and Civil Engineering

August 2012

ABSTRACT

Fatigue is emerging as a significant concern in the workplace principally focused on its relationship to accidents and lost productivity. Construction work exposes workers to many hazards and if safety programmes are not effective, accidents will result. Based on the sector's safety performance, workers are not being adequately protected and improvement is needed. Fatigue-related impairment has been identified as a subject of concern for all workplaces yet it is not yet a focus within construction and few operational studies have been undertaken to develop tools to assist with identification and control of this workplace impairment. This research started with an assessment of the management of impairment within the global construction industry as well as an evaluation of tools that might assist in identification and classification of fatigue levels. In particular, cognitive tests were studied and shown to have sensitivity to natural changes in alertness in an operational setting. A small battery of cognitive tests was compared and showed that cognitive tests based on reaction times were possible candidates to help identify fatigue-related impairment in real time.

The top performing tests were then used as possible surrogate measures for fatigue. To finally assess their performance capability their output was compared to estimations from an advanced actigraph-fed fatigue model. 100 volunteer workers each wore an actigraph for a month each to collect information on their personal sleep/wake cycles and activity whilst periodically doing the cognitive tests. The data from the actigraphs was analyzed by proprietary software to determine individual performance effectiveness. It was found that output from these simple, quick, and low cost tests significantly correlated with the most advanced actigraph-fed fatigue model.

It is concluded that cognitive tests can be used as screens for fatigue-related impairment in the workplace. All primary parameters used for modelling showed extremely high significance (Pr(>Chisq) < 2.2e-16) in correlation to fatigue-based effectiveness results and could be developed into a screening tool for fatigue-

related impairment in the construction industry as part of a fatigue management programme.

ACKNOWLEDGEMENTS

The author is most thankful for the time and effort Dr. Alex Copping, Director of Studies for Construction Management in the Department of Architecture and Civil Engineering at the University of Bath, put into guiding this work. His oversight and assistance in completion of this project is greatly appreciated.

The author thanks Dr. David Darby, Associate Professor at the University of Melbourne for his on-going support over several years that this research spanned. His insight, knowledge, patience, and prompt responses were immensely helpful in guiding the study.

The author gratefully acknowledges the contributions of Dr. Gavin Shaddick, a Reader in Statistics within the department of Mathematical Sciences at the University of Bath, as well as Holly Tovey, a student of Dr. Shaddick's. Their insight and knowledge of statistics was most instrumental in ensuring proper treatment of the collected data to draw valid final conclusions.

The author is most appreciative of the support given by Larry Ward, Chief Operating Officer of the author's place of employment. From the outset of starting this journey, Mr. Ward provided his full support to the author including permitting the research within the workplace. This research would not have been possible without his support.

Additional thanks go to all the participants who volunteered their time to either conduct repetitive cognitive tests twice a day or to wear an actigraph and conduct random cognitive tests over a month of their life. Their contribution formed the foundation of this work which will hopefully improve others' lives.

The author would like to thank and acknowledge the financial contribution of WorkSafeBC without which this research would not have completed in this time frame or scale.

Finally, the author offers many thanks to his wife Pascale who endured the countless hours that this work preoccupied him. Without her support, the journey would not have completed.

TABLE OF CONTENTS

APPENDICES

2. Powell, R. I., Copping, A.G.A. 2010. Construction Worker Sleep Deprivation and Its Effects on Personal Safety; Conference – ARCOM, September 6 – 8, 2010. Leeds
3. Powell, R. & Copping, A. G. A. 2010. Sleep Deprivation and Its Consequences in Construction Workers. *Journal of Construction Engineering and Management,* 136**,** pp1086-1092.
4. Powell, R. 2010. Impairment in the Workplace ....Fatigue Gets a "Free" Pass. 28th International System Safety Conference - Building Bridges to a Safer Future, August 30 - September 3, 2010, Minneapolis, Minnesota. Minneapolis: International System Safety Society.
5. Powell, R. I., Copping, A.G.A., Shaddick, G., Darby, D., 2012, Using cognitive tests as screens to detect fatigue-related impairment in operational settings; *Pending Publication*
6. Powell, R. I. 2012. Detection of Fatigue-Impairment in the Workplace; Final Report. Richmond: WorkSafeBC; (IFP #00422; Contract #5071).

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATI0NS AND ACRONYMS

Acronyms used in this thesis:

BAC – Blood Alcohol Concentration
DET – Detection task
EDS – Excessive Daytime Sleepiness
EEG – Electroencephalographic
FAA – Federal Aviation Administration
FACTS – Fatigue Accident Causation Tracking System
FAST – Fatigue Avoidance Scheduling Tool
FMP – Fatigue Management Plan
GML – Groton Maze Learning task
ICM – International Construction Management
IDN – Identification task
LME – Linear Mixed Effects
ONB – One Back card task
OCL – One Card Learning task
OSA – Obstructive Sleep Apnoea
PVT – Psychomotor Vigilance Task
REM – Rapid-eye Movement
SAFTE – Sleep, Activity, Fatigue and Task Effectiveness
SBT – Sleep Before Task
SCN – Suprachiasmatic Nucleus
TATT- Tired All the Time
TWB – Two Back card task

Symbols used to abbreviate scientific terms and statistical terminology throughout the thesis:

AIC – Akaike Information Criterion
ANOVA – Analysis of Variation
BIC – Bayesian Information Criterion
Chisq – Chi-squared
cos – cosine
Dev - Deviance
e – exponential
h – hour
Hz - Hertz
ln – natural log
m – MESOR
mg – milligrams
min – minutes
ms - Milliseconds
Pr – Probability
π – pi
REML – Restricted Maximum Likelihood
μV – microvolt

CHAPTER 1 - INTRODUCTION

1.1 Research Overview

A construction worker was killed at the University of Bath working on a project developing new student residences. While presenting his update on projects, the Director of Estates announced the R5 Student Residences had suffered a considerable set back in November 2007 as the result of this fatality on the site (Vice-Chancellor's Office 2008). On top of the tragedy itself, the project was stopped for several weeks putting it behind schedule and increasing costs. This unfortunately occurs far too often in the construction industry. This fatality was one of over 800 deaths from injuries resulting from construction work over a ten year period through 2008 in Great Britain (GB); a rate over 4 times the GB average for all industries (Walters and Bolt, 2009a). While showing significant improvement since the 1980s, the rate has levelled off in the last 5 years and the industry remains the cause of more worker fatalities than any other sector in GB. This has prompted the Rt. Hon. James Purnell MP as Secretary of State for Work and Pensions on December 4, 2008 to commission an inquiry into causes due to the great concern presiding over the number of construction industry accidents.

This research is supportive of the Phase 1 Report findings released in July, 2009 which identified "a long hours culture in the industry results in fatigue, compromised decisionmaking(sic), productivity and safety" (ibid, p92) as one of eleven issues contributing to non-fatal accidents and recommended for subsequent investigation. Additionally, the report found "Many veins of research contribute to understanding the underlying causes of construction fatal accidents including: 1. Deep research into factors generically and in specific contexts (such as fatigue, non-compliance etc) to understand how they occur and can affect safety" (ibid, p67).

Great Britain has no monopoly on a high rate of accidents and a high severity of accidents in construction. The European Union average is almost three times the Great Britain rate and globally the construction sector is a poor performer. A risky construction environment is accepted and rules and a safety focus have evolved to make improvements but not all factors that may help the industry's safety and resulting productivity have been studied in detail.

Within construction and other sectors the pervasiveness of the potential problem of fatigue-related impairment is not well understood. To date most research has focused in sectors such as transportation, aviation and the medical field and has tried to understand the contribution to accidents and lost productivity and the potential connection to fatigued workers. However, an understanding of the degree fatigue-impairment may even be an issue is underdeveloped. An important premise of this research is that this understanding is best addressed by measuring fatigue-impairment in operational settings.

Powell (Powell, 2007) found the presence of sleep deprivation in a construction operational setting and revealed significant fatigue-impairment confirming speculation of its presence. Actual sleep measurements of construction workers fed an impairment model and coupled with other relevant indicators from extensive survey pointed to a fatigue-impaired workforce with higher workplace risk. This research provided confirmation and actual data to support wide-held beliefs regarding fatigue in the construction workplace. The magnitude of the estimated decrement in performance lead to questions associated with what organisations might do to better manage fatigue-impairment. This includes how to identify it, how to measure it, developing fatigue-impairment plans and establishing effective guidelines to take appropriate action as necessary.

Fatigue-related impairment is a factor that impacts all of us thereby making it a factor that we all can relate to on a very personal level. However, there are derived problems from this. As it often impacts us as an imposed aspect of living in the 21st century, Powell (Powell, 2010) showed that in Europe and North America we view and treat it differently from other forms of selfimposed impairment. We are more accepting of it and tolerant of its effects regardless of the increased danger it may pose. Additionally, we have no reliable techniques to identify fatigue-impairment levels. This is supported by a surprising lack of strong knowledge associated with sleep, fatigue-impairment and its impact on a personal or organisational level.

Having identified fatigue-impairment in the workplace with higher worker risk of accident, a driving question now is can it be measured or quantified in real time? If it can not be measured, controlling it will be difficult or not possible. Tools have been identified which purport to have this ability but have not been used in an operational setting or tested in an operational setting with uncontrolled factors affecting individuals and their alertness. This research intends to add to the body of knowledge associated with fatigue-impairment in the workplace by testing new tools for use in detecting and measuring fatigue and mental effectiveness levels. Further, all factors that might potentially defeat the measurements are also included in the study to develop understanding on practicality of use. Finally, results are put forward as recommendations for change and mitigation with potential solutions for identifying and quantifying fatigue-impairment.

1.2 Statement of Purpose

Identification of fatigue-related impairment in the workplace raises new questions which are a focal point for this research to address. Answering them has huge relevance for individuals and organisations as it could potentially not only be the solution to part of the accident/productivity issue but could also potentially save lives. The fact that workers rate fatigue as the most prevalent form of impairment in the workplace with no broad testing for it and a general acceptance of it suggests that the consequences of fatigue and impairment are not understood well enough to drive the changes needed to address them. This research is intending to assist with this problem by helping expose the problems and solutions. The need is clearer but the solutions are not. The purpose of this research is to take a small step in that direction.

1.2.1 Importance of Research

'Western' world societies have accepted that individuals have certain rights pertaining to their health and safety at home and at work. All workers have a right to expect to go home from their workplace in the same condition of health they arrived in. Certain abuse and misuse of substances such as illicit drugs and alcohol have been identified as impairment agents which can increase risk for individuals and others. Their use is either banned or strictly controlled by laws in high-hazard workplaces and in safety critical activities such as driving. Initial work (Powell, 2009) associated with this field of study has presented strong evidence that sleep deprivation was prevalent amongst construction workers leading to workplace fatigue-impairment and psychomotor-performance decrements with increased workplace risk. This risk is not different from that of alcohol and illicit drug use but it is tolerated.

While the work was revealing as the first large study to estimate the extent of the fatigue factor from a lack of quality sleep in construction, it raised derivative questions associated with the findings including:

- How is fatigue-impairment viewed in the workplace as an issue?
- Have organisations come to anticipate decrements in performance from fatigue-impairment and thus develop longer schedules and coping mechanisms for it or are they creating the problem with the long and variable schedules?
- How would an organisation know if its workers were too fatigue-impaired to work? What measurements could be used?
- What mechanisms exist for organisations to detect and screen out fatigue-impaired workers?
- Is there any evidence to suggest that a fatigue-impaired worker is less concerning than other forms of impairment in the workplace?
- How should fatigue-impairment be addressed in the workplace? What policies and tools are available to organisations to assist them in coping with fatigue-impairment?

- How do and how should individuals and organisations cope with fatigue-impairment?

Many factors can affect impairment and one's 'fitness-for-duty' including fatigue-impairment but employers' and societies' focus continues to be alcohol and illicit drug use. From an impairment perspective, regardless of source, employers should be more concerned with the degree of impairment than its source. This research continues to expose fatigue-impairment in the workplace as a serious problem but intends to help organisations understand its cost and deal with it. Having tools to recognize it, evaluate it and remove it from the workplace when unsafe thresholds are reached would be invaluable.

1.2.2 How this Work is Original

This work is original on several fronts:

1. It was conducted in an operational setting with many concerning elements of the construction industry such as 24 hour shift work, long work hours, safety-critical roles, variable weather conditions, long commutes and pervasive workplace hazards. As field-based research in an operational setting, information was collected from actual workers in their work environment as opposed to controlled clinical and lab studies. Robust research attempting to correlate fatigue-impairment from work and sleep habits with alertness tests at this level has not occurred in an operational setting.
2. It is focused squarely on addressing means of detecting fatigue-impairment in an operational setting as follow-on to finding pervasive fatigue-impairment in the construction workplace.
3. It is assessing the risk associated with fatigue-impairment in a construction environment. The construction environment has not been assessed before for fatigue-impairment and has no tools to deal with it.
4. It is mapping out an identified gap in workplace impairment management which must be bridged to practically deal with fatigue in the workplace.

5. It has identified new evolving tools that may assist detection of workplace impairment but which have not been assessed in an operational setting and correlated to fatigue-impairment.
6. It addresses the significance of worker fatigue and how it relates to the over-all area of workplace impairment. There are no known effective practical screens for fatigue-impairment in the workplace but there are potential solutions which this research evaluates.

To effectively conduct this research requires the support of several large institutions and organisations. Obtaining such support illustrates the interest level generally existing. Perhaps it exists because of the unanswered questions requiring exactly this type of work to address them.

1.3 Aim, Objectives, Hypothesis

1.3.1 Aim

This research aims to determine an effective means of screening for fatigue-impairment and declaring worker fitness-for-duty to reduce risk of accidents and improve productivity in all places of work.

1.3.2 Objectives

Several objectives have been set for this research including:

1. To detect and provide actual objective measurements and assessment of fatigue-impairment to confirm fatigue-impairment exists.
2. To provide an understanding of the impact of fatigue-impairment on construction workers' risk, safety and productivity.
3. To provide understanding on issues and impediments associated with dealing with fatigue-impairment in the workplace.
4. To review and critique means of assessing organisational and individual fatigue.
5. To test possible fatigue-detection solutions in an operational setting and assist with development to allow introduction into the construction workplace as a tool to screen out fatigue.
6. To validate the effectiveness of tested fatigue-detection tools with proven means of assessing fatigue-impairment.

7. To contribute to improved productivity through the reduction in work-related injury, illness and working days lost by providing a path forward for the industry in dealing with fatigue-impairment.

1.3.3 Hypothesis

The Hypothesis for this research is:

- Fatigue-impairment can be identified in the workplace and the associated risks from individuals' fatigue-impairment can be reduced by knowing about the condition and taking appropriate actions.

1.4 Ethics

The foundation of this research is set on extraction of data from workers in multiple locations and over several years to gain understanding of a real issue which was concerning for them in their workplace and personal lives. Protection of individual rights and safeguarding personal information was of paramount concern and necessitated emplacement of a Memorandum of Understanding (MOU) signed by each participant and the researcher prior to participation. Additionally, all participants chose to participate freely, without remuneration including participation in completion of surveys, testing and sleep monitoring. A confidential personalised sleep report was prepared and returned to each participant at the completion of wearing an actigraph for a month. A copy of the MOU is included in Appendix 8.

1.5 Boundaries and Limitations

The sector of interest for this research was the global construction sector. However, this is a broad space with numerous definitions associated with activities, roles and organisations that make up this sector. It is also an evolving sector. Some of the largest global engineering and construction companies, including the world's largest, control assets and operate and maintain them for the duration of contracted periods; for instance as the completion of a concession agreement under a Public-Private Partnership. This then by extension takes these organisations into non-traditional construction spaces such as equipment servicing, operation and maintenance. However, due to the very nature of the sector which employs workers across a broad spectrum of activities from the office front end design, engineering and procurement, to project execution and site activities to workers who now complete the cycle by maintaining the construction assets, extremely diverse activities fall within construction. As much as was possible within the boundaries of worker access, financing and practical research project management, data has been collected from personnel who share some of the same hazards and conditions as others in this sector. A very real limitation must be understood in terms of the general use of the term construction as it relates to the participants in this study and the millions who are employed in the sector globally. Whilst construction workers face varying hazards within the sector across sites, geography, cultures and projects the one thing they all have in common is human physiology which is at the heart of any discussion on fatigue-impairment. This makes the study of fatigue relevant to anyone regardless of their activity and without boundaries. Where activity becomes important is in the discussion on implementation and focus of solutions.

Increasingly, large global organisations are playing a role in establishing their own common culture on top of the local geographic cultures. This is especially true in the space of health and safety. Common standards must be established for all workers regardless of where the project is being completed. As such, when it comes to health and safety, not only are common standards sought for organisational efficiency, but best practices are sought. This allows leveraging developments in any location for roll out and application globally after fitting to local circumstances. This research fits this circumstance completely. Except for the global surveys, the research was conducted in a high-hazard, highly regulated operating unit which was part of a large global engineering and construction company in Canada. It should not be thought of as a simple case study but rather a pilot study of new technologies and methodologies for roll out as applicable to a larger diverse community globally. Chapters 5, 7 and 9 report on data derived from different groups affiliated with this sector and to the degree any concerns are raised on the boundaries of applicability, they are discussed, but in general results are not seen with boundaries and except for the surveys, findings should be generally applicable even outside of construction. The relevance of results extracted from a real world workplace is as opposed to clinical or laboratory research which may impose controls not normally practiced in the field. For new technology to succeed it must pass the final hurdle of field acceptance and resolve issues which can and are only seen in the real workplace. This made the results from this research all the more powerful. Finally, while construction is filled with a large number and variety of work tasks, nothing in this research is specific to a task that might rule out its applicability to any worker anywhere. As such there were no restrictions placed on individual participation and there is no concern over whether an individual was representative of other's work activities. As long as they share independent normal human physiology, results obtained from the small population will have some applicability to others anywhere in the world. The relevance of task again comes back to selection of application of changes and one can imagine some roles are more safety-sensitive than others and therefore would be a priority to address with any solution. Figure 1, illustrates the focal point of the research and its position within this global sector.

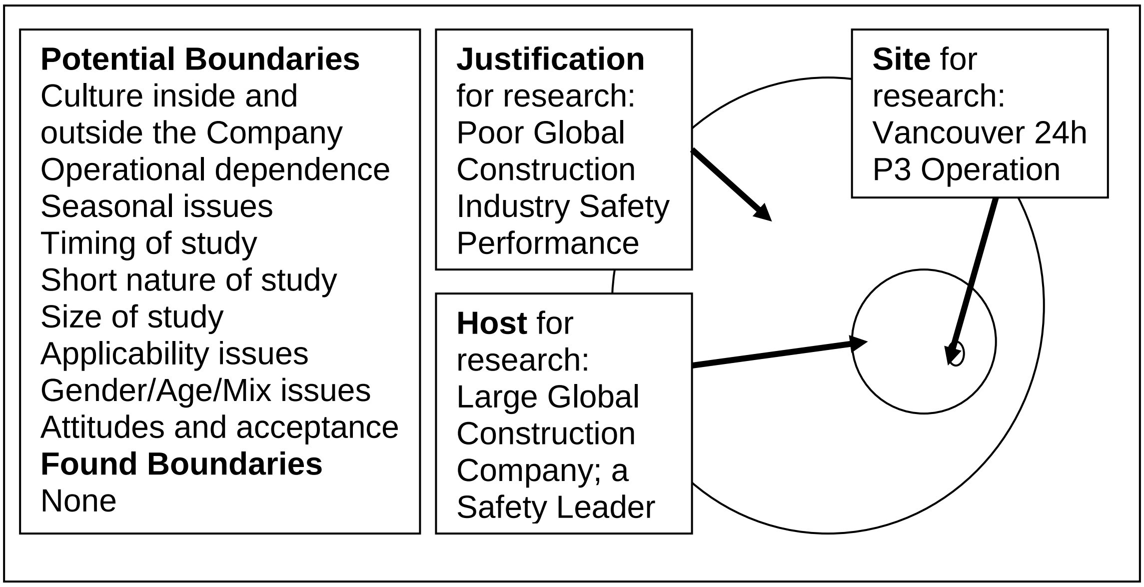

Figure 1 – Boundaries of the Research

CHAPTER 2 – RESEARCH METHODOLOGY

2.1 Introduction

To understand fatigue-impairment, the research starts with a literature review associated with sleep and studies pertaining to the derived impairment in the workplace. This research uses surveys, cognitive tests, actigraphy and results of other research to develop an understanding of attitudes and ranges of cognitive performance from tests of regular workers in the workplace. Normal worker fatigue levels will be assessed in a population of workers over a one month period with cognitive tests concurrently taken to determine levels of performance in different areas of cognition. This research will attempt to draw correlation between the two. As the tests are testing human cognition and how it is influenced by fatigue, the tests are not task or worker dependant other than a desire to have some workers with normally occurring and real fatigue in their jobs. This provides broad flexibility on recruiting candidate participants. Even fatigued, sufferers of sleep disorders are candidates as they are naturally found in work places.

2.2 Methodology

This research focuses on the fatigue component of workplace impairment. To do this the research builds on five levels using various means to:

1 - Understand fatigue-impairment and assess it.

2 - Review evidence that fatigue-impairment even exists in the workplace or is a workplace problem.

3 - Review what is being done and why more is not being done in this area if fatigue-impairment exists and is harmful.

4 - Review possible techniques to detect fatigue and results from testing with new tools. Possible candidates for further testing are selected.

5 - Review results of new testing to validate the new tools with proven fatigue-assessment tools.

2.2.1 Fatigue-impairment

To understand fatigue-impairment, a review of existing studies associated with the general field of sleep, fatigue and fatigue management in the workplace was undertaken covering:

- The important basics of fatigue and sleep required for understanding this subject;
- How fatigue affects motor and cognitive skills and how it is measured. This was required to understand the options available currently as tools.
- What has already been learned in construction and other sectors regarding applying tools to control fatigue-impairment in the workplace.

2.2.2 Search for Evidence of Fatigue in the Workplace

Research was undertaken to determine if there is any evidence from other studies associated with fatigue in the workplace. Sectors with results were reviewed including construction. Significant evidence of fatigue and fatigue-impairment in the workplace becomes a driver for the need for controls. Results from surveys of actual workers and finally actual measurements of sleep help triangulate on the issue.

2.2.3 Review of Attitudes and Action Associated with Fatigue

Surveys were used extensively to extract a profile of attitudes from the construction industry associated with fatigue-impairment. A profile of the industry's fatigue-impairment consciousness is revealed from actions taken including policies and training. The results assist in understanding what is or is not being done to combat fatigue and how big a problem this issue is perceived to be.

2.2.4 Techniques Used to Identify Fatigue

Research was undertaken to understand the types of technology used to identify fatigue. This leads to an understanding of which techniques could be used to identify or measure fatigue in the workplace and forms a basis for the possible means of implementing systems in the workplace. In particular:

- Different techniques are reviewed and critiqued as potential solutions.

- Actual results from available tools are reviewed and assessed for applicability. This narrows the search on possible solutions.
- A project was established to extract preliminary results from a particular cognitive testing technology to determine suitability for workplace application and further testing.
- Finally, the chosen technologies were used to collect actual data.

2.2.5 Comparing Technologies

As a penultimate test, identified and selected technologies from the initial workplace research were tested against results from identified leading and proven techniques for assessing fatigue impairment using actigraphs. The latter part and majority of this research utilized a quantitative approach requiring the support of actual workers to collect data for analysis. The means of the collection was four-fold:

1. Surveys on attitudes in the workplace.
2. Sleep logs and reports were completed by participants.
3. Results from computer-based cognitive tests. To assess this technology a project was established to test at two different times of day at times known to coincide with natural changes in alertness level governed by the circadian cycle. If they could detect significant change between nodes, they might be candidates for further testing and development.
4. Normal sleep patterns of a sample of 100 workers were established by collecting data for a one month period per participant with an actigraph. The actigraphs' technology is now developed such that it permitted continuous use and data extraction for over a full month without battery recharging. While cost effective and practical for this research actigraph use still carries considerable cost which was covered by a small research grant won for the work. Used in conjunction with an associated software program to analyse the data and feed a model to determine individuals' sleep patterns, an estimate of the quantity and quality of their sleep was made. Results allowed a minute-by-minute extraction of mental effectiveness levels

to compare to results from cognitive tests that were randomly taken concurrent to wearing the actigraph per Figure 2.

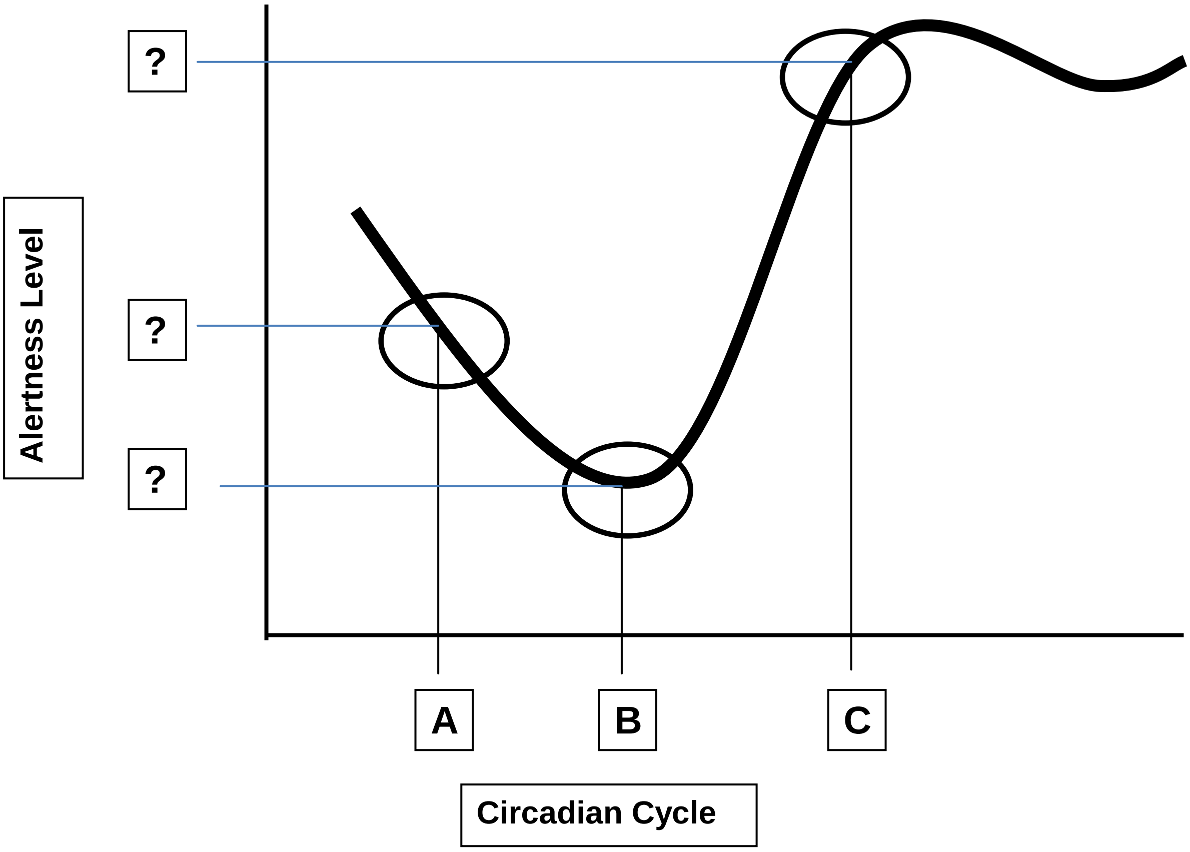

Figure 2 – Using the Circadian as the basis for testing.

Data analysis from the selected technologies allowed a detailed comparison for validity of use of both and established a recommendation to go forward.

2.3 Key Questions

Some key questions served as drivers for the research including:

Is fatigue and sleep deprivation really a concern or should it be? How is it defined and possibly measured? If a lack of proper sleep is important, why is there not more being done? Is the issue one of just not understanding the issue or one of not being able to measure and therefore not control?

What are the effects of sleep deprivation on cognitive capabilities? On motor skills? On general health? On risk in the workplace? On worker mood? Why aren't there more techniques being used? Is fatigue or lack of sleep investigated or tracked? Why or why not? If a technique existed to measure or estimate fatigue would it be used? What would be required to develop it and implement it? How does the industry compare to other industries in terms of what is being done? How would one classify the effects of improper sleep in the workplace? Is there a cost known? Can guidelines for best practices be recommended? How have other industries handled this issue? (e.g. Aviation, Transportation workers, Medical workers, Shift workers). Are there possible solutions for this issue? If yes, what would it take to implement it and how? If yes, what next?

2.3 Summary

The research methodology is summarized in Figure 3. It starts with a literature review to understand current status of the sector and factors that might affect fatigue in the sector. This is followed by surveys to validate perceptions of the current status and concurrently, a pilot study to determine if a potential found solution is worthy of further testing. Finally, results from the pilot study feed the main study for validation of the tests in the work place.

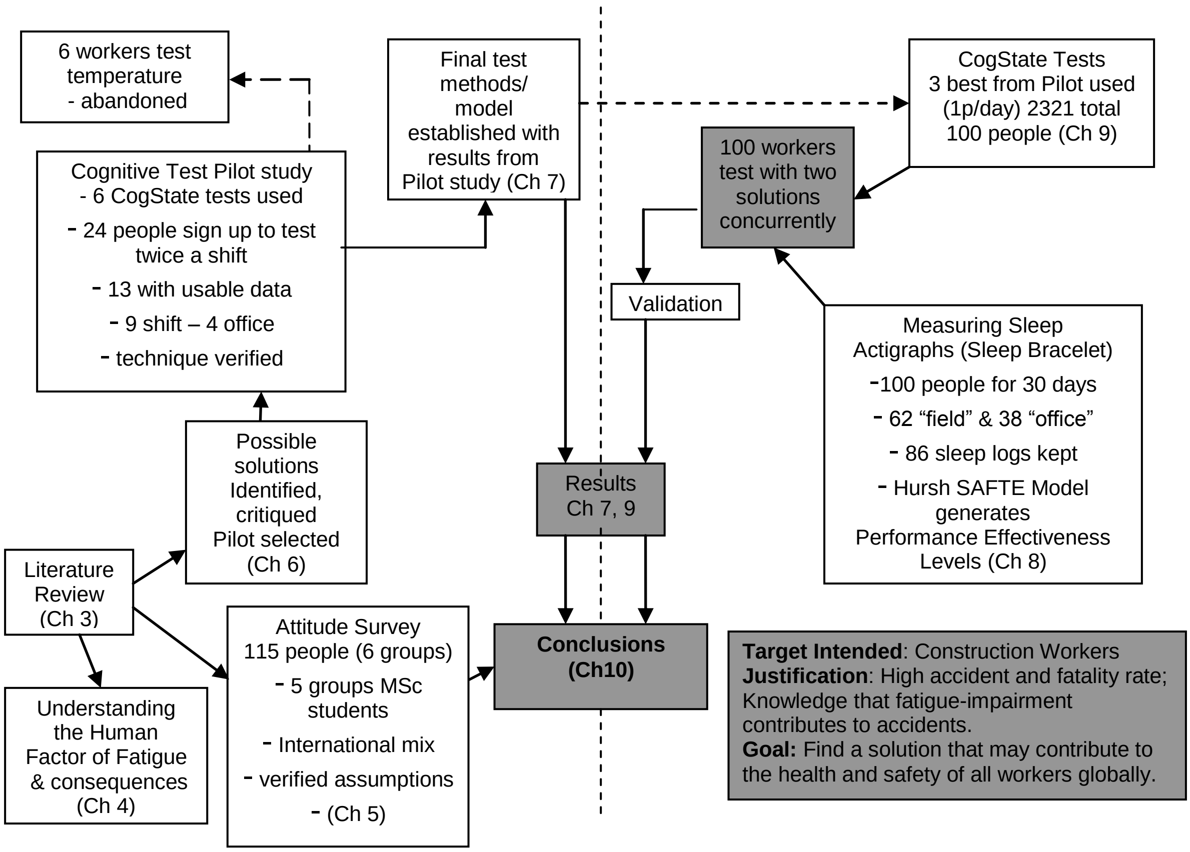

Figure 3 – Research Methodology

The various reviews and research are set out in the chapters that follow.

2.3.1 Chapter 3 – Background Literature review

The literature review establishes boundaries for the issues studied and develops an understanding of what drives fatigue-related impairment; understanding fatigue as an impairment and what has been done to date; what the possible background technologies and solutions are; defining what is and what is not known to frame this research.

2.3.2 Chapter 4 – Managing Fatigue-impairment as a Human Factor in Safety

With a focus on Construction, Chapter 4 reviews how we account for risk and how fatigue-impairment impacts risk in the work place. Fatigue–impairment as a Human Factor is discussed.

2.3.3 Chapter 5 – Fatigue's Free Pass

Chapter 5 deals with attitudes towards fatigue-impairment in the construction industry. Results of surveys are discussed to expose what views are held regarding fatigue as an impairing agent and what is and what is not being done in the construction industry. Key roles are assessed for concerns for fatigue-impairment due to safety-criticality, schedule, or length of shift. A review of what has been done to combat fatigue in the work place is undertaken to understand an apparent lack of action suggesting fatigue-impairment is not viewed as an issue.

2.3.4 Chapter 6 – Identifying and Measuring Fatigue - Related Impairment in the Workplace

Possible means to detect fatigue-impairment are reviewed with results from studies which have attempted to do this with different technologies. Possible means to screen workers and identify fatigue-related impairment are assessed. In particular, actigraph-based effectiveness models and cognitive tests are explored to see if they can be adapted to screen for worker fatigue.

2.3.5 Chapter 7 – Using Cognitive Tests to Screen for Fatigue-Related Impairment in the Workplace

Practical use of the tools includes an attempt to measure workers' alertness using a tool designed to do so. To test sensitivity the tool is used to measure individual normal changes in alertness correlated to the circadian rhythm. The results accommodate other factors normally found in the workplace which may confound results such as caffeine consumption. Chapter 7 reviews research with individual candidates using the cognitive screens in an operational setting. The screening results are examined to learn over-all effectiveness and issues associated with their use including how they may be beaten. Models are developed and discussed to help understand results relative to organizational confounding parameters. Recommendations are made on best choices for the intended use.

2.3.6 Chapter 8 – Measuring Sleep to Validate Fatigue Measurements

Chapter 8 focuses on establishing grounds for using actigraphs and sleep models as a basis for validating cognitive screening results. The selected technology is explained and reviewed.

2.3.7 Chapter 9 – Correlating Real-time Measurements with Sleep

Chapter 9 includes results from workers on different shifts who had their actual sleep/wake patterns measured 24-hours per day for a month. In conjunction with the use of actigraphy and proven sleep-based fatigue models their fatigue levels were assessed and compared to results from random cognitive tests. Results for correlation between a screening tool reading and fatigue-impairment are established and multiple models are developed.

2.3.8 Chapter 10 – Recommendations and Limitations

Chapter 10 reviews findings of the research for validation. Results of the modelling from the tests conducted are discussed. Conclusions supported by the findings and analysis are presented as well as the limitations of the research. Suggestions for further work are presented.

CHAPTER 3 - LITERATURE REVIEW

3.0 Introduction

Four areas are reviewed as core subject background for this thesis.

1. Sleep and fatigue;
2. Fatigue-impairment in the workplace;
3. Impairment detection technology;
4. Detecting and measuring fatigue-impairment in the workplace.

The thesis explores relevant literature in each area as a foundation for subsequent levels. The first two areas above are covered in this chapter and subsequent chapters also review relevant literature.

3.1 Background theory review – Sleep and Fatigue

3.1.1 Introduction

A basic understanding of sleep is foundational to this thesis as it pertains to all aspects of the work conducted. The subject of sleep is broad but the review herein is narrowed to focus on those aspects that are most relevant to the research conducted.

There still remain, many unanswered questions about sleep, and one of the most interesting aspects of sleep is that in the current year 2012, researchers do not know the reason we sleep. In an interview, Dement, the co-discoverer of rapid-eye movement sleep stated that after 50 years of studying and researching sleep, “the only reason we need to sleep that is really, really solid is because we get sleepy” (Max, 2010, p). Additionally, there is no strong reason for the high levels of brain activity during some phases of sleep; the significance of dreaming; or even the basis of the restorative effect of sleep (Gorman, 2004). Since introduction of medical scanners more insight into brain activity is now possible during sleep, resulting in significant progress in the last decade in understanding what happens during sleep (Siegel, 2003). This builds support for some theories explaining sleep but the key point is that while we may not have a definitive answer on why we sleep we understand a lot about what happens during sleep and increasingly, what happens when we don't get adequate sleep.

3.1.2 Sleep Basics

Our knowledge of sleep is derived from relatively recent research into sleep. The beginning of modern sleep research is generally associated with the 1953 discovery by Nathaniel Kleitman and Eugene Aserinksy of rapid-eye movements tied to dreaming during sleep which defined for the first time the presence of at least two states of sleep; non-rapid eye movement (NREM) and rapid-eye movement (REM) sleep. Prior to this it was generally believed sleep was just a period of low brain activity. Later, Kleitman and another of his students, William Dement, showed by means of electroencephalographic (EEG) recordings from normal subjects, that the first state, NREM, can be further characterized in different stages that occur in a sequence (Siegel, 2003). Our very definition of sleep today is tied to this early research which defined sleep in terms of this measured brain wave activity shown in Figure 4.

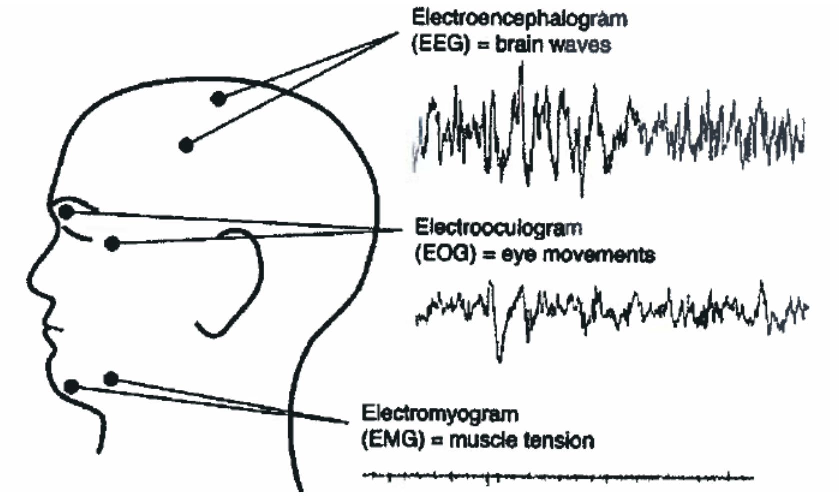

Figure 4 –Placement of electrodes to study and define sleep (adapted from (National Institutes of Health, 2003))

Our normal waking state with the eyes open is characterized by high-frequency (15–60 Hz), low-amplitude (~30 μV) brain activity. This pattern is called beta activity. As we fall asleep into stage 1 NREM sleep the brain wave activity is characterized by decreasing EEG frequency (4–8 Hz) and increasing amplitude (50–100 μV), called theta waves. Moving into stage 2

NREM sleep is characterized by 10–15 Hz oscillations (50–150 µV) called spindles, which occur periodically and last for a few seconds. Stage 3 NREM sleep is characterized by slower waves at 2–4 Hz (100–150 µV). Stage 4 sleep is defined by slow waves (also called delta waves) at 0.5–2 Hz (100–200 µV). After reaching this level of deep sleep, the sequence reverses itself and a period of rapid-eye movement (REM) sleep, ensues. REM sleep is characterized by low-voltage, high-frequency activity similar to the EEG activity of individuals who are awake.

This complicated process of sleeping consisting of these five stages of falling asleep (stage 1), brain slow-down (stage 2), deep sleep (stages 3 and 4) and REM (stage 5) are illustrated in Figure 5. In each stage, the brain and body act differently.

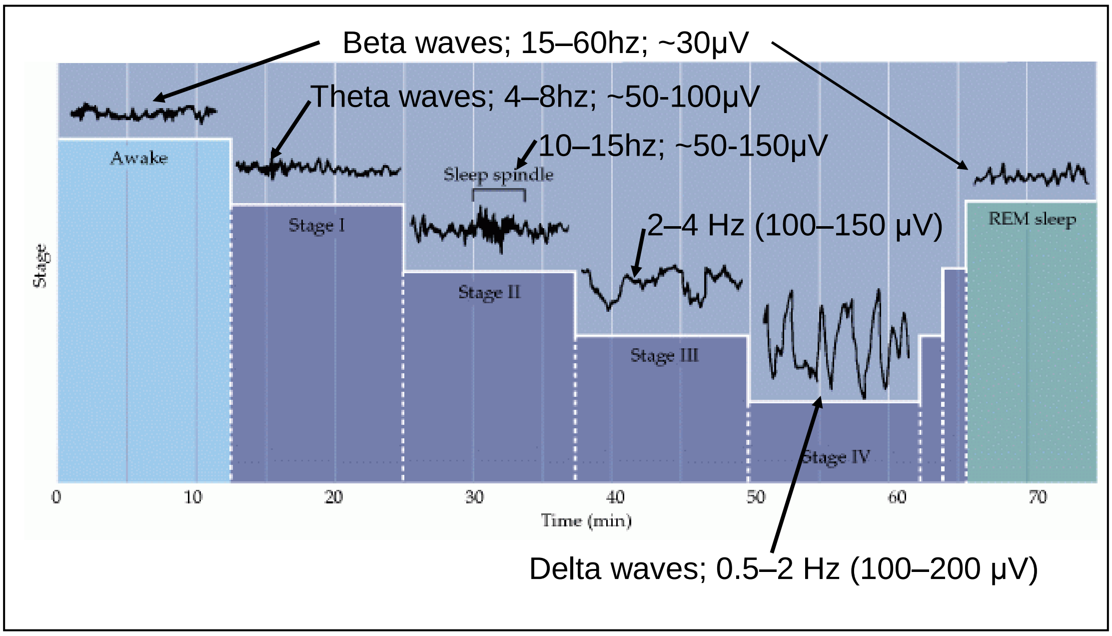

Figure 5 – Stages of Sleep with EEG Recordings of each Phase
(Adapted from Hobson, 1989 in Purves (Purves et al., 2001)

A night of sleep therefore consists of progressing from stage 1 – 4 then reversing into REM then cycling again through all five stages about four to six times per night as shown in Figure 6.

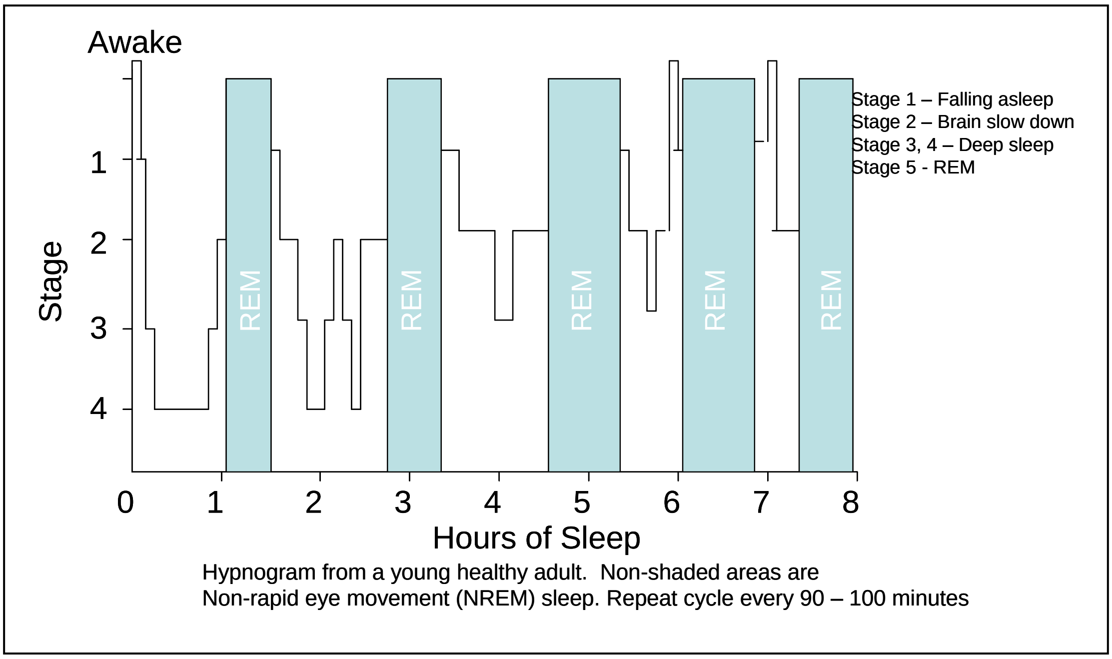

Figure 6 – Hypnogram of a normal night's sleep (adapted from (National Institutes of Health, 2003))

However, the profile of each phase with these five stages changes throughout a night's sleep. The beginning of a normal night's sleep has more stage 4 (deep) sleep as a percentage of total phase sleep time than the last phases of sleep where more sleep time is spent in REM. It is also common for individuals to wake several times during the total sleep period. The REM period of our sleep characterized by brain activity similar to when we are awake and coinciding with dreaming is potentially impactful for this research. Different theories can be found regarding the role REM sleep plays, such as brain development, maintenance, learning and memory development and consolidation (Siegel, 2003). Sleep researchers also know but can not explain why REM sleep changes with age. Babies have about 8 hours of REM sleep at birth which drops to about 2 hours at 20 years to only about 45 minutes at 70 years of age (Purves et al., 2001). This research will focus on areas that are potentially impacted by REM sleep such as learning and memory. Additionally, because REM sleep is always the last natural stage of a phase of sleep, if inadequate sleep is attained, it may often be the REM that is clipped. The significance of this is not fully understood (Siegel, 2003) but it could have both short term cognitive effects as well as long term.

Researchers, now aided with new technologies to map out brain activity during sleep have gained a better understanding of NREM sleep. During this time, different parts of the brain are found to do quite different things. The brain-stem neurons just above the spinal column are inactive whilst those of the frontal regions of the brain or cerebral cortex are very active. However, this activity changes during sleep from patterns while awake. When awake, the cerebral cortex neurons appear to 'fire' somewhat randomly in response to the activity of the moment but during sleep, researchers have observed that "adjacent cortical neurons fire synchronously, with a relatively low frequency rhythm" (ibid, p93). Further and in contrast to the inactivity that many believe is associated with sleeping, this synchronous firing produces higher voltage brain waves than when awake but consumes less energy. Breathing and heart rate tend to be quite regular during NREM sleep, and reports of vivid dreams during this state are rare.

Researchers (The Canadian Institutes of Health Research, 2012) also believe a very small group of brain cells at the base of the forebrain is maximally active only during NREM sleep. These cells called the 'sleep-on neurons' appear to be related to inducing sleep believed to be triggered by levels of adenosine. Researchers believe increased body heat influences these neurons too as evidenced by the drowsiness many experience after a warm bath or during a hot day. Recent research found adenosine from a group of helper, non-neuronal brain cells called astrocytes regulate sleep (Science Daily, 2009) and may one day explain why we sleep.

Understanding sleep and fundamental sleep research has benefitted from use of technology for brain-imaging and sensor data processing which allows scientists to record activities of numerous individual neurons and cells simultaneously. What is emerging is a better understanding of brain neuron and cell activity patterns during sleep. The technology has moved a great distance from the first continuous brain recordings that led to the discovery of REM sleep in 1953. Adequate sleep is the simple naturally occurring cure for fatigue. Every person is unique but one thing we have in common is the need for sleep.

3.1.3 Sleep's Control Centre and Programme

The state of sleeping is not the result of a simple reduction of brain activity; rather, sleep is a dynamic and cyclical process of moving through precisely controlled brain states, and while doing so, some parts of the brain are as active as when we are awake. The sequence of sleep states is governed by a group of brainstem nuclei that project widely throughout the brain and spinal cord. The need for continuous sleep that becomes restorative and results in feeling refreshed and alert for the day ahead is inescapable and controlled by specific regions of the brain dedicated to this function.

Analogous to dedicated processors looking after critical control functions of any process, the brain's dedicated processors are located in specific brain regions. The hypothalamus region of the basal forebrain is an important area controlling NREM sleep and generates the NREM EEG patterns (Figure 7).

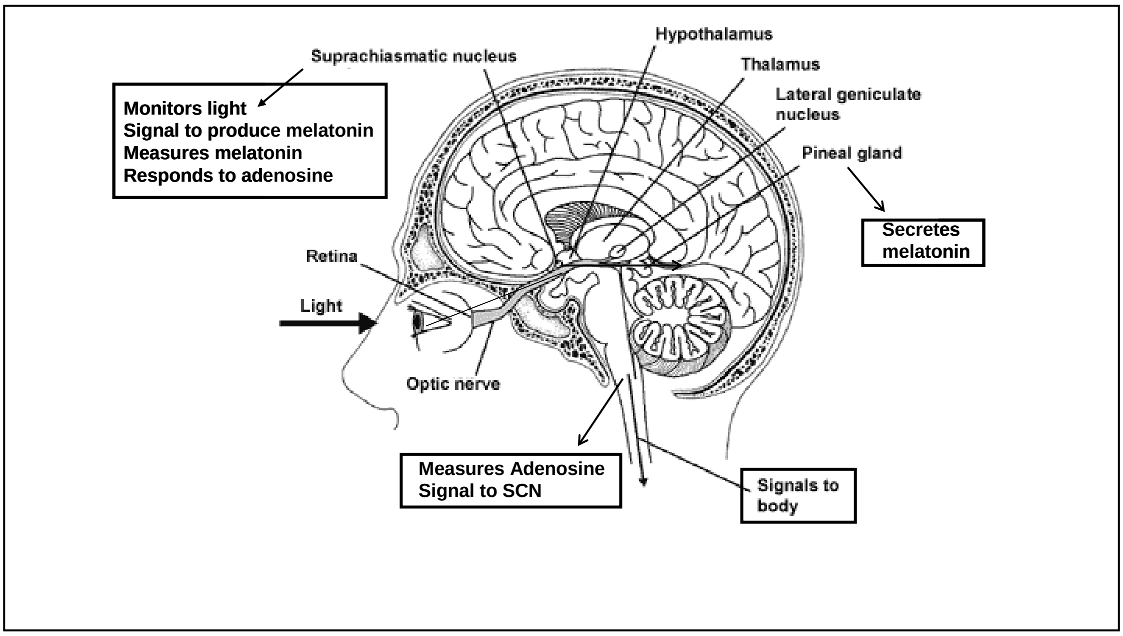

Figure 7 – Regions of the brain and their sleep functions (adapted from (National Institutes of Health, 2003))

An area known as the pons in the brainstem is tied to REM sleep. An area called the Suprachiasmatic Nucleus (SCN) in the hypothalamus region of the brain plays an important role in sleep and it is heavily influenced by natural light (Sleepfoundation, 2007). The SCN is the processor that runs the programme governing our sleep/wake cycle. The programme by which it works is biologically hard-wired in. This means we do not have to think and decide if we should sleep. Sleep is not an option, it is a physiological requirement determined by the many sensors, molecules and the genetic programme of the SCN.

3.1.4 Circadian Rhythms

The internal programme regulating sleep in humans is called the circadian rhythm or cycle. The Sleep Foundation (ibid, p4) reports that there are actually two sleep processes active. One is the restorative process when sleep occurs naturally associated with our time awake and build-up of adenosine that increases the pressure to sleep with time awake. The second is associated with the timing of sleep and alertness after waking during a 24-hour period. The circadian rhythm influences both. It is believed that humans "… like all mammals have a circadian (circa=around; dies=day) pacemaker in the brain that regulates physiological and behavioural functions on a 24-hour basis" (Rosekind et al., 1996, p159).

Circadian rhythms are physiological and behavioural processes controlling such things as sleeping and waking, digestion, hormone secretion, core body temperature regulation, mood and activity.

Unlike some animals, humans are daylight active (diurnal) not darkness active (nocturnal). Light is our signal to get active. During the course of a day, our circadian rhythm defines our natural alertness levels and it varies over the course of a day in a fashion similar to Figure 8.

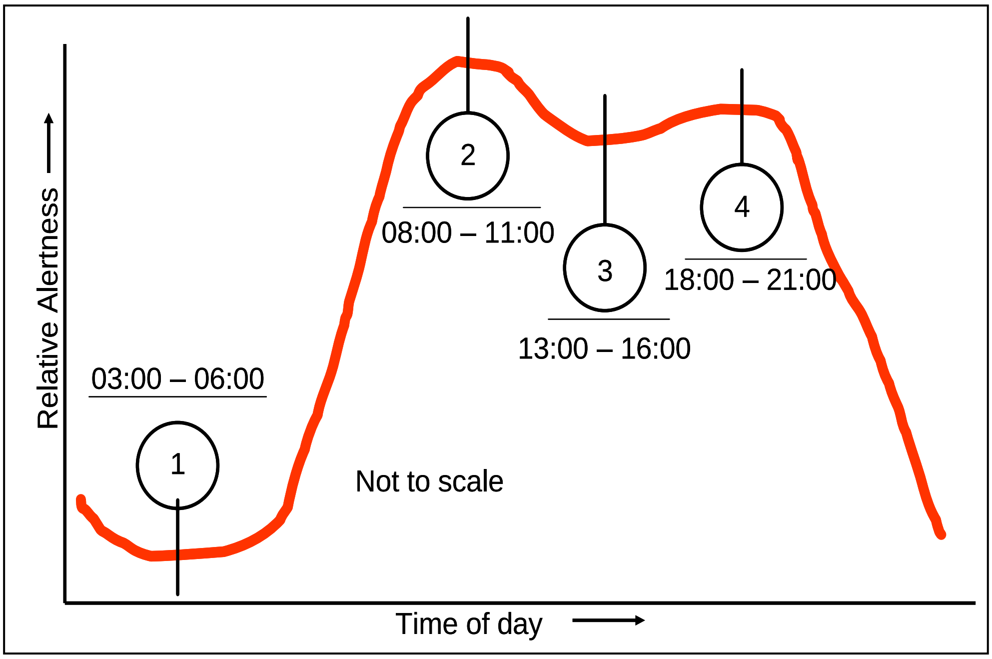

Figure 8 – Profile of alertness governed by the circadian cycle. (Not to scale)

Attempts to measure the circadian in real-time have shown the result to be quite noisy with individual changes in alertness. However, with more data, the circadian rhythm governing our sleep/wake cycle and alertness is generally shown to have two peaks for alertness within 24 hours and two low points (Wang et al., 2000). The shape of the circadian cycle does not have an instrument to directly measure it but is uncovered by measuring what is controlled by the circadian such as core body temperature and cell secretions such as hormones like cortisol and adenosine. The natural shape of the circadian cycle is important for this research as it defined timing periods for test taking.

It is believed that the rhythms naturally oscillate on a period slightly longer than 24 hours. 24 hours is commonly stated as the periodicity but longer cycles have been cited up to and including 25 hours (Strauss, 2003 , p3) however, Charles Czeisler's team completed a definitive study demonstrating that the human circadian clock runs at about 24 hours, 11 minutes (Rosekind, 2000,Czeisler et al., 1999) which is closer to plants and other species used in research to define this cycle length. Having the circadian with a natural period length slightly longer than 24 hours means, in the absence of anything else, the clock of the circadian would not match the external clock of 24 hours per day suggesting that if this is all that governed our sleep activity we would daily drift forward on the timing of when we slept and woke and were active.

3.1.4.1 Entrainment

The solution to drift is that the internal clock must be reset or entrained each day to match the length of the external day-night cycle and this is done by monitoring light conditions. It is not our normal receptors, the rods and cones that are used but rather, dedicated photoreceptors in the retina that feed signals directly to the SCN. (Berson et al., 2002). The cues, called "Zeitgebers," (Strauss, 2003) keep the circadian clock set to the appropriate time of day. Besides daylight, common Zeitgebers include meals and work-rest schedules and other things that form repetitive daily patterns.

Contemporary research is now exposing the influence of aging on maintaining the circadian rhythms. Prior to Berson's (Berson et al., 2002) work, the SCN maintenance of the circadian rhythm from a dedicated feed from the eyes was not known to exist. Now, research is uncovering the influence of aging eyes and how deteriorating eyes filter out more and more of the blue part of the light spectrum needed by the SCN to maintain proper circadian regulation (Mainster and Turner, 2011). We know that sleep changes as we age but these new emerging age-related factors are helping with understanding of the significance of all aspects of health associated with sleep disruption. In most western work places this is an important issue presently as baby boomers increase the average age of workers. This was a factor of interest for this research.

Small shifts in the circadian rhythm normally occur to accommodate our changing schedules and as seasons change. Air travel has allowed fast and large changes to our environment but human physiology is not designed to change as quickly as we can move our bodies around the globe. When we make large shifts in our schedule such as crossing time zones or changing work from a day-shift to a night-shift we need to adjust to the new conditions. To adjust, shifts in the circadian cycle are required to keep our internal feedback loops controlling sleep and wake-alertness periods synchronized with the new schedule or time zone. This takes some time to fully adapt and the amount of time depends on the number of hours one's schedule is shifted, and the direction of the shift (Waterhouse, 1999).

During a transition, the conflict between the internal circadian and external clocks is not synchronized resulting in what is commonly known as 'jet lag' when travelling. It affects all rhythms and can produce effects similar to those of lost sleep. Trans-meridian flights of at least three time zones can result in significant circadian rhythm disruption. Thus the physiological time and local time can vary by several hours. When flying in a westerly direction the traveller's day is lengthened. The traveller's day is shortened when flying east, against the direction of the sun. Symptoms of jet lag are usually worse when flying from west to east as the day is artificially shortened. Strauss (2003, p3) believed "it takes about one day for every time zone crossed to recover from jet lag. When circadian disruption and sleep loss occur together, the adverse effects of each are compounded". Others (Rozell, 1995) have provided a range of recovery times for travel east versus west from one day per hour difference going west and one and a half days per hour of difference going east to 3 to 7 days for westward travel, and 5 to 14 days after an eastward flight (Gillingham, 2012) to 1 day per hour time difference going east to 1 hour per 1 and half time zones going west (Logan, 2012). The ranges given are indicative of actual recovery times being tied to many individual factors that will influence the time for re-adjustment making all of these suggestions possible. The relevance for this study is associated with the shift-workers as it at least temporarily, causes de-synchronization in the same fashion as jet lag and adversely impacts sleep. Some of this research has also fed into the debate on which direction shifts should rotate to minimize impact on fatigue. Additionally, there is a persistent presentation within the workforce of individuals who have travelled and returned from business and personal trips with jet lag. Desynchronosis was an important factor associated with this research as the circadian cycle and its stability were central to one of the research studies with shift workers.

3.1.4.2 Hormone Secretion

Sleep cycles follow our circadian rhythms and it has been found that the natural hormone melatonin plays a large part in mediating sleep. Its production is linked to the light-dark conditions. As darkness falls, enzymes in the brain controlled by the SCN stimulate the release of melatonin from the pineal gland into the blood. Melatonin levels are detected by receptors in the SCN to induce sleep. Melatonin is expelled from the body according to wake time and light detection by the SCN interpreted as start of a day. Other secretions of the body such as serotonin which is linked to mood, prolactin which is closely related to growth hormones and prostaglandin which is secreted at multiple cells in the body and linked to the control of blood pressure have influence often associated with the sleep cycle (Cauter et al., 2005). The functions of many organ systems are linked to the circadian cycle including the endocrine (all hormone secretion), renal (kidneys) and alimentary (gastric acid secretion) systems. Disruptions that affect the circadian rhythms affect many aspects of a properly functioning and healthy body. The full influence of all hormones such as melatonin are not fully understood but are strongly tied to strength of the immune system. It is an important aspect of understanding the relationship between sleep deprivation and worker performance and was a key factor of this research especially for a shift worker population.

3.1.4.3 The Pressure to Sleep

Sleep is further regulated by the SCN directing the secretion of molecule markers and monitoring their levels with time spent awake. A key marker is adenosine (Figure 9). Its sensed build-up causes increased pressure to sleep.

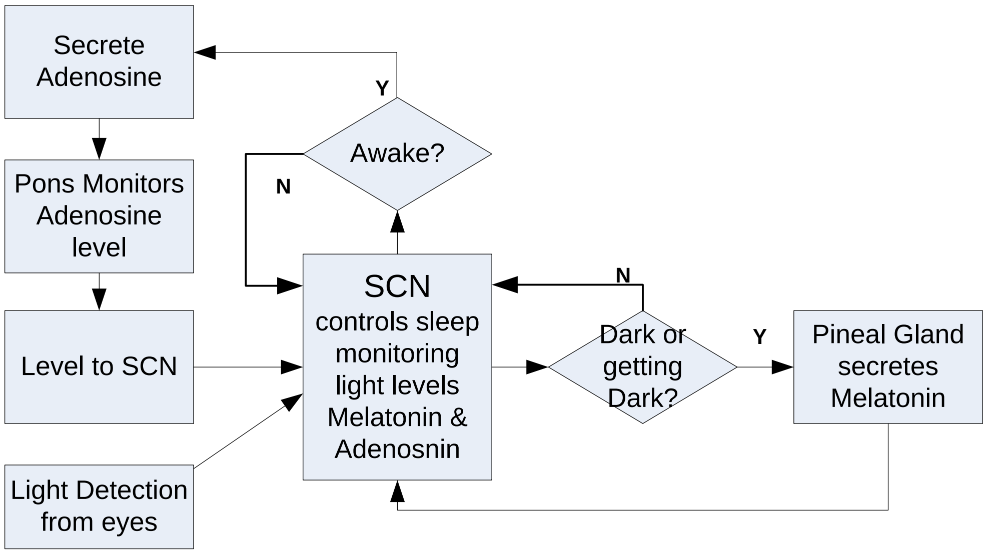

Figure 9 - The circadian cycle maintained with feedback loops

Adenosine is secreted into the blood from the point of waking and is detected by neurons in the brainstem which transmit to the SCN as a kind of sleep-signal based on hours one has been awake and how much sleep debt has been tracked. This molecule and its control, sometimes referred to as the homeostatic regulation of sleep, reinforces the circadian cycle and makes it increasingly difficult to resist sleep after as little as 16 hours awake. The 16 hours is tied to the 8 hours we need to sleep and the 24 hour limit in a day. This is the foundation for several well referenced research projects which compared hours awake to alcohol intoxication starting with Dawson and Reid (1997). Adenosine levels drop while we sleep and then rebuild when awake. If inadequate time is allowed for sleep, there will be residual adenosine starting the next day. This is an important feed-forward factor in more complex models of sleep that attempt to account for cumulative effects of inadequate nightly sleep.

Another important factor for this research is that caffeine has been identified as a block on the adenosine receptors preventing proper nerve cell function for the prevailing levels of adenosine. Instead of normal adenosine binding to the receptors causing nerve cells to slowdown and blood vessels to dilate, the caffeine binds and speeds up the cells (Boutrel and Koob, 2003). Nehlig's research concluded caffeine "…cannot be considered a `pure' cognitive enhancer. Its indirect action on arousal, mood and concentration contributes in large part to its cognitive enhancing properties” (Nehlig, 2010, p1). As the caffeine is expelled, the receptors detect the true level of adenosine and the effect can be a rapid large swing in performance sometimes referred to as the ‘caffeine crash’. In the work place, caffeine consumption is common and must be accounted for in any discussion on alertness and performance as with this research.

3.1.5 Sleep Required

Research continues on the fundamentals of sleep but relative to the understanding sleep’s contribution to workplace fatigue-impairment, there is ample understanding and an abundance of literature that illustrates the average amount of sleep needed by healthy adult populations is 8 hours per 24 hours for full restoration (Van Dongen et al., 2003, Russo, 2005, Rosekind et al., 2000, Wehr et al., 1993, Roth, 2006, Dinges et al., 1996, Levine et al., 1988). Several studies suggest we get seven or less hours of sleep per 24-hour period resulting in a sleep deficit (Dinges et al., 1997). There is also understanding that population norms do not necessarily fit any individual. Additionally, the sleep required is not static. There is a basal need and a nightly need which may include additional hours for debt repayment, illness, stress, or body repair. However, as sleep need is believed to be a distribution about 8 hours individual basal needs will fall within a range above and below 8 hours. It is often listed by sleep centres as a need in the 7 – 9 hour range (National Sleep Foundation, 2012) this perhaps contributes to confusion for workers who may conclude their sleep taken is adequate as long as it is in the range. Others depart from giving a number. Jim Horne of Loughborough University's Sleep Research Centre in a BBC report suggested "The amount of sleep we require is what we need not to be sleepy in the daytime" (BBC News, 2012, p2).

Regardless of individual need, sleep is a common physiological function as necessary for survival as food and water. Researchers believe sleep occurs in all mammals and probably all vertebrates. As little as one hour short of an individual’s natural need is marked and retained as deficit. There is still some debate associated with sleep debt and how long it is carried and whether it must ever be repaid. Indeed, without repayment, concerns pertain to long-term functioning and normalcy whilst carrying sleep debt (Horne, 2004). Ideal sleep becomes a very individualistic need but for healthy human populations it is governed by boundaries which shift for humans from birth until adulthood. There is also another important component relative to meeting one's sleep requirements and that is the quality of sleep.

Quality of sleep is a measure of how closely an individual follows an uninterrupted natural sleep pattern. It has two components; the first is associated with the number of wake periods and sleep loss due to wake periods and the second is associated with what stage of the sleep cycle is lost. It is normal to wake during the night and multiple times, but excessive waking or movement during a night's sleep detracts from total hours trying to sleep. Additionally, every waking is an opportunity to stay awake which occurs if one does not fall asleep again.

While contemporary guidelines call for the need of eight hours of sleep per 24-hour period, it is not universally agreed how that sleep should be taken. There is growing evidence that trying to take the sleep in one session is a product of the modern world with artificial light and fixed work schedules but is not a normal mammalian trait. Evidence suggests our ancestors without the artificial night environments as recently as 100 years ago would commonly sleep in at least two sessions with activities between sleeps (Hegarty, 2012). There is suggestion that prior to broad application of artificial light, the average person slept about 9 hours per 24 hours but some such as Horne (2008) refute this suggesting there is simply no credible evidence for it and what is presented as fact is a misunderstanding of an early 20th century study on sleep of children.

Sleep disorders affect quality of sleep and this may not be a small issue. Many individuals estimated as high as 20% of the total North America population (HealthCommunities.com, 2012) have sleep disorders (more than

70 described) which prevent them from getting quality sleep. Obstructive sleep apnoea (OSA) is a common and serious one characterized by a repeated collapse of the upper airway during sleep, causing breathing pauses. When this occurs, OSA sufferers wake often, preventing any sustained proper sleep. It has been characterized as a night of 1000 naps leaving these individuals in a perpetual world of excessive need for sleep. Sleep disorders are relevant for this research as their commonness should statistically cause them to surface in research in operational settings. Those with sleep disorders are also a particular group with higher risk of accidents (Garbarino, 2008) in all activities and most importantly, are treatable once identified.

We crave sleep when deprived of it, and from animal studies such as those conducted on rats (Rechtschaffen et al., 1983) we know continued sleep deprivation can ultimately be fatal. If individuals do not obtain their nightly quotient of sleep then feelings of sleepiness can result when awake. If sleepiness comes at an inopportune time relative to mental or physical demands, the resulting impairment can lead to errors. If errors occur in a critical, demanding, or high-risk activity the outcome can be very consequential. Too little sleep can be acute or chronic and can contribute to excessive daytime sleepiness (EDS). Anyone for any reason not meeting their sleep requirements at night will exhibit a level of fatigue-related impairment. Understanding levels of impairment and the source was an important factor for this research. As such, understanding individual sleep/rest and wake periods relative to a normal circadian cycle was required.

3.2 Fatigue-impairment in the Workplace

3.2.1 Introduction

The word ‘fatigue’ is in standard everyday use, describing a number of conditions varying from general lethargic states to specific work-induced aches, burning, and pain of one's muscles. Fatigue has a physical and a mental component. Physical fatigue is a daily phenomenon clearly known to most after periods of heavy physical work or exercise which pertains to the depletion of the energy required to sustain performance. However, physical fatigue also carries a strong mental component as in a sense of more required effort (Enoka and Stuart, 1992) when fatigued which can inhibit one's ability to continue functioning at the level required.

When we introduce the concept of impairment related to fatigue and as used throughout this research we are implicitly including the mental component which can manifest itself both as sleepiness or somnolence as well as the general symptoms that arise in this state. Some psychologists call this reduced willpower or ego depletion, which occurs when one's energy drops and the mental strength required for self-control is impaired or depleted (Baumeister and Vohs, 2007). It practically manifests as decision fatigue (Tierney, 2011). As fatigue can have multiple causes and is not directly observable, it is usually a reported symptom of a cause. Berrios believed that fatigue is often confused with feeling fatigued which is a companion to various medical and psychiatric conditions which when understood can assist in defining fatigue (Berrios, 1990). Others have identified the space between psychiatry and medicine as a space where disorders fall that are not easily classified as either mental or physical illnesses. This space is inclusive of chronic fatigue (Gruber MD et al., 1996). Fatigue has also been described as a level of decreased consciousness (Giannini, 1992). If extreme sleepiness occurs, an individual may unknowingly lose consciousness temporarily in a condition called a micro-sleep.

In addition to inadequate sleep which feeds fatigue-related impairment, there are other factors that feed this condition including time on task, type of work and how stimulating it is. Other personal factors can also lead to fatigue such as illness, stress, disease, and depression. Fatigue can have just a daily component as in a lack of sleep, or a weekly component due to a temporary minor illness like a cold or flu or a longer more chronic component resulting from sleep-related disorders and other diseases. Some of the diseases include anaemia, allergies, depression, hypothyroidism, infections, diabetes, and heart disease. Recent research (Dyken and Im, 2009) also cites a very high incidence of sleep apnoea with strokes. Other health issues have been linked as well, including heart disease, diabetes, and sexual problems. What is not understood is whether the sleep disorders and diseases are simply coexisting in predisposed individuals or whether there is a causal relationship. Other cited non-medical issues causing fatigue include poor dietary practices and caffeine. Caffeine, previously presented as a stimulant to fight off fatigue paradoxically can lead to fatigue if too much is taken as it is known to increase heart rate, blood pressure, and activity level. (WebMD, 2012).

In all cases, impairment from fatigue can show as slowed reaction times, loss of attention and inability to focus on details. This can be dangerous when performing any tasks that require vigilance, such as driving. It can also be disastrous if real-time mental agility is needed drawing on memory or other learned responses that are now impaired.

In general, fatigue can be considered to be a longer term condition than sleepiness but inadequate sleep versus individual need is considered to be a primary cause of fatigue-impairment (Beaulieu, 2005). Strauss (2003) who has operationally studied the effects of fatigue in real world work environments, defined fatigue as “a non-pathological state resulting in a decreased ability to maintain function or workload due to mental or physical stress” (ibid, p1). Strauss further believed that there are “…two major physiological phenomena that have been demonstrated to create fatigue: sleep loss and circadian rhythm disruption” (ibid, p1) previously introduced as key variables to understand relative to an operational study.

Tiredness is certainly a part of the discussion on fatigue as in a new medical term in use, tired-all-the-time (TATT) syndrome. TATT is driven by a world which no longer rests, symbolized by services, entertainment, communication, and work schedules to satisfy around-the-clock service leaving one chronically tired. Tiredness has been used to relate to both sleepiness and fatigue.

The consequence of living in a world which no longer rests has created this new problem for humans. The magnitude of fatigue-related impairment and lack of sleep did not exist at this level 100 years ago. Van Dijk and Swaen (2003, pi1) believed “approximately 20% of the working population report symptoms that fall under the concept of fatigue. Other surveys have reported prevalence rates of fatigue varying from 7% to 45%...”

There is some opinion that there has been a steady decline in the average sleep taken each night over the past 100 years. Jim Horne, in Brennan, did not agree and stated “There is no evidence the Victorians slept longer. I suspect many people slept less in the early 1900s because they worked longer hours and had poorer sleeping conditions” (Brennan, 2012, p2).

Indeed new historian work is revealing more of a changing sleeping pattern than validating declining sleep (Hegarty, 2012).

3.2.2 Identifying Fatigue-related Impairment

As fatigue is a non-specific symptom, self-assessment and diagnosis is the only practical way of identifying mental impairment due to fatigue. This is usually expressed as feelings individuals have come to know with respect to their own condition prior to the onset of sleep. They can include:

Somnolence (Physical Signs)

- Generally feeling tired and desiring sleep;
- Nodding off, dropping head or falling asleep;
- Continuous yawning;
- Having trouble keeping one’s head up;
- Feeling lethargic;
- Having eyes close or go out of focus by themselves; drooping eyelids;
- Rubbing eyes;
- Digestive issues;

Cognition (Mental Signs)

- Slowing reaction times;
- Reduced vigilance and poor concentration;

- Poor decision making and judgment; accidently doing the wrong thing or not doing the right thing;
- Becoming fixated;
- Unusual forgetfulness; difficulty remembering tasks being done;
- Inability to recollect the events of the last few minutes;
- Wandering, disconnected thoughts; lapses in attention;
- Inability to focus or learn;
- Becoming apathetic;
- Failure to communicate important information;
- Reduced hand-eye coordination;
- Reduced visual perception;

Hormonal (Emotional Signs)

- Bad moods and irritability;
- Reduced capacity for interpersonal communication;
- More quiet or withdrawn than usual;
- Loss of motivation or willpower to do a task well;
- Low energy levels.

These may also just be symptoms of underlying health conditions as opposed to poor sleep. When we refer to fatigue-related impairment we are specifically talking about a decrement from 'normal' cognitive functionality. It is also important to note that our levels of cognitive capability that are tied to fatigue naturally change over the course of a day according to the circadian cycle. We do not maintain alertness at 100% continuously during a day so impairment must be expressed relative to normal levels in cognitive performance which naturally cycles over the course of a day.

Those factors which are associated with measurable cognitive functions such as reaction times, memory, and learning were of interest for this research. In the absence of any other factors explaining these symptoms, experiencing changes in any of these may mean an individual has experienced fatigued-related impairment. However, self-assessment of these factors can be impaired too. For instance, Rosekind (Rosekind et al., 2000, Rosekind, 2000) found sleepiness had two distinct components; one physiological and one subjective. The physiological component is driven by sleep loss with one's drive to sleep fuelled by getting less sleep. The second component is one's assessment of how sleepy they feel. Depending on circumstance there can be large variation in physiological need and subjective sleepiness. The nature of work activities, the work environment, and conditions can exacerbate or conceal sleepiness leading to an inaccurate estimation of one's own alertness. This might have significance on completing tasks properly or the judgment exercised in a task required to ensure it is done safely and correctly.

3.2.3 Consequences of Fatigue-related Impairment

The consequences of fatigue-related impairment may include a general diminishment in certain cognitive functions, affecting tasks performed. It is most evident if someone falls asleep but most often the result is more subtle or undetectable such as slowed reaction time or loss of attentiveness. If someone falls into a micro-sleep, they may not even be aware they were asleep.

If the tasks being performed are not safety critical (reading a book) no concerning consequence is expected. If the task being performed is safety critical (driving), slowed reaction time, nodding off, or drifting attentiveness can have deadly consequences. Worker fatigue-impairment has been found in numerous work cases to be associated with significant loss of life and property. Many researchers and institutions such as Transport Canada (2007) and the National Transportation Safety Board (1999b) in the USA report fatigue has been cited as contributing to or causing some of the world's most visible accidents including the grounding of the Exxon Valdez (National Transportation Safety Board, 1990). Restarting a large oil refinery after maintenance construction work led to an explosion killing 15 and injuring another 180. Fatigue-impairment was a factor (U.S. Chemical Safety and Hazard Investigation Board, 2007).

The National Transportation Safety Board (1999a) of the USA cites 21 percent of reports in the Aviation Safety Reporting System are related to fatigue; 1.6 percent of all highway crashes (100,000 per year) and 3.6 percent of all fatal crashes (1,500 fatalities per year) are caused by drowsy drivers; 16 percent of critical marine vessel casualties and 33 percent of personal injuries and; that fatigue is a major factor in transportation accidents including railroads.

Heavy truck accidents are a particular problem with an estimated 30 – 40% having truck-driver fatigue as a contributing factor (ibid). The incidents do not need to occur at work but can be caused by the work. Co-workers of Brent Hershman petitioned to limit hours of film shooting after Brent was killed after he fell asleep at the wheel driving home after a 19-hour day (Hassen, 2002, p1). 'Brent's Rule,' called for a 14-hour shooting limit on film and television sets". Canada's first legal challenge [Gartner v. 520631 Alberta Ltd.] associated with employer liability for fatigued employees' injuries involved an Alberta worker who was severely injured with permanent disabilities from an accident driving home after a 19-hour work shift, which exceeded the Alberta Employment Standards Code which limits workers to 12 consecutive hours of work a day. The case ultimately failed the 'chain of causation' but in hearing the case, Canadian employers were served notice they could be held liable for a traffic accident caused by an off-duty, fatigue-impaired worker under different circumstances (OHS Insider, 2010). This is further verified by two USA cases where the courts found employers negligent for letting workers drive home after long work shifts which resulted in deaths driving home. [Escoto v. Estate of Ambriz]; [Faverty v. McDonald's Restaurants of Oregon, Inc.]. In both cases the courts concluded the employers knew the workers were exhausted and it was negligent to not recognize the risk of allowing the workers to drive home (ibid).

On October 19, 2011 at 0400, returning home from an emergency patient relocation, two paramedics from BC's Ambulance Paramedics died after their ambulance plunged into Kennedy Lake. One was found in the back apparently sleeping and it is believed the driver fell asleep and left the road.

(Quinn, 2010). A quick search for ambulance accidents reveals numerous similar cases with explanations associated with long shift hours and on-call schedules 24-hours per day. Additionally, they are individuals required to make quick assessments and take proper action to save lives, which is most important for all the construction workers requiring their services due to the high accident rate in construction. The accident statistics and others from shift-working drivers such as the police and taxi drivers may be included with their safety performance statistics but many others are not especially if in the commuting phase.

This is especially important for construction workers who frequently drive to and from job sites in inclement weather and long distances. Construction workers of British Columbia's power utility BC Hydro, reported at least one vehicle death in 2010 and multiple accidents from its membership due to early morning starts (0330) required for long commutes to get to remote sites for contracted work. Workers believed their fatigue was to blame as no shift-schedule allowance or relief was given for the long commutes required. This resulted in extremely long days in some cases perhaps contributing to commuting accidents. However, the significance in this case and throughout many jurisdictions is that for the commuting construction workers they would not show up as a workplace accident statistic but rather a roadway motor vehicle accident statistic. There is no monitored statistic for workers injured or killed commuting to or from work (Walters and Bolt, 2009).

Fatigue-impairment safety concerns in the workplace are the focal point but another insidious aspect of fatigue is that it does not recognize reporting boundaries. Such examples reveal that fatigue-impairment has safety consequences in our personal and work lives.

3.2.4 Fatigue-impairment Operational Research

There is evidence that workplace fatigue was of some interest as early as 1929. Page (1929, p137) wrote, "Fatigue, to the average man, means simply the state of being tired. He needs neither management engineer nor industrial physician to tell him when he is fatigued, for that is a matter of his own immediate, knowledge." Page stopped short of trying to differentiate the consequences of being tired as potentially having significant personal or organizational consequence. We might know when we're exhausted but fatigue as a form of impairment is not well understood by the average person. Nor is it controlled or properly acted on. This is rectified with understanding both the magnitude of the workplace problem and the risk it presents.

Operational research and managing fatigue at work has increased in the last 10 to 15 years, based on research papers available. Operational research in Aviation (Flight Safety Foundation, 1995,Rosekind et al., 1996) has led in understanding the levels of risk associated with fatigue-impairment and establishing operating guidelines for fatigue-impairment to improve safety. This has been followed in other sectors including Medical (Howard et al., 2002), Trucking (Wylie, 2005), Rail (Gertler et al., 2006), and Construction (Powell, 2007). Other industrial sectors (Petroleum Industry Training Service (PITS), 2003) have recognized that fatigue is a key component of workplace safety and employee wellness and therefore a factor which must be managed. In sectors such as Transportation, Aviation and Medicine, a lack of proper sleep has been identified as a key contributor to accidents over the past 30 years. Much has been learned from their operational investigations and studies where the consequences of succumbing to fatigue-impairment are more direct and potentially dangerous to the public.

When the amount of sleep taken is coupled with other individual factors the desired level of worker cognition may be impaired, and in severe cases, individuals may be put at unnecessary levels of risk due to the consequences resulting in accidents.

3.2.5 Fatigue-impairment in Construction

"Construction is a high hazard occupation" (National Institute for Occupational Safety and Health, 2007) which, coupled with factors such as poor working conditions due to inclement weather, mobile equipment, different work sites necessitating travel as well as changing and demanding schedules requiring additional work hours, inadequate training and impaired workers contribute to a risky work environment (Chapter 4).

Construction workers not only carry with them normal societal pressures and stresses but also can have project schedule and work pressures added. Construction workers need to be alert at all times to perform their work effectively, efficiently and safely. Pressure from construction project schedules leads to working long hours and sometimes working around-the-clock. This creates the physiological challenges resulting in performance-impairing fatigue, leading to higher worker risk. Continuous improvement of worker performance, productivity, and safety is possible but starts by acknowledging and managing these factors. Rosekind (Rosekind et al., 1996, p157) noted, ignoring them "…can lead to decrements in performance and capability as well as the potential for incidents and accidents that can result in tremendous societal and individual costs".

The potential impact of construction worker fatigue-impairment has not been studied well. Dong (2005) looked at accident data and hours worked from the construction industry in the United States and compared it to other industries. Specifically she was looking for a correlation between risks of work-related injuries and hours worked. Dong was able to show that "construction workers had a higher work-related injury rate than did their nonconstruction counterparts" (ibid, p332). The study also found when workers put in more than 8 hours a day or 40 hours per week their injury risk increased slightly. This risk doubled after 50 hours per week. Other factors tied to higher injury risk were shift work, starting work early and ending late.

Kecklund et al (2001) investigated how 'double-shifts' (15.5 hours) affect sleep, fatigue and self-rated health. The study was carried out on male construction workers of which 80% were long-distance commuters. The results showed that sleepiness, and to a certain extent, mental fatigue increased during double shifts and accumulated across days. It was concluded that a shift system involving double shifts has a negative effect on fatigue, recovery, and health-related wellbeing.

Powell (Powell and Copping, 2010, Powell, 2009) conducted an extensive study on a group of construction workers associated with a large multi-faceted construction project in Canada. For the first time, actual sleep data was collected from construction workers and performance decrements estimated. The study found the workers on average were getting inadequate sleep resulting in decrements in performance solely due to sleep by about 10 percent. More relevant for accident prevention and lost productivity, several workers were found to be working at an impairment level equivalent to a blood alcohol concentration equivalency of 0.08 percent. Fatigue-impairment was exposed as a problem in construction and another concerning source of workplace impairment. Recommendations included establishing a fatigue-management programme within construction focused on the most hazardous or safety-critical activities as well as screening mechanisms for fatigue-impairment.

In 2009, an extensive report was produced on the causes of accidents in the construction industry (Walters and Bolt, 2009). The motivation was similar to this research; to help make improvements in an otherwise poorly performing sector. The problem is similar to what the other sectors face. Symptoms of fatigue-impairment are all that is reviewable and beyond establishing the linkage between events and time-of-day, fatigue-impairment and its associated risks can only be speculated on and suggested as contributors. The actual data from research becomes scarcer the further the impairment is pursued into each sector. There remains a vacuum for research to understand contemporary issues related to practical aspects of measuring fatigue-impairment and dealing with fatigue in the construction workplace.

3.3 Summary

Within this chapter, the concept of fatigue-impairment was introduced. Fatigue-impairment is maintained in a balance between wake and sleep cycles. Fatigue-impairment has been shown to be not only a common workplace issue, but also a consequential factor that has contributed to a huge number of accidents and deaths. Most recently, there is a movement to employ technologies that detect potentially the most serious symptom of severe fatigue and that is falling asleep. In some sectors such as transportation, the impact of falling asleep is immediately evident and consequential. In other sectors, the potential impact may lie dormant until called upon to wreak havoc with impaired minds. Detection of symptoms or even implementing plans to deal with identified fatigue-impairment is not obvious.

CHAPTER 4 – MANAGING FATIGUE- IMPAIRMENT as a HUMAN FACTOR in SAFETY

4.0 Introduction

Many safety programmes focus on hazard identification and risk management. Significant focus has recently been put into understanding and training associated with human factors as a means to improve safety (HSE, 2009). This chapter will review how we account for risk and accidents associated with fatigue-impairment as well as review the mechanisms by which fatigue-impairment manifests as a human factor in safety. A new systems-based approach for addressing the risk of worker fatigue in the workplace is emerging and is reviewed.

4.1 Risk in the Workplace

An important aspect of safety management systems is identification of hazards and ensuring workers are protected from them. This typically follows a programme of identification, classification, awareness, and mitigation. If hazards are identified early enough, work processes can be designed to avoid or eliminate the hazards, thus reducing risk for the worker. Effective training programmes go beyond awareness and help prepare workers to not only be alert for hazards but to take proper safeguards when encountered. Mitigating steps are intended to be in place to protect workers from all hazards.

4.1.1 Hazards

As it pertains to the human element a hazard can be thought of as something presenting a level of increased threat to one's health and safety. In most environments hazards are classified as potential for harm or injury and associated with a theoretical risk based on a probability assessment of one being exposed to the hazard and the resulting severity of occurrence of the incident. Hazards or potential for harm can be very broad from biological, to chemical, mechanical, electrical, natural or automated systems (MacCollum, 2006). In a construction environment, hazards are constantly shifting as the projects evolve. They can be as diverse as second-hand smoke from co-workers to rebar moved overhead by a crane, to working at height. As projects evolve, there are typically different activities at a site, which change from the initial site preparation which may include demolition and excavation through to peak construction periods where multiple activities occur at once. Constantly changing landscapes around construction sites is a large part of the safety challenge. Another aspect of hazard management is preparing workers to take appropriate action when faced with hazards (ibid). A fundamental safety approach in site or job design is to attempt to isolate workers from potential hazards. If this cannot be done, then the worker must be protected with appropriate protective equipment. Assessments including possible consequences and probability of occurrence determine the actions to be taken. By isolating workers from hazards or limiting their exposure reduces the risk on site. It is important to note that risk is managed but seldom eliminated.

MacCollum (ibid) believes that the industry needs to borrow a concept from systems engineering and assess the probability of failure of mitigating steps. This goes beyond risk assessment to incorporate a reliability study which, in part, provides a quantitative assessment of a failure mode for identified hazards. This is to establish that the mitigation put forth for the elimination of, or protection from each identified hazard, will be effective and to what degree it will be effective. MacCollum's suggestion is that at the end of the hazard-control hierarchy, consisting of four steps, to eliminate the hazard; guard the hazard; apply safety factors to minimize the hazard and apply redundant parallel safeguards, a fifth step should be added to calculate the probability of eliminating or controlling the hazard.

A contemporary view is that fatigue itself is an operational risk (Lerman et al., 2012) which requires management.

4.1.3 Fatigue Risk Management

Strong construction site safety programmes attempt to maintain awareness on updated lists of hazards and take appropriate actions to help mitigate potential harm. The relevance of this for this research is that despite best efforts to put in place a system to manage risk on sites and alert, inform, train and protect workers, there are risk elements that are not site controllable. Some of the risk elements are imported with the workers, such as stress and fatigue-related impairment. In part this is because their contribution to the risk profile is not known or understood; in part it is due to a resignation that inadequate sleep-induced fatigue-impairment is a fact of life (Powell and Copping, 2010). It has now been found to be a key component of the risk profile facing construction workers of British Columbia (ibid). Another aspect of this risk element is that it too is constantly shifting and indeed the job site and work can be a contributor to individual stress and lack of sleep. Over the course of a day, the circadian cycle will also influence and change the level of workers' alertness and their effectiveness. In this circuitous fashion, fatigue is born and dies. Something more than understanding is needed at the job site to break the cycle.

4.1.4 Managing Fatigue

Fatigue risk management systems in the workplace are not different in concept to a general system managing operational safety and can be incorporated into overall health and safety management. Such systems usually start with a policy statement to ensure all workers understand the importance of the programme and responsibilities of stakeholders. This is then followed by developing a plan focused on addressing the risks of fatigue in the workplace; which roles and activities are most susceptible and vulnerable; and what controls might be needed to mitigate the risks.

Prevention of accidents associated with excessive fatigue is difficult especially when the condition of workers is unknown. The most troubling aspect of fatigue-related accidents is that the fatigue may have originated due to work. Organisations typically have little control over what happens in a worker's personal life after work. They cannot schedule or control an individual's sleep anymore than they can control diet, exercise or lifestyle choices. However, employers can ensure that their work schedules do not push workers over the top and otherwise prevent adequate and proper rest and sleep. Employers can assist with training to help workers understand the importance of sleep and lifestyle. In sectors where fatigue-management was recognized as a priority, such as transportation, including heavy trucking, rail and aviation, regulations emerged to set minimum acceptable standards for work and schedules. Organisations responded with what is commonly known as a Fatigue Management Plan (FMP) to illustrate compliance and address the practical aspects of resource planning required to ensure compliance.

FMPs are not built around any actual measurements of individuals' fatigue but rather attempt to take a pro-active step to prevent or reduce fatigue-impairment by limiting continuous work hours and allowing adequate time from work for adequate rest and sleep. Most importantly FMPs ensure that work schedules are not contributory to the workers' fatigue.

FMPs might address some common elements:

1. Staffing is reviewed to maintain adequate balance with work requirements;
2. Shifts are reviewed with consideration of time-of-day, length and rest periods;
3. Check-in points with fitness-for-duty and alertness monitoring;
4. Review of work environment, workplace design and ergonomics;
5. Fatigue and personal sleep management training.

4.2 Evolution of Fatigue Management Plans

Different sectors place different emphasis on the importance of fatigue management planning. It is core to the aviation industry and transportation industry in general. In more complex work environments such as aviation, where workers are in safety critical roles and may be crossing multiple time zones with circadian disruptions, great care is taken to ensure adequate rest. This includes prescribing on-job napping with time parameters around the length of naps to prevent workers entering into stage 4 deep sleep. This is all done to a plan but even the best of plans can go wrong. In April, 2012, the Canadian Transportation Safety Board filed its findings relative to a trans-Atlantic flight from Toronto to Zurich, Switzerland, during which Air Canada reported severe turbulence caused extreme altitude changes on the flight causing passengers to be injured whilst thrown from their seats (Gillis, 2012). The findings indicated pilot fatigue and broken rules controlling mid-flight napping were among the causes of the January 2011 incident. The rule limited time napping to prevent pilots entering stage 4 deep sleep.

There is an increasing understanding that while Fatigue Management Plans help, they are not enough to fully address workplace fatigue. They typically only address the work-side of fatigue management and lead to a realization that more than establishing guidelines for work and rest must be done to address fitness-for-duty and long-term health. They fall short of addressing the concept of fatigue as a hazard in the workplace and therefore do not embrace failure modes associated with the FMPs. To fully address the elevated risk from this hazard means that it must be integrated into safety management and it means that individuals must share, as partners, the full scope of managing fatigue as all time away from work is in their control.

4.2.1 Fatigue Risk Management Systems

A Safety Management System can be thought of as a formal framework for integrating safety into every day operations. It will typically include defining safety policy, setting safety goals, implementing plans, conducting risk assessments, setting responsibilities and authorities, defining rules and procedures, training, monitoring, reporting, reviewing, and evaluating processes for corrective action and improvement. Such a system is shown in Figure 10.

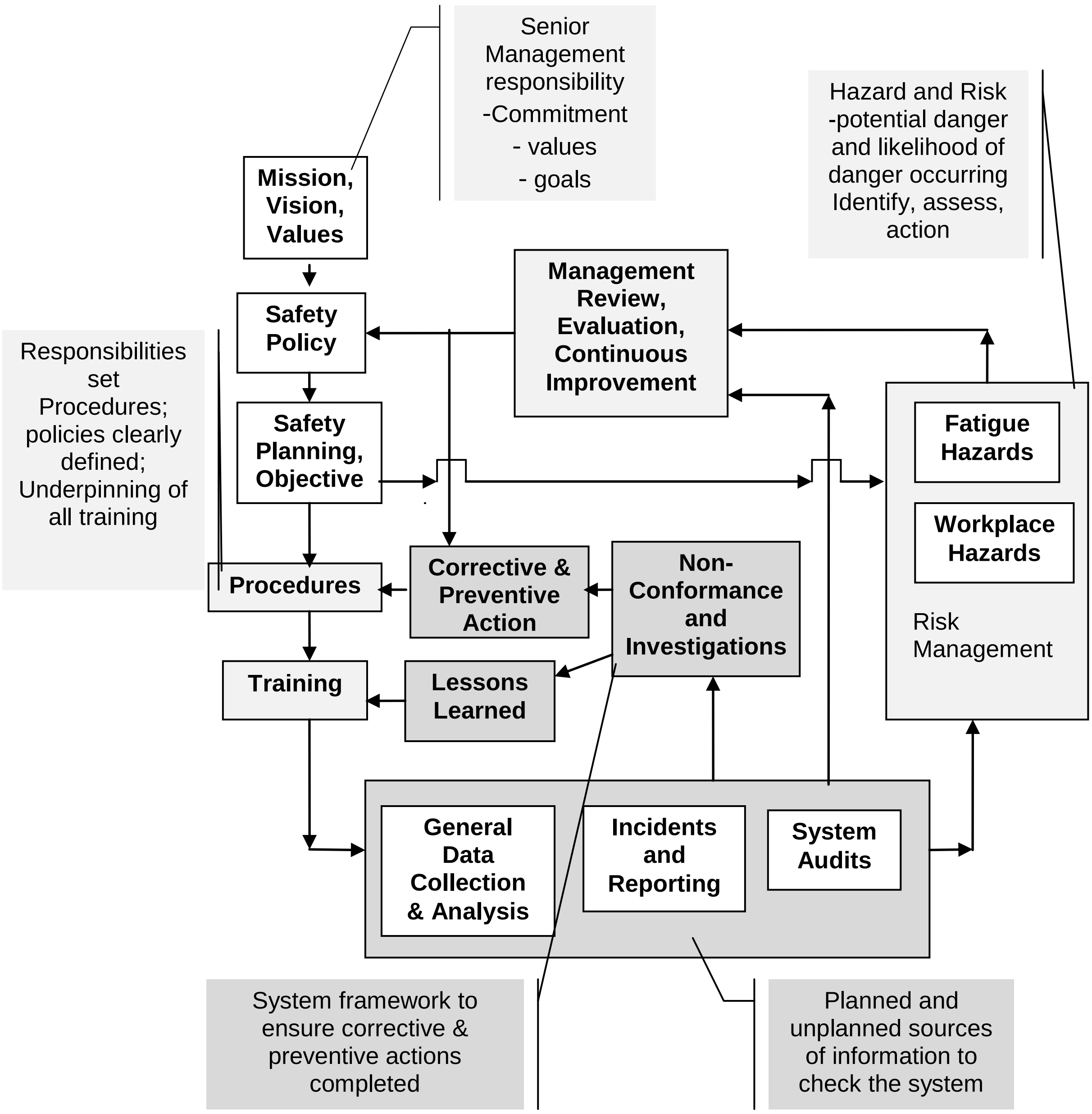

Figure 10 - Interconnected activities and flow making up a Safety Management System

Recognizing worker fatigue contributes to elevated risk means that fatigue risk management can be integrated into this system to measure, monitor, mitigate, and manage. It means that worker fatigue is dealt with in a manner commensurate with the level of risk it presents. For instance safety critical roles such as heavy equipment operation would be assessed and treated differently from office clerks. A system such as shown in Figure 10 has a goal of constant review, acting and measuring for continuous improvement.

A 2012 Presidential task force on fatigue risk management (Lerman et al., 2012) found that one departure from a standard safety management system would be the inclusion of a new area specific to fatigue, sleep disorder management. This is obviously tied to how common sleep disorders are in society and the research which has shown the clear association with increased risk with those who suffer sleep disorders. Sleep disorders are all treatable, so in addition to reducing workplace risk, there is a social benefit as well. However, it requires full support and participation of employees to be effective.

Adopting a safety system approach for managing fatigue allows organizations to evolve past prescriptive, plan-only positions which are really only part of the equation. The aviation industry once again leads in this area with new legislation that will force the adoption of Fatigue Risk Management programmes (FAA, 2010).

A challenge that organisations face with fatigue risk management is enlisting full support of all workers. Without other means, procedures, training, and adherence to procedures are required for monitoring, detecting and reporting fatigue.

Despite best efforts to design a work space, train workers, and maintain vigilance towards hazards, accidents still happen. An appropriate response from a safety management system would be to conduct a full investigation to understand root cause of the accident. This then allows identification and implementation of improvements to prevent recurrence. Too often, the root cause is left as human error with corrections directed towards individuals involved stopping short of real root causes with proper corrective action taken. Stopping at human error suggests that human error is purely behaviour-based actions that will always occur and have consequence but otherwise cannot be controlled.

4.3 Human Error

Dekker (2006) lists two points of view on human error as the cause of accidents. The first view is that failures regardless of how they manifest are the result of individuals in the system who are the key components to more accidents than anything else. A second opposing view is that human error is not a cause of accidents but a symptom of something deeper; it should never be considered the cause of accidents but the start of an investigation. This concept suggests that errors and accidents are only remotely tied together in the complexity of the systems in which they live. Reason (2008) used the analogy of holes in slices of Swiss cheese to explain accidents suggesting that good management of hazards puts multiple layers between individuals and accidents (Figure 11).

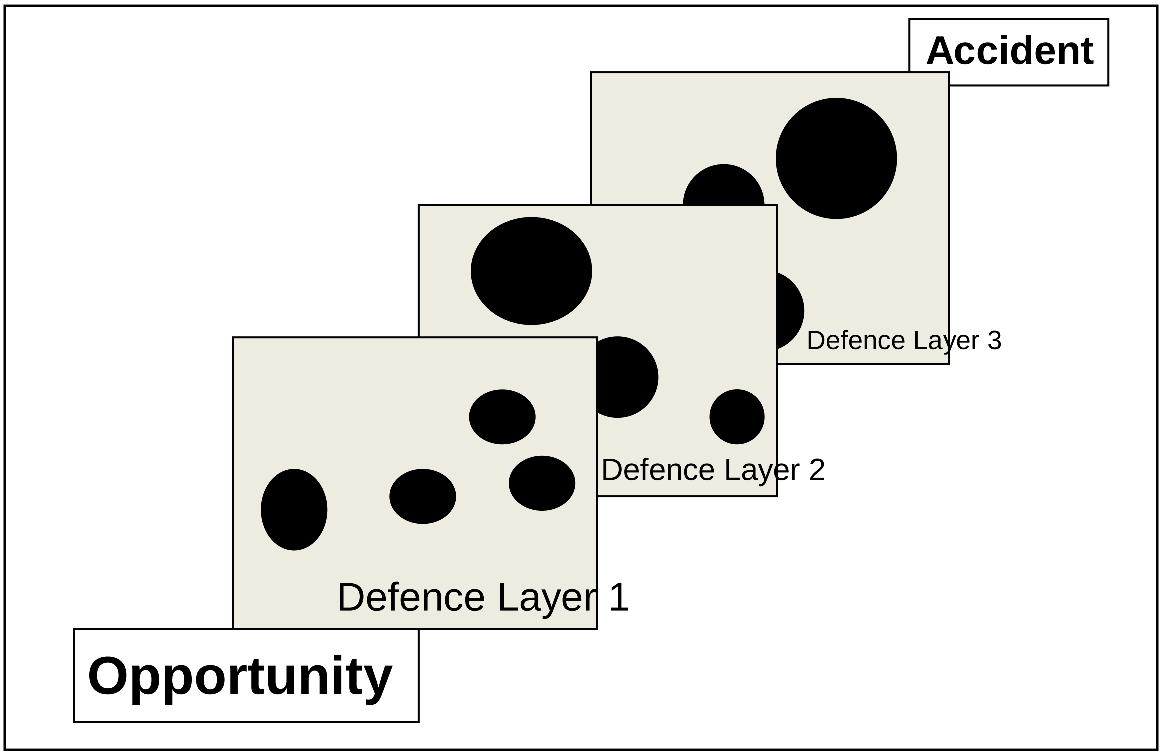

Figure 11 - Reason's Swiss Cheese. (Adapted from Reason 2008, p97)

For an accident to occur Reason suggested several factors need to align simultaneously, like holes in Swiss cheese. Not only must the hazard be present and active as an opportunity for accident but awareness must be missing, as well as protection from it.

The holes in each unique slice represent failure but with multiple layers of protection, chances are if a hole appears in one level, another layer's control behind it will protect. Accidents then are the result of the improbable alignment of the holes in multiple layers creating a line-of-sight from opportunity to accident. The concept suggests it usually takes many jointly-linked factors to align or fail simultaneously to cause accidents. More and different levels of defence will improve reliability of mitigation. In this analogy, each slice of cheese is different with unique hole formations independent from the others. The holes which represent failure, may be fixed or shift, contract and expand. Reason further suggests the holes in the layers of defence occur for a couple reasons, latent defects or an active failure.

This concept was supported in a report addressing the underlying causes of fatal accidents in the UK construction industry (Brace et al., 2009). In this report, the layers of defence are considered to be at a macro level (societal, legislative, union); a mezzo level (project, management, organizational); and a micro level (worker, supervisor, worksite). At each level, the report suggests examples of breakdowns leading to accidents and suggesting controls. These controls can be at multiple levels and consist of multiple defences at multiple levels but all have intent of preventing accidents.

Support material for Rita Donaghy's report to the Secretary of State on an investigation into causes of construction fatalities suggested a modification of Walker's failure categorization template (Walker, 2007). A fourth order component was added which relative to this research aligns with fatigue being an underlying factor of adverse physiological states, explaining personal factors which are preconditions for unsafe acts. These preconditions are the latent conditions waiting for an opportunity to align with. These preconditions appear and disappear as holes in the defensive layers as individuals' conditions change. This tool is helpful in standardizing an investigative approach to identify underlying factors to explain causes of accidents which recognize human factors. Figure 12 illustrates the progressive categorization of factors that can be used to explain accidents. The shaded areas have an association with fatigue.

1st Order	2nd Order	3rd Order	4th Order
(C) Preconditions for Unsafe Acts	(C.1) Environmental Factors	(C.1.1) Physical Operating conditions	(C.1.1.1) Welfare facilities
			(C.1.1.2) Uncontrolled/hidden hazards
			(C.1.1.3) Noise/ Lighting/ Ground conditions
			(C.1.1.4) Space
		(C.1.2) Technological Environment	(C.1.2.1) Availability of suitable resources
			(C.1.2.2) Quality of inspection & maintenance
			(C.1.2.3) Equipment operability
	(C.2) Personal Factors	(C.2.2) Adverse Physiological States	(C.2.2.1) Fatigue
			(C.2.2.2) Health
			(C.2.2.3) Situational awareness
			(C.2.2.4) Intoxication
			(C.2.2.5) Effects of medications
			(C.2.2.6) Stress
		(C.2.3) Training/ Experience/ Ability	(C.2.3.1) Competence
			(C.2.3.2) Reaction times
			(C.2.3.3) Information overload
			(C.2.3.4) Inadequate experience
			(C.2.3.5) Aptitude for task
			(C.2.3.6) Literacy/ numeracy
		(C.2.4) Personal Readiness	(C.2.4.1) Teamworking
			(C.2.4.2) Compliance
			(C.2.4.3) Vulnerable/ migrant workers
			(C.2.4.4) Training/ experience
	(C.3) Task Factors	(C.3.1) Task type	(C.3.1.1) Quality of comms
			(C.3.1.2) Availability of info/ advice
			(C.3.1.3) Minor/ one-off jobs
			(C.3.1.4) Distractions
		(C.3.2) Equipment	(C.3.2.1) Poorly designed PPE
			(C.3.2.2) Hazardous materials
			(C.3.2.3) Tools not designed to'fit' user
		(C.3.3) Task Tempo	(C.3.3.1) Busy phases of project
			(C.3.3.2) Long/ antisocial hours
			(C.3.3.3) Incentives for fast work
			(C.3.3.4) Repetition/ boredom
			(C.3.3.5) Rescheduling of work without planning
(D) Unsafe Acts	(D.1) Accidental	(D.1.1) Skill-based Errors	
		(D.1.2) Decision Errors	
		(D.1.3) Perceptual Errors	
	(D.2) Deliberate	(D.2.1) Routine	
		(D.2.2) Exceptional	

Figure 12 - Underlying Causes of Accidents (adapted from Brace, Gibb et al, 2009)

Additionally, the active failures of Figure 12, ‘D’ Unsafe Acts, are classified as accidently or deliberately deviating from a safe practice to make mistakes. They are why multiple layers of defence are put up; they are not things in a system design that can be eliminated. When considering fatigue as a hazard it may not convert to an accident but may be present due to workload or personal factors. It may be present and latent for long periods or short periods of time; it may be transient. It may be present at certain times of the day coincident with fluctuations in the circadian cycle and may appear and disappear in the same work shift. It otherwise may not impact a site except for productivity.

4.3.1 Fatigue as an Active and Passive Component

Not yet presented is the concept that fatigue can be both a driver and weakened defence for an accident, at least at the site level. It can align with active failures (Category 'D') to cause the unsafe act and it can also weaken defences of another in harm's way. Latent conditions are usually the result of the design of the work or procedures and training governing how the work will be done. As a component of adverse physiological states which define personal factors, it must be noted that fatigue, like other impairing agents such as alcohol and drugs, are not narrow in their impact. They can impact several categories directly and indirectly due to their influence on multiple areas of cognition and at multiple locations in the planning of the defences to hazards. In the design of a construction project and worksite, factors such as staffing levels and overtime can lead to latent conditions of fatigue in workers that may lead to slowed reaction times or poor judgment. One worker's departure from safe work procedures may require another's vigilance and fast reaction time to avoid an accident, but fatigue weakens that defensive layer.

Moore-Ede (2012) believes fatigue is under–reported and under-investigated, meaning that most businesses do not know what percentage of their incidents have fatigue as an underlying cause. Additionally, they would not know whether their safety programmes or fatigue-management programmes are effective or not. Further, as there are no currently tested bio-markers for fatigue, the process is highly subjective and dependent on input and accuracy of the input from individuals relative to those factors that influence worker fatigue level. Moore-Ede sees this as resulting in a complex process of scientific deduction requiring experts to get it right. The problem is evident in not having tools to assist with assessment of fatigue.

Analysis of collected data as advocated by Moore-Ede's FACTS (ibid) programme intends to provide a consistent approach to the collection and interpretation of data that will explain in a probabilistic fashion the likelihood fatigue-impairment was a factor underlying the human error.

Impairment can be thought of as a veil that falls over individuals' cognition and acts in a fashion to elevate risk associated with all factors related to the impaired cognitive functions. If tools can be used to measure normal cognitive functionality they may be able to provide insight into the fall of the veil. This research attempted to gain insight into this factor. Laughlin stated "If there is one thing we know about fatigue, it is that it leads to human error, and the consequences are seen everywhere" (Laughlin, 2003, p1). The challenges do not change however, when considering how one might know if fatigue was a factor or not.

4.4 Preparing the Defence to Fatigue

Accepting Reason's concept means that the solution to human error is found in taking a systems approach. The systems approach accepts human errors will occur as consequences of the environments humans work in and therefore attempts to design layers of controls as defence for the conditions. Defences can be technology-based to monitor conditions and alarm, and even take initial control actions when limits are reached; defences can be human-based relying on procedures and training at the micro level or legislation and compliance at the government and organizational level. Lerman (Lerman et al., 2012) suggested there are five levels of defence required that impact two similar but different aspects of fatigue management; sleep and alertness shown in Table 1.

Workload- staffing balance	Sleep Management
Shift scheduling	
Employee fatigue training and sleep disorder management	
Workplace environment design	Alertness Management
Fatigue monitoring and alertness for duty	

Table 1 - Five levels to defend against fatigue in the workplace (ibid)

If there is a weakness in the model from a practical perspective, it is the reliance to be placed on individuals as partners in managing their fatigue. In Lerman it is stated " ….employees should be encouraged to monitor their own level of fatigue and inform their supervisor if they believe they are too fatigued to safely perform their work" (Lerman et al., 2012, p28). Back-up to this is peer-to-peer safety observations. Regardless of training, large variation will exist in abilities and willingness to do this. Additionally, it is not suggested how organisations would deal with non-conformance to requirements. Also, training cannot overcome an impairment issue that may itself cause one's assessment of their level of fatigue to be off as well as their judgment on proper action to take. Reliance on individual assessments and organizational observation to detect fatigue is especially concerning in schedule-driven environments like construction. Supervisors may not want to hear from their workers that they are too fatigued to continue working safely and are therefore apt to act negatively to that suggestion. Additional resources may not be available and the over-all schedule may suffer putting pressure on all to continue working. There may be individual financial disincentive to report. Construction work can fluctuate with seasons and the economy, so when work is available, there may be individual pressure to maximize hours worked rather than book off due to fatigue. Figure 13 shows Dawson and McCullough's framework for managing fatigue and the control required for each step as well as positioning the opportunity space for this research.

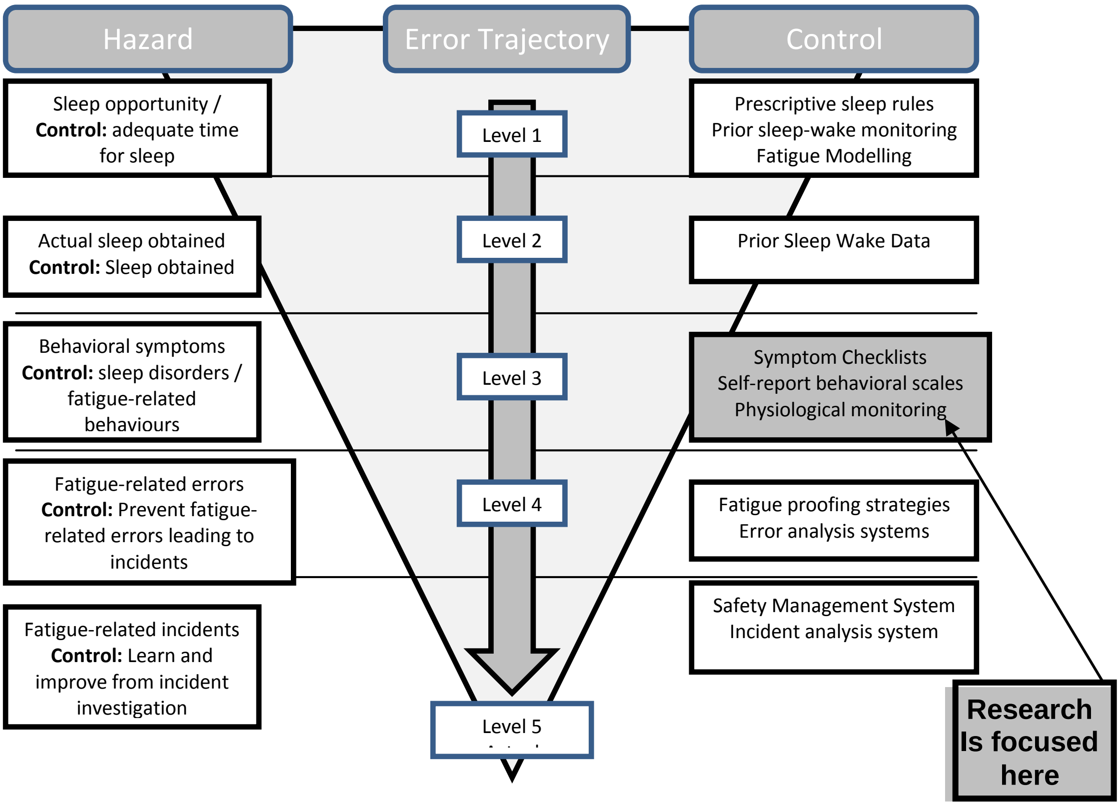

Figure 13 – The Fatigue Error Trajectory (Adapted from McCullough et al. (2008, p9))

The hazards and their suggested controls are listed in Figure 13 for each of 5 levels progressing from sleep obtained through to fatigue-related incidents. The objective is to arrest the hazard at the lowest possible level. On the control side, Dawson and McCullough suggest that to control for level 3 fatigue-related behaviours, self-reporting behavioural scales should be used as well as checklists and monitoring. The opportunity is to develop tools that may provide more objective means of assessing fatigue-impairment.

4.5 Summary

This chapter illustrated the recent evolution of managing fatigue in the workplace. Heavily regulated sectors such as aviation, where fatigue has long been identified as a significant factor affecting worker performance, placed emphasis on Fatigue Management Plans in attempts to ensure adequate rest was available for safety critical roles to prevent fatigue-related incidents. More recently, fatigue has been recognized as a human factor that is itself a workplace hazard that varies depending on the safety sensitivity in the role. The suggested approach for managing workplace hazards is a risk-based systems approach. Safety programmes that focus on hazard identification and risk management by applying such a systems-based approach for addressing the risk of worker fatigue in the workplace allow individual defence layers or controls to be put in place commensurate with the hazard's risk. The weakness still remaining is the lack of tools available to objectively assess fatigue and establish working limits.

CHAPTER 5 – FATIGUE'S FREE PASS

5.0 Introduction

Significant evidence has already been presented on concerns that researchers hold relative to fatigue-related impairment. Whether at work or in general life, fatigue-impairment should be of great concern to everyone. Not only are immediate performance-based risks elevated by impairment but long-term health effects, albeit not fully understood, are also a concern. With existing levels of knowledge and understanding on sleep and fatigue, it is not evident that a commensurate focus is being put on managing fatigue-impairment in the workplace. In many job sites there appears to be disinterest in fatigue impairment or reluctant acknowledgement. This chapter investigates whether this perception is correct or not and explores potential reasons for our views.

5.1 Worker Awareness of Fatigue-Impairment

To gain understanding of the general state of focus on fatigue in the global construction industry, a survey was constructed for release to representatives of this global industry. The results found have been put forth in a peer-reviewed paper and presented at a global conference (Powell, 2010).

5.1.1 Construction of the Survey

The intent of the survey was to explore the level of awareness and focus the construction industry was putting on fatigue in the work place. The results were not intended for statistical inference but rather to confirm and help explain observed performance. Questions were put forward to expose individual and organizational positions and attitudes relative to fatigue in comparison to other impairment factors. The feedback had reach into four continents and to a target group who had good knowledge of operations.

The survey was designed to be completed quickly and comprised twenty-five (25) questions (Appendix 1) grouped into four categories; identification; knowledge of impairment; organizations' activity relative to impairment; and coping with impairment. The survey was only released on-line to facilitate analysis and because all target participants had daily access to computers.

5.1.2 Participants

The target for the survey was workers of the global construction industry. The survey was released 6 times over the course of three years, 2008, 2009 and 2010 to 6 different groups. 5 of the 6 groups were affiliated with the University of Bath, International Construction Management (ICM) programme. One release was to Safety Managers/Directors of a large global construction company. Releasing to the ICM students carried a unique opportunity. The programme brought together students from different parts of the Americas, Europe, and Africa. The reach of the survey was therefore beyond the boundaries of the classroom and allowed opportunity for comparison of view points from a global perspective.

5.1.3 Alcohol Comparison to Fatigue

Fatigue comparisons to alcohol are deliberately made. Alcohol is consumed world-wide and has numerous performance-based studies existing. All respondents could easily relate to alcohol and sleep. Performance issues have been studied and laws have changed to address safety critical activities that can endanger others if not performed fully alert. Driving is the best example. Thanks to Dawson and Reid's pioneering study (Dawson and Reid, 1997) which compared performance changes due to alcohol-impairment and sleep deprivation, the pragmatic leap has been made in quantifying fatigue-impairment. Several research projects in the past decade have refined this area resulting in correlations between lack of sleep and Blood Alcohol Concentration (BAC) levels (Fletcher et al., 2003, Lamond and Dawson, 1999, Falleti et al., 2003, Maruff et al., 2005). Numerous pre-existing studies revealed the effects of alcohol on psychomotor performance and cognition, thus allowing a bridge from fatigue-impairment to numerous studies related to BAC.

Associated with workplace ‘fitness-for-duty’, the contemporary issues pertain to employers continuing to focus on alcohol and drug tests as a means to remove impairment from the workplace. Some workplaces have moved beyond this but it remains an issue in the global construction industry. In part it is due to the fact that the industry is an apparent magnet for drug and alcohol use by its members (Powell, 2010) which sets up opportunities for major work site impairment. Drug and alcohol testing actions are driven by legitimate desires for safety and productivity (Smithson, 2008) but are running into a host of concerns including that there is no evidence that existing tests ever prevented an accident. Worker rights’ advocates have raised strong opposition to the fact that workers may be tested even if they don’t use alcohol or drugs. This may violate individuals’ privacy and civil rights and effectively require workers to continually prove their innocence at the expense of exposing their personal life-style.

Without clear evidence that urine or blood testing produces workplace benefits that justify the costs of the testing programmes it raises serious concerns regarding the effectiveness in improving safety and productivity. Invasive drug and alcohol testing has many flaws and from the perspective of overall impairment, suggests other impairment factors do not matter. This is important when considering fatigue-impairment. There is no biological test for fatigue so in contemplating tests, consideration immediately moves to those performance-based areas for testing that are the result of impairment, as opposed to evidence of the impairing agent. This keeps the focus properly placed on fitness-for-duty.

Many organisations address impairment factors in guidelines and mention them in personnel policies but it is not clear if they are doing more than establishing a boundary for potential liabilities as opposed to really trying to manage them. This survey addressed some of these issues.

5.2 Results

115 individuals completed the survey and represented operations in Europe, North America, South America and Africa. Table 2 summarizes the participants.

Release	Female	Male	Total	Years Worked				
				1 - 5	6 - 10	11 - 15	16 - 20	20 +
First	3	32	35	6	6	9	8	6
Second	1	21	22	3	5	1	4	9
Third	4	17	21	10	6	3	1	1
Fourth	0	14	14	1	6	0	2	5
Fifth	3	11	14	6	0	0	3	5
Sixth	0	9	9	2	1	3	1	2
TOTALS	**11**	**104**	**115**	**28**	**24**	**16**	**19**	**28**

Table 2 - Survey Participant Profile

Collectively, the respondents represented over 1500 work-years of experience. The respondent workers tended to be managers, professionals and owners, reflective of the distributions to mature students typically coming from managerial and professional roles as well as a group of safety managers. For the purposes of this survey, this was an advantage as individuals who had visibility on current practices in their organisation were responding. Table 3 illustrates the results from 3 questions associated with which impairment factor observed in the workplace was the most serious, prevalent and biggest priority for the industry. The results show that overall, fatigue is viewed by 6 separate surveys as the most serious problem in the construction industry, the most prevalent in the industry and the number one priority to deal with in the industry.

Survey Observation	Impairment Factor	Survey Release						Weighted Response
		1st	2nd	3rd	4th`	5th	6th	
Seriousness (1 is most serious; 6 least serious)	Alcohol	#2	#2	#2	#2	**#1**	**#1**	2.86
	Fatigue	**#1**	**#1**	**#1**	**#1**	#4	#3	**2.68**
	Illicit Drugs	#3	#3	#3	#3	#2	#2	3.07
	Illness	#4	#4	#4	#3	#3	#5	3.49
	Medicinal Drugs	#5	#5	#5	#5	#6	#4	4.24
	Other factors	#6	#6	#6	#6	#5	#6	4.67
Prevalence (1 is most prevalent; 6 is least prevalent)	Alcohol	#3	#3	#4	#3	#3	#4	3.67
	Fatigue	**#1**	**#1**	**#1**	**#1**	**#1**	**#1**	**1.98**
	Illicit Drugs	#6	#5	#5	#5	#6	#4	4.14
	Illness	#2	#2	#2	#2	#2	#2	2.91
	Medicinal Drugs	#5	#4	#3	#5	#4	#3	3.99
	Other factors	#4	#6	#6	#4	#5	#6	4.31
Problem Priority (10 is the biggest problem)	Alcohol	#2	#2	#3	#3	#5	#3	5.78
	Fatigue	**#1**	**#1**	**#1**	**#1**	#3	**#1**	6.41
	Illicit Drugs	#4	#3	#4	#5	#6	#2	5.27
	Illness	#3	#4	#2	#3	#3	#4	5.55
	Medicinal Drugs	#6	#5	#5	#6	#2	#5	4.73
	Other factors	#5	#6	#5	#4	**#1**	#6	4.92

Table 3 - Results of Surveys

The importance of each of the impairment factors in construction is evident from Table 4 showing the percent of organisations represented in the survey that had emplaced impairment policies and impairment training, as well as the views held on actions to take for impairment.

Impairment Factor	% w/ Policy	% w/ Training	Zero Tolerance	Only if risk to others
Alcohol	73.9%	47.8%	81.7%	8.7%
Fatigue	15.7%	27.8%	20.0%	42.6%
Illicit Drugs	68.7%	37.4%	85.2%	4.3%

Table 4 - Percent of Companies represented that had Policies and Training for each type of impairment

The group estimated an industry productivity loss at 14% due to fatigue and 79 of the 115 respondents (69%) were aware of an occupational injury in the last 12 months where impairment was a contributing factor.

Table 5 shows 49% of the respondents believed fatigue-based impairment would allow better job performance than alcohol/illicit drug-based performance whilst 44% saw it as the same. 47% believed fatigue-impaired workers would be more productive than alcohol/illicit drug-impaired workers whilst 44% saw it as the same.

Question	Survey						All Average
	1st	2nd	3rd	4th	5th	6th	
Fatigued job performance vs Alcohol/Drugs	Better (61%)	Better (50%)	Better (33%)	Better (42%)	Better (50%)	Better (44%)	49%
	Same (28%)	Same (45%)	Same (67%)	Same (50%)	Same (35%)	Same (56%)	44%
Fatigued productivity vs Alcohol/Drugs	Better (58%)	Better (50%)	Better (48%)	Better (35%)	Better (29%)	Better (44%)	47%
	Same (28%)	Same (45%)	Same (48%)	Same (57%)	Same (64%)	Same (44%)	44%

Table 5 - Comparison of Fatigue-impairment to Alcohol and Drugs on Performance and Productivity

Detection techniques were similarly undeveloped. ‘Observation’ was the technique receiving the highest response for ‘Alcohol use’, ‘Fatigued’ and ‘Illicit drug use’ from all respondents which suggests that even though there are controversial attempts to test for drug and alcohol use, at least amongst this group surveyed, there was no significant testing in their spaces .

Finally, when asked when action should be taken against impaired workers, 82% believed anyone using alcohol should be immediately removed from the workplace with zero tolerance whereas 9% felt they should be removed only if a risk to others. This was 82% and 4% for illicit drug use. However, for fatigue-impairment, only 20% believed they should be immediately removed with zero tolerance whilst 43% believed they should keep working unless a risk to others. A further 30% of the respondents did not view fatigue to be a problem unless the individual could not do their job.

5.3 Discussion

There was general consistency in the responses across the six surveys with some shifts in views on particular questions over the three years the survey was used. The differences shown in some responses in the latter surveys appeared to be a shift towards more economic-based factors perhaps due to the global economic downturn from better conditions pre-2008. Thus stress from job-loss for instance was appearing in comments of the survey associated with ‘other’ factors.

Most importantly, the surveys confirmed that the individuals surveyed and the organisations they represented have different views of impairment based on the cause of impairment and do not view the issue as one of ‘fitness-for-duty’. Taking the pooled responses weighted by the number of participants from each survey provided clear results. Fatigue and fatigue-related impairment was a clear leader in the construction space represented by the survey despite construction having leading incidents of alcohol and illicit drug use compared to other sectors (Powell, 2010). There was surprising agreement regarding fatigue as a source of impairment in the workplace such as viewing it as the most serious factor, the most prevalent factor, the number one impairment problem priority. It was also rated as the number one factor associated with productivity loss.

5.4 Double Standards

Some interesting views of fatigue as an impairment factor emerged. Despite the findings above it was not seen as a factor worthy of testing for, even if simple tests were available. This would have to be clearly addressed in introducing any programme incorporating testing in an operational setting. Another interesting response was associated with impaired worker performance. While worker fatigue was viewed as the most serious problem responsible for the most productivity loss, respondents felt if two workers were impaired to an equal degree, one from fatigue and one from alcohol or illicit drugs, the fatigued worker could still perform better than one using alcohol or drugs. One question which clearly brought out the suggestion that being 'under the influence of alcohol or illicit drugs' was viewed quite differently than other forms of impairment was associated with the actions that should be taken on the impaired. The results of Table 4 illustrate different standards for impairment in the workplace. The 'zero tolerance' views for alcohol and illicit drug use held by over 80% of respondents were in stark contrast to fatigue where only 20% held a 'zero tolerance' view. This is supported at the organizational level by the large differential in policies and training in place for illicit drugs and alcohol versus fatigue-impairment. The low percentage of organisations addressing fatigue versus alcohol and illicit drugs perhaps speaks to a lower concern for it in the construction workplace despite the ratings of Table 3 suggesting it is the most relevant. It also speaks to the state of readiness of the industry to move to where the leading sectors already are with fatigue-risk management programmes. A possible explanation for the individual views may be associated with the lack of fatigue policies and fatigue training in the industry. Organisations focus on what is important to them and a lack of focus conveys a message on importance to workers. There is a clear parallel in the low level of policies and training and the individual views on when to take action.

This leads to the suggestion that fatigue-related impairment gets a 'free pass' in the workplace and raises questions as to why. These results do help illustrate a built-in bias of the industry, which on the one hand believes fatigued workers are the biggest impairment problem in construction but on the other hand not only tolerates them but protects them. From the comments section, there was a prevalent belief that 'strong' individuals can rise above human physiology and disregard the desire and need for sleep to adequately address their fatigue. Other viewpoints put forward are characterized as choices made.

- ‘Fatigue is a fact of life in the 21st century. It’s not something we choose to do but a consequence of not having enough time to do what we have to do’;
- ‘It’s naturally occurring therefore not as bad as someone choosing to impair themselves’;
- ‘Nobody gets enough sleep so you just have to learn how to cope with that fact’;

If there is a clear message emerging from this sample, perhaps it is associated with the inference that how one got impaired is believed to be more important than their level of impairment or whether they are fit-for-duty.

Laughlin (2003, p1) wrote, “Fatigue. Everyone knows it is bad but excessive fatigue is often worn as a badge of honour – as if to show how hard we’re working instead of how dangerous we have become”. If it is true and everyone knows it is bad, we have not evolved in many sectors to embrace how bad it is. Perhaps Laughlin’s comment would have been better worded as “Fatigue. Everyone should know how bad it is and then it wouldn’t be worn as a badge of honour.”

5.5 Summary

This chapter reviewed attitudes from individuals working in different capacities in the global construction sector on four continents. The views confirmed a consistent message that was independent of culture, country borders or size of company. Fatigue is regarded as the most prevalent and most serious impairment factor facing the industry. This view arose in the face of the construction sector having the worst or second worst incidents of alcohol and drug abuse of all industries.

Despite the views towards fatigue, the survey also confirmed little to nothing was being done about fatigue-impairment as opposed to its companions, alcohol and illicit drug use. Perhaps we are closer to understanding why no action is being taken; it is accepted as a fact-of-life and given a free pass.

This has large ramifications for implementation of any solutions as it suggests attitudes will need to change first and this may require some heavy lifting.

Many of the beliefs are associated with a general lack of knowledge and understanding associated with sleep and human physiology. Until changes are made, fatigue will continue to get a ‘free pass’.

CHAPTER 6 - IDENTIFYING AND MEASURING FATIGUE-RELATED IMPAIRMENT IN THE WORKPLACE

6.1 Introduction

Previous chapters have revealed that all science suggests that fatigue-impairment in the workplace can have huge multi-dimensional consequences. Additionally, it is also perceived by construction employees to be a major issue in their workplace. Chapter 4 introduced that advocates of very contemporary Fatigue Risk Management Systems believe employee awareness and observation are critical defences to fatigue-related impairment in the workplace. A persistent question has not been addressed and that is, how does one, let alone an organization know when they are experiencing fatigue-related impairment. Indeed, not knowing may be a significant reason that there is so little done in the workplace associated with fatigue. This chapter reviews attempts to detect fatigue-impairment with different technologies. Possible means to screen workers and identify fatigue-related impairment is critiqued as well as reviewing concerns and factors for consideration in real world operational settings. In particular, actigraph-based effectiveness models and cognitive tests are explored to see if they could be adapted to screen for worker fatigue in the workplace. To assess practical use of tools, other alertness-impacting factors such as caffeine consumption, which is normally found in the workplace, are reviewed for potential confounding effects. Reports on attempts to measure fatigue-impairment in the workplace are restricted to pilot or research projects using some of the technologies of this chapter.

6.2 Employee Input

Detecting fatigue in the workplace is dependent on self-reporting and vigilant observation. It is the symptoms that are observable, not the fatigue. Surveys are tools that can be used to gain insight into fatigue conditions in the workplace. Additionally, there are estimates made based on surveys and extrapolations from medical data and accident reports. There are two surveys in common use which can be used for assessing levels of fatigue-impairment and as part of screening for sleep disorders in the workplace. In particular the Epworth Sleepiness Scale (Johns, 2012) has a high degree of specificity for some sleep disorders which manifest as extremely high daytime sleepiness. Another is the Stanford Sleepiness Scale (Dement, 2012) which compares a self-reported condition of sleepiness at different times of day with those from normal populations to help identify abnormalities. While surveys may provide useful diagnostic insight, they have limitations as a tool for managing fatigue-impairment in the workplace on an on-going real-time basis.

6.3 Impairment Detection Technology

Many contemporary operational practices do not even focus on impairment detection but rather the detection of impairing agents such as alcohol and illicit drugs from which impairment assumptions are derived. Relevant for this research is fatigue-related impairment and a search for practical, economic real-time objective tests for it in the workplace. Even when fatigue-impairment is acknowledged, there is no understanding of its severity or how one might quantify it to deal with it in a cost-effective manner (Moore-Ede, 2012).

6.4 Classifying Tests for Fitness-for-Duty

Miller (1996) classified possible approaches for monitoring worker fitness-for-duty in the workplace in three areas, "performance, neurological, and biochemical measurement". Biochemical tests are suitable for drug and alcohol testing but their use in the workplace has received substantial critique. During the past decade, advocates for alternative means of testing for these sources of impairment in the workplace argued legal and human rights issues. This has helped the development of technologies focused on performance testing as opposed to the detection of individuals' use of particular impairing agents. The suggestion is that the mental and physical condition of the employee as a measure of 'fitness-for-duty' is the only correct measurement as opposed to testing whether certain, perhaps illegal, substances have been used, sometimes in the distant past.

Conventional testing for alcohol and drug use still has application but for general workplace screening for fitness-for-duty new tests for impairment such as computer-assisted performance tests which measure hand-eye coordination and response times are emerging (Belenky, 2012, Gunzelmann et al., 2008). By extension, tests can and have been developed for all facets of cognition that are affected during impairment (Cogstate, 2012). Lerman (Lerman et al., 2012) suggested there are two types of fatigue monitoring technologies; operator task monitoring and task-independent tests. A third approach that does not fit either of the above is associated with modelling or estimating fatigue per Table 6.

Type	**Application**	**Tool**
Monitoring	Physical	Ocular or eye-based monitoring; head position; EEG brain wave reading
	Performance	Monitoring deviation from expected path/response; lane tracker
Cognitive	Test	Reaction times; psychomotor vigilance; memory; learning
Model	Sleep	Sleep-based effectiveness model; actigraphy

Table 6 - Types of Approaches for monitoring fatigue

6.5 Biomarkers for Fatigue

Unlike some other impairment factors in the workplace such as alcohol, illicit and medicinal drug use, there are no convenient biochemical elements that can be tested that correlate to the condition of fatigue. There are no biological tests that workers can take to determine impairment status. Fatigue-impairment is more similar to illness in that the symptoms are the first clues that one is impaired. Only recently, research is surfacing the existence of biomarkers as potential links to some diseases (chronic fatigue syndrome) (Hokama et al., 2009). Others have found promising links to biomarkers for sleepiness but we are still at the fruit fly level (Purdy, 2006).

We are not yet close to having any biological tests to assess fatigue levels for workers.

6.6 Monitoring for Fatigue

While fatigue may not be directly measurable there are detectable symptoms of fatigue. Somnolence or the on-set of sleep is one. When we explore the possibilities for real-time, non-task specific fatigue measurements we quickly find that a surrogate for fatigue is needed as there are no direct biological or physical measurements. As shown in Chapter 3, a key measurement tied to our circadian rhythm is alertness, which can be thought of as how quickly we perceive, process, and act on our processing. Research has shown this is in part related to eye activity which can be monitored to detect fatigue-related changes in performance (Van Orden et al., 2000). However, monitoring eye activity requires roles with fixed head positions and activities allowing monitoring multiple eye measures to correlate to real-time estimates of performance. Research has also been conducted using electroencephalographic (EEG) measurements of brain wave activity tied to performance by EEG power spectrums found to fluctuate with changes in alertness (Makeig and Jung, 1995). These research studies assist in leading to solutions but have no practical outlet presently for most real world workplaces and would clearly not meet needs in a construction environment.

6.6.1 Eye-monitoring based systems

There are two types of application of ocular-based products. One group uses technology to actively monitor the eye in work task to detect somnolence; the other is equipment that delivers an ocular-based alertness test.

Research has been conducted in several aspects of perceptible eye changes during fatigue including pupillary diameter, pupil reaction to light and saccadic velocity (Morad et al., 2009). Techniques for detecting somnolence utilize camera-based systems to monitor eye and eye-lid movement and feed companion computers and software to make declarations about the state of drowsiness (Everding, 1999). Drowsy states set off alarms for the driver. Applications tend to be focused on activities which require on-going fixed gaze and alertness in performing duties such as driving where it appears to be well suited. Cameras can be mounted with unobstructed view of the eyes to feed images to small in-vehicle computers. Images are constantly interpreted as normal or abnormal states of drowsiness based on reflection of detected light frequencies typically in the infrared spectrum.

Systems are now being deployed in higher end motor vehicles and some other heavy equipment. Using Reason's (2008) model, such systems would act as a final line of defence to prevent the final step in a succession of causal links that have all failed leading to somnolence and an accident. However, they can only be viewed in this regard and not as a solution to address any of the other symptoms of fatigue. Knowing that a sedentary activity such as driving can promote somnolence, this monitoring is all the more important. However, in the state of decreased alertness, other important factors such as safe judgment and reaction times are also impacted; neither is detectable. Any worker roles that allow a fixed monitoring of the eyes and where the operators' uniform gaze is required would be candidates for application of this technology. This has already been implemented in heavy equipment and will most probably continue its penetration in this regard into construction work sites. These task-specific monitoring tools are otherwise not suitable candidates for general assessment of fatigue-related impairment or for any activities which do not fit the base requirements of fixed gaze.

Intended for general workplace use, other ocular systems have been developed (Eyedynamics Inc., 2012, PMI, 2012). Techniques used vary but capitalize on a concept of a responsive ocular system being coupled to an unimpaired mind. One example is Safetyscope© from Eyedynamics which gives readings of impairment by monitoring how well the eyes smoothly track an object moving horizontally ("horizontal gaze nystagmus") and monitoring the pupils' response to changes in lighting. PMI markets the FIT© system, which is also based on the eyes' involuntary responses to light stimuli. The eye is monitored and if responses fall outside of defined ranges for normalcy a reading of impairment is given. Stimulants which are commonly taken in the workplace are not accounted for in the performance of these systems or explained with the results. The brevity of the tests, reported as being as fast as 30 seconds, is an advantage. Operating motorized equipment in construction would be candidate roles for this technology.

6.6.2 Performance-based systems

Performance-based systems require tasks that have outputs that can be compared to a norm or expected output. Driving is another example where systems are emerging to monitor the drivers' performance and alarm when there is too much deviation from a norm. Camera-based systems again have been used within the truck to monitor drift from a defined line or track. Degree of drift is associated with fatigue and thus provides feedback to the drivers on their performance if it slips. Any other activities which have defined levels of performance which can be modelled are candidates for using this technology. It appears to be most suitable for activities requiring sustained and defined performance which is completely different from construction. While there are a lot of safety critical functions in construction such as operation of equipment, the feedback system requires a programmed and planned operation to monitor operators. Highway heavy-haul trucking provides this opportunity. It is not clear where else this would have application in construction. As a final safety-barrier for the driving roles, like camera-based systems, it holds some promise in part due to the non-selective application of the monitoring. Whether one has lost focus due to a lack of sleep or due to other impairment factors such as drugs or alcohol, it will provide a degree of final protection. However, after trial and review in a large pilot study, it was not identified (Dinges et al., 2005) as a successful candidate for acceptance by the operators.

6.6.3 Cognitive-based systems

Associated with changes in brain wave activity and alertness are the mental processes or cognition. This research is intimately tied with 'cognitive' processes or that which relates to cognition in a psychological sense. The cognitive processes are numerous and include our ability to react to stimuli, make decisions, learn, recall from memory, focus and be attentive. If impairment leads to decrements in cognition, our ability to test psychological functions then becomes interesting from a perspective of assessing impairment. In particular, memory is an important component of cognition as it supports higher levels of cognitive functions including rational thinking and decision making. It has been shown to deteriorate with poor sleep (Turner et al., 2007).

Henry Bowles and Theodore Langley were perhaps the first to develop a dedicated alertness checking device made available in 1998 (Bowles-Langley, 2012). The assessment was based on a computer-based test of mental alertness which stored results as baselines for expansion into other areas of applied cognitive science. Other systems have emerged including CogState Research (Cogstate, 2012) which was used extensively in this research. The underlying concept is similar. Stimuli are presented for various aspects of cognition and subjects are tested relative to a baseline which may be either database norms or relative to an individual's past performance. CogState is oriented to research applications and as such has significant numbers of peer-reviewed papers on its use for a variety of cognitive applications to facilitate new studies.

6.6.4 Sleep-based effectiveness modelling

Coupled with understanding sleep and its impact on human physiology, numerous attempts have been made to model sleep's effect on cognition (Van Dongen, 2004). Modelling the sleep process and accounting for all aspects that affect alertness via sleep is probably a natural outcome of researching and understanding sleep. Personal computers greatly facilitated the practical application of effectiveness modelling based on sleep. There are still unknowns however that limits the application of these models. An important limitation is the amount of sleep taken versus the amount of sleep needed. One might think it is easy to determine how much sleep we get each night but we only generally know how much sleep was actually obtained on any given sleep session. This is associated with none of us fully understanding the quality of our sleep and only at best, being able to estimate our time spent sleeping. Understanding the amount of sleep needed is also problematic. There are generally accepted ranges that define the basal need for healthy populations but there is a moving daily target for our sleep based on accumulated variance from need due to a host of individual reasons and circumstances.

Actigraphs are small wrist-worn devices that measure the motion of the wearer and provide input to companion software to assess the state of the individual; ultimately via filters and software providing an interpretation of the movement as a sleep/wake state. These devices are now commonly used in conjunction with computer-based models or increasingly with firmware embedded on chips within the actigraph unit. The intent is to display information associated with an interpreted alertness state based on the sleep/wake cycle now measured, not estimated. These devices have proven to be capable of interpreting sleep accurately (Morgenthaler et al., 2007) thus providing extremely sensitive non-invasive measurements of actual sleep. While they are capable of tracking alertness based on the sleep/wake cycle a disadvantage is they must be worn and worn continuously. This is impractical in normal workplace settings and due to the need to monitor sleep and personal activities this quickly can become problematic as a privacy issue for broad continuous deployment. Removal of the units during sleep breaks the valid modelling they provide. Finally, they cannot detect effects of stimulant use or the actual state-of-mind as they model effectiveness solely based on sleep.

There are other sleep-based models in use such as Circadian's Fatigue Accident Causation Tracking System (FACTS) (Moore-Ede, 2012). Structured inputs to the model are made by individuals interested in assessing whether fatigue might be a factor in incidents. The intended use is as part of an incident investigation as opposed to prevention.

6.7 Workplace results

There have been some attempts to study the different technologies associated with measuring impairment including the Workrights Institute (Workrights, 2009) and Transport Canada (Rhodes and Vincent, 2000).

Dinges (Dinges et al., 2005) reported on a large study conducted on behalf of the US Department of Transportation and Transport Canada that looked at the potential use of another group of 'fatigue management technologies' (ibid) to assist truck drivers. Eyelid movement (PERCLOS) and a wrist-worn actigraph with associated algorithms to provide alertness information were reviewed with other performance-based monitors. The Dinges study focused on comparing and understanding the suitability of these technologies specifically for the trucking industry to determine if they could have positive effect on fatigue-based driving. The other two performance-based technologies reviewed were specifically for the trucking industry. As previously illustrated, the heavy haul trucking industry is a very problematic space for operators falling asleep while driving which has led to this focus to develop and apply new technologies, which operate in real-time to assist operators prevent accidents. Preventive application of tools is most desirable as opposed to simply being able to assign probabilities that fatigue was a factor post incident. Whilst models are in use which could be helpful as part of a longer range plan and fatigue risk management system to identify and target problem sites, they provide no immediate defence on a construction site. As a broader issue, actual reports and measurement results for general 'fitness-for-duty' and measuring fatigue-impairment levels in the workplace are scarce.

As disclosed in Chapter 5, the common tool construction organisations use is monitoring and interaction with co-workers to observe any forms of impairment at sites. This is not really different from other sectors, with the exception noted above. There are numerous issues with this. Supervisors may only actually interact with a worker a few times a shift and with no, little or poor training may not even be able to detect any concerning symptoms of impairment. Worse for fatigue, even if symptoms are detected, supervisors may choose to disregard them and act as the survey found and take no action believing they are only potentially impacting productivity and that productivity and schedules may suffer a larger blow if the employee is removed from duty. Thus fatigue gets a free pass. Fundamentally however, workers simply do not know their state of impairment from fatigue and observation is highly unreliable. Even if the obligation is put on the workers to monitor and report their own condition, just as with other impairing conditions, workers' subjective reports do not show correlation with objective measurements associated with their levels of alertness (Lerman et al., 2012).

The essence of this research is that a missing piece in the defensive strategy is an objective real-time measurement for fatigue-related impairment. Lerman et al. (ibid) acknowledge the emergence of tests designed to assess fatigue and alertness independent of task performed, but state "It should be noted that although real-time objective assessment of fatigue is very promising, it is not widely used in the field. It is likely that field trials would be needed to validate their use in an organisation's circumstances" (ibid, p249). This research, specifically addresses this need.

6.8 Summary

There are now new tools and technologies which have strong potential application for certain workplace activities; however, there are clear limitations to using most. All systems monitoring workers outside of the work environment will be affected by personal intrusion issues. Eye monitoring or sleep monitoring via actigraph or using performance tracking devices have limitations for general application. Ocular-based tests have possibility but are limited by the access to the system and cost. An ability to test the real time state of cognition in individuals as a surrogate to their alertness state and ultimately their expected performance would be advantageous, especially if the tests were web-based or otherwise easily portable. However, they lack practical real workplace applications and results. This will be expanded on in Chapter 7.

CHAPTER 7 – USING COGNITIVE TESTS TO SCREEN FOR FATIGUE-RELATED IMPAIRMENT IN THE WORKPLACE

7.1 Introduction

Chapter 6 introduced the emergence of cognitive tests as one of the possible techniques for identifying fatigue-related impairment. Results from these tests have been shown to produce results which correlate with changes in cognitive functions and support the use of cognitive tests for general screening. There are numerous appealing aspects of these tests for screening for fatigue-related impairment including their simplicity, cost, and speed of administration. Unfortunately, there are no known results from research in real work environments to assess their suitability, especially in the face of real world confounding agents. This chapter reports on the research into selection of cognitive tests as a means of screening for fatigue-impairment in the work environment. The types of tests available are reviewed and results from this research's attempts to find a possible technique for the workplace are reported. Consideration is given to confounding factors found in the workplace such as circadian influence and caffeine use. Linear mixed effects modelling was employed to understand and rank the results. This research verified cognitive tests could be used as screens for fatigue-impairment and those focused on reaction-time components performed best.

7.2 Cognitive Tests

Cognitive tests are designed to test aspects of an individual's cognition. For a complete assessment a battery of tests would be required. For specific areas of concern, tests may be used individually or in combinations. Common tests include attention tests; judgment tests, reaction time tests and memory tests. Many of the tests are to assist diagnosis of disease. Memory tests can be designed to test short term, working and long term memories and have assisted in diagnosing Alzheimer's disease (Darby et al., 2011,Ellis et al., 2009). Applications for medical clinical use have been studied and abnormal conditions associated with attention may also lead to diagnosis of attention deficit disorder (Collie et al., 2007).

By the late 1970s Professor Wesnes of the University of Reading had introduced computerized cognitive testing for repeated use in clinical trials (BioSource, 2012). Task stimuli were presented to participants with anticipation of “YES” or “NO” responses. Accuracy of responses and time to react to the stimuli were recorded. Computer-based cognitive testing has significant appeal for screening for fatigue as tests easily run on all computers with exception of smart-phones and tablets. However, it is easy to imagine they will evolve as an application for these devices. They can be stand-alone or web-based. They are simple to administer, sensitive to impairment of cognition and there are enough tests to be very specific with the cognitive functions screened.

This concept is identical to that used in this research as developed by CogState Research (CogState Ltd, Melbourne, Australia) (Cogstate, 2012). When approached about this research project they were willing to fully support it by providing complete access to their battery of tests with technical support for the project without fee. CogState introduced their first test in 1999 to help diagnose sports-related concussions. This evolved to a full battery of 14 tasks for web-based applications of testing all aspects of cognition.

The computerized cognitive tasks were chosen with the support of CogState based on prior research. They consisted of cognitive tasks known to be influenced by impairment and in particular, fatigue. The tasks use a game-like format to assess the cognitive domains of psychomotor processing, visual attention, learning and working memory. These domains have previously been shown sensitive to the effects of mild head injury and concussion, (Collie et al., 2006, Collie et al., 2003a, Makdissi et al., 2010, Maruff et al., 2009, Moriarity et al., 2004, Straume-Naesheim et al., 2009, Makdissi et al., 2001) as well as fatigue and drug use, (Falleti et al., 2003a, Collie et al., 2007, Snyder et al., 2005a, Snyder et al., 2005b) psychiatric

(Pietrzak et al., 2009a, Pietrzak et al., 2009b) and neurodegenerative disease (Darby et al., 2002, Ellis et al., 2009, Maruff et al., 2004, Darby et al., 2011). Similar to other computerized approaches, each task requires a “YES” or “NO” response to stimuli, which in this case is upon display of a central playing card. Each time the test is taken, the visuomotor requirements remain identical, but equivalent alternate forms of each task are randomly generated. Reliability, stability, practice effects, validity and correlations with conventional neuropsychological tests have been reported previously, (Falleti et al., 2006, Fredrickson et al., 2010, Maruff et al., 2009) as have other psychometric properties (Collie et al., 2003b) which drives the interest in researching whether these simple, brief tests could be applied with success in a real world work environment.

7.3 The Opportunity for Cognitive Tests

As reviewed in Chapter 6, progress has been made recently on introducing systems into work environments to identify or monitor for fatigue-impairment or its symptoms but they have limitations. The most effective are narrow in their application, such as the camera-based eye monitoring while other performance-based systems are too specific in their applications. Sleep-based actigraph systems and other sleep-based modelling do not account for confounding elements in the workplace and require considerable administrative time to maintain contemporary data necessary for validity of reported results. Additionally, actigraphs require continuous usage to get data to feed their models and they raise significant issues with on-going use associated with personal privacy. Purpose-built ocular systems have cost/distribution issues and have not developed full understanding on workplace confounders.

In 2003 Falleti (Falleti et al., 2003) used a cognitive test developed by CogState Research (Cogstate, 2012) to validate findings of Dawson and Reid (Dawson and Reid, 1997) who first introduced findings associated with fatigue-related impairment and Blood Alcohol Concentrations (BAC). The research confirmed a correlation between performance and cognitive declines with both fatigue and alcohol impairment but also validated the use of cognitive tests to detect the performance decrements due to fatigue.

Others (Belenky, 2012, Balkin et al., 2002) were concurrently focused on developing tools to assist managing sleep to sustain performance in military operating environments, including using actigraph-based predicted performance and using a new test, the psychomotor vigilance task (PVT) which was ported to a hand-held device. The PVT presents to a subject a lighted dot in different locations on a screen. This is repeated for 5 to 10 minutes and measures responses from the subject. The subject presses a button when they see a light. Reaction times and more importantly, misses are determined, to give a numerical measurement on sleepiness. The PVT has evolved to become even more interactive in spotting and touching the lit area. It has however, been found to be susceptible to individual motivation of subjects affecting results (Loh et al., 2004) and limiting repeated use, especially in operational environments. For commercial operations, conducting regular tests up to 10 minutes in length are difficult to implement and maintain. Even if highly effective, only the most critical of roles would justify this time commitment burden. Additionally, the length of test will affect their acceptance by organizations and individuals and if easily defeated based on poor participant motivation, they would be less valuable in work operations than in labs to diagnose problems.

In Falleti's research (Falleti et al, 2003), there was no consideration given to the possible influence of the circadian cycle on alertness during the tests at different points of the same day. Additionally, other factors shown to be relevant in alertness research, the length of time awake before test and sleep activity in days prior to the test were not mentioned. While the amount of sleep the night before the test was estimated, accumulated sleep debt from previous nights was not. Sleep debt has been suggested to have an influence on alertness even after one good night of sleep (Dinges et al., 1997). The research also controlled stimulants including caffeine. In work environments caffeine consumption is very common and uncontrolled. However, Falleti's results were strong enough to suggest that further testing was warranted in an operational setting. Processing speed, attention and working memory results from a brief computerized cognitive test are the surrogate markers of interest in a work environment. Testing for changes in alertness and looking for significant changes will define whether they could be used in the presence of confounders. This defined the objective of this phase of the research.

7.4 Design of Experiment

To support the objective, the goal of this phase of the research was to design an experiment with simple cognitive tests, to measure aspects of cognition affected by fatigue, followed by analysis to determine how sensitive the tests are in the workplace environment. To determine whether cognitive tests would be capable of providing a strong enough signal on changes in cognition to allow a declaration on a worker’s fitness-for-duty, the focus for testing was the circadian cycle. With an understanding that the circadian has two naturally occurring peaks and two naturally occurring lows (Figure 8), the design of this experiment selected the highs and lows as test points. Ability to measure differences in alertness at these points with a cognitive test would suggest they might be candidates for workplace screens for fatigue-impairment. Other benefits of testing the circadian oscillation included the possibility of not just obtaining a reading on the sensitivity of the screens but also a reading on the magnitude of fluctuations in alertness over the course of a work shift.

Initially, as part of the experiment’s design, it was thought that it would be beneficial to know where individuals’ circadian peaks actually were. To confirm individuals’ circadian cycle aligned with expectation, core body temperature which is also governed by the circadian cycle would be measured (Wirz-Justice, 2007). Once confirmed that core body temperature could be measured to expose a participant’s circadian, it was planned to do so as part of the testing with the cognitive tasks. Six infrared ear thermometers were purchased from Almedic and given to six volunteers.

The request was simple; take a measurement every 10 – 15 minutes and record the temperature.

The work place for this research was an operations and maintenance unit of a large global engineering and construction company based in Vancouver, Canada. The workers were responsible for operating and maintaining an automated commuter rail service as the private arm of a public-private partnership. The operation started in 2009 and operates around the clock, 365 days a year. Additionally, several workers in roles involved in the movement of rail-borne equipment are classified as safety critical roles under Canada's Railway Safety Act, creating additional corporate interest and support in fatigue management. Unlike the controlled environments of previous research, this research did not screen anyone from the study but rather took individuals as they came and took what came with them. This necessarily meant no control of stimulants; no control of sleep; no control of sleep patterns, smoking, exercising, or anything else which may impact one's cognition. Therefore, the study would assess the practicality of using the computerized cognitive tests in a work place environment including assessing impact of potential confounders reported in other research (Van Dongen, 2004).

7.4.1 Selecting Cognitive Tests

Lim and Dinges (2010) summarized research associated with attempts to understand the effects of sleep deprivation on cognitive performance. Their summary reported on a large number of research papers (70 papers) and illustrated disagreement on interpretation of data. Their finding revealed that the variation in sample sizes in each project produced limitations on consistent outcomes. A meta-analysis was conducted to add clarity on short-term sleep deprivation effects of speed and accuracy in six different aspects of cognition affected by fatigue as identified in the others' research. Results from the previous research were classified in simple and complex attention, short-term memory, working memory, processing speed and reasoning. These were the targeted areas for the selection of cognitive tests for this research. Tests were looked to for applicable measures of speed and accuracy. Speed and accuracy relate to processing reaction times on test responses and percentage of correct responses for a given number of responses.

CogState Research (CogState Ltd, Melbourne, Australia) provided access to their full battery of computerized cognitive tests as well as their data base for data storage and retrieval. Additionally, access was given to comparative data for norms on populations tested. Six tests were suggested for this study which addressed areas of learning, working memory, visual attention, and psychomotor processing speed. The latter measure was present in all tests. These were the domains that previous research has shown to be sensitive to fatigue (Falleti et al., 2003b,Collie et al., 2007). They have also been shown sensitive to concussion (Makdissi et al., 2010) and drug use. Differences in responses from different sources of impairment may ultimately provide a response signature and be useful information to develop a tool which can determine source of impairment as well as magnitude of impairment. This was not however, under investigation with this research.

Five of the tasks presented participants with a series of cards face-down which were then turned over with an expectation of a “YES” or “NO” response to the posed question. A response was given by hitting the “K” or “D” key on the computer keyboard respectively. Each task has a different question to respond to and when the task was repeated, the same questions were asked but each task varied slightly in form due to random generation of the cards and timing. The sixth task was a hidden maze requiring participants to find their way from start to finish by clicking on tiles to reveal the path. The task repeated to test learning and how much was retained from each trial.

The tests selected were called the Groton Maze Learning (GML), Detection (DET), Identification (IDN), One Card Learning (OCL), One Back (ONB), and Two Back (TWB). These tests made up a battery which was completed at one sitting. There was no criteria for order other than the GML had to come first and last to maximise time between the first and the last rounds to provide the best test of memory. They were given in the following order:

1. The Groton Maze Learning (GML) task tested working memory and learning via speed-of-response and error rates. A 10 X 10 grid of tiles with a hidden 28-step pathway was presented. Through trial and error, the hidden pathway was found from the upper left corner to lower right corner. Once completed the test repeated five more times with the path again hidden to participants. At the end of the complete battery the GML task was done one more time to measure the same variables as a measure of learning retention. This task took around 40 seconds per round dropping in time each round to about 20 seconds for the last test.

2. Detection (DET) Task is a simple reaction time task used to test reaction times via speed, accuracy, hits and misses. The test presented a playing card and required a "YES" response when it turned over. Any anticipation, counted as a response time less than 100 ms after a card turned over was not accepted as a valid response. The task stopped after 35 correct responses and lasted around 100 seconds.

3. Identification (IDN) is a simple test that tested choice reaction times with speed, accuracy, hits and misses when assessing whether a card was red. All cards were either red or black and presented face down. When it turned over, a "YES" or "NO" response was required. Participants had to decide if the card was red before responding. This task continued for around 100 seconds.

4. One Card Learning (OCL) tested memory and visual learning by tracking speed, accuracy hits, misses, and anticipations. A playing card was presented in the centre of the screen and turned over to reveal its identity. Participants had to decide whether they had seen this card previously in the session or not and then respond "YES" or "NO". There were 42 presentations, with 6 repeating cards of different suits and

denominations mixed with additional randomly selected cards repeating as well. This task took about 3-4 minutes.

5. The One Back (ONB) task was a test of working memory that tracked speed of response, accuracy, hits, misses, and anticipations. A card was presented in the centre of the screen and turned over to reveal its identify. Participants had to decide if the new card was the same as the previous card seen. A “YES” or “NO” response was required. This task stopped after 30 correct responses. This task duration was about 90 seconds.

6. The Two Back (TWB) task was also a test of working memory similar to ONB. It tracked speed of response, accuracy, hits, misses, and anticipations after viewing a card which had turned over on the screen. Participants had to decide whether the card was the same as the card which was presented two cards previously. It required a “YES” or “NO” response. The task stopped after 30 correct responses. The total time to complete this task was around 110 seconds.

7. After completing TWB, participants did another round of the Groton Maze Learning to test working memory and learning.

All tests were run on personal computers which participants had access to with their work duties. All tests were conducted on sessions hosted by CogState and all results as well as other comments were encrypted and uploaded directly to their database for extraction and analysis.

7.4.2 Confounders

As this research was field-based in a real work environment, it was anticipated that there would be numerous sources of uncontrolled factors that might confound alertness of individuals including sleep efficiency, season, day of the week, shift schedule, medicinal and illicit drug use, alcohol or stimulant use such as caffeine, illness and stress. Further, these factors may affect individuals differently and may not be predicted. All of these are sources of variability that can lead to fluctuations in cognitive performance and hence may lead to an over- or under-estimation of cognitive decrements if only one factor is considered.

Caffeine in particular was a concern for this research as it is often consumed in the work place as part of common beverages such as coffee, teas, colas, and energy drinks. To be effective, work place solutions for measuring fatigue-related impairment must account for such sources of systematic and random error and be strong enough to see the signal through all sources of noise. Van Dongen (2004) listed some of these as factors in work places affecting fatigue and performance and made further note that they could be expected to be 'noisy' measurements in the field . Lim and Dinges (2010) cited hours awake as a common element explaining variation between groups.

7.4.2.1 Caffeine

Participants were asked to record all caffeinated beverages consumed. Consumption of caffeinated beverages was known to be common in this work place. Caffeine affects alertness by blocking receptors in the brain stem that detect adenosine; one of the key sleep markers (Brain et al., 2012). How caffeine affects us and its influence in general is known. Caffeine is also known to stay in the body for long periods and has an estimated half life of at least 5 hours. How it actually affects individuals is specific to the individual and the big challenge for any research is getting an accurate reading on how much is in the system. Content in caffeinated drinks (Energy Fiend, 2012) and some foods such as chocolate has been estimated (Amano, 2011). It is also known that the amount of caffeine varies significantly within the same drinks and drinks from the same source. With these caveats which would fall into the category of noisy field measurements, it was planned to record caffeinated drinks consumed and estimate the amount of caffeine in the participants' system at time of testing.

Caffeinated drinks such as coffee and tea were recorded in quantities of 250 ml cups. Colas and other canned caffeinated drinks were recorded at actual size. The amount of caffeine per serving was estimated from published tables (Energy Fiend, 2012).

A simple equation was developed to estimate the amount of caffeine in participants at the time of testing.

$$C = C_0 - C_0 * (t - 1)/10 \quad \textbf{Equation 1}$$

- **C** - the estimate of residual caffeine affecting alertness in mg;
- $\mathbf{C_0}$ - the estimated amount of caffeine consumed with each particular drink in mg;
- **t** - the number of hours the drink was consumed before the test.

Total caffeine in a participant was determined in this way based on the input on quantity, size, timing and type of drinks consumed. Caffeine consumed more than 10 hours before the test was not considered to have influenced test results in this research and so was considered to be at zero.

7.4.2.2 Other factors

It was planned to track other factors found in other research to be significant in predicting performance (Lim and Dinges, 2010). Time awake before performing the task was important. Another was hours slept the night before the test. This information would be provided by the participants as comments to the researcher. Additionally, participants were instructed to provide any other comments post test to the researcher which they felt might be relevant in interpreting a result.

Alcohol and illicit drug use are strictly prohibited in the work places associated with this research.

7.5 Testing

An email was distributed to a target group of personnel associated with participating in this phase of the research (Appendix 2). The group was identified as good candidates because they were on fixed rotating schedules and at a computer for the full duration of their 12-hour shifts. With rotating shifts the target individuals would be in a position to do tests at all identified peaks and valleys of the circadian cycle (the nodes). Testing twice per shift at different nodes would produce the data for analysis. The target group was augmented with requests to a larger general population that included workers whose shift did not rotate.

Participants were asked to volunteer out of pure interest and no payment of any kind was offered for participation. Participants were asked to conduct two tests for comparison per shift at nodes of the circadian cycle which would have to be managed in conjunction with their regular work activities. The total time for the battery was 15 – 20 minutes so the request amounted to a 30 – 40 minute time request on days the tests were done, which was supported by the employer. For this research, it was assumed that there were no major shifts in the circadian rhythm of any workers and that without further information it would align with highs and lows reported from other research (Eddy and Hursh, 2001). This seemed to be a reasonable assumption as none of the evening 12-hour shifts lasted for more than 2 days and were followed by days off then day shifts. Supporting this, Waterhouse (1999) has reported that it takes around one day per hour of shift in the circadian and it further takes a persistent change of at least two hours to cause the shift. This means that the assumption was that individuals on the rotating shifts would experience abnormal and poor work/rest conditions for a couple of days but then revert with a normal day-shift schedule for 5 days.

Selected participants were given access to the tests and provided with instructions on use. Of primary importance was to conduct two tests per day at predefined times and to work through the tests promptly and with as much accuracy as they could. The software allowed selection of a practice mode so participants could practice as much as they wished before conducting an actual test. Additionally, the software had a ‘trial run’ prior to each test just to prepare individuals for what was coming. For infrequent tests this is probably a good feature but for frequent repeat users this is more of an annoyance. The trial run could be skipped by hitting ‘Ctrl’ and ‘Tab’ on the key board, which is what individuals were doing after a couple of rounds due to familiarity. The challenge associated with this phase of the research was finding enough participants willing to do this test and remaining motivated enough to continue doing them.

The times participants were given to test were four preselected nodes of the circadian cycle:

1. Node 1 was 03:00 to 06:00
2. Node 2 was 08:00 to 11:00
3. Node 3 was 13:00 to 16:00
4. Node 4 was 18:00 to 21:00

The times were broad enough to allow flexibility for the individuals to test in a node around their work activities but it was requested that they fix their particular test times within the node to within ± 30 minutes.

Once a test session was completed data was transferred via the internet to CogState’s database in Australia, where the results became visible. The results were frequently monitored to check progress and to provide assistance if needed. Upon completion of the battery, participants were taken to a screen which provided them with an opportunity to submit ‘notes to the researcher’. This is where they reported the sleep obtained in their last sleep session, their wake time, as well as the details regarding any caffeinated drinks consumed. Additionally, any other information they felt was relevant could be typed in and submitted with the data transfer. All tests were automatically time stamped which provided needed information to calculate hours awake and validate node information. While all participants were taken as they came, some confounders such as alcohol and illicit drug use were not evaluated or identified. Influence from their use is prohibited in the work place participants came from but their use was not reported. It is not suspected of being a factor based on on-going observation of the participants.

7.6 Data Analysis

As shown in section 7.4.1, there were specific measurements of interest with each test. Of primary interest was the speed of processing and accuracy measurements. The processing speed translated into a reaction time in milliseconds and accuracy was a ratio of correct responses versus the total number of responses. The data also underwent integrity checks built into the CogState system to ensure reasonable efforts were made and that the data fit expected norms. These have been described by Fredrickson (Fredrickson et al., 2010) and include invalidating data as incomplete if less than 75 percent of the expected correct trials were attempted. Additional built-in integrity checks included that DET accuracy be 90% or greater; IDN accuracy be 80% or greater; ONB accuracy be 80% or greater; and OCL accuracy be 53% or greater. Using the CogState normative population for review also eliminated results if a measured speed fell below 2 standard deviations of the mean performance of age-based controls. Any results that did not pass integrity checks or which were labelled practice were not taken as data for analysis. The analysis of the data was to evaluate if statistically significant shifts in measured results existed between any two nodes tested at.

7.7 Results

Prior to starting the cognitive tests, a small group of employees agreed to try and map out their core body temperatures with the infrared ear thermometers. The simple request to take a measurement every 10 – 15 minutes and record the temperature was a full failure. Not a single worker was able to sustain the temperature mapping with enough frequency to generate any usable information to graph. This was a clear limitation of conducting research of this nature in conjunction with regular work activities.

It was not pursued further for the cognitive testing phase and all testing was conducted according to that outlined in section 7.5.

In total 24 employees volunteered to participate in the cognitive testing and took initial training with the cognitive tests, however, 11 were not able to provide any usable data. Some of this was simply due to failing to do 2 tests per day because of work constraints. Some was due to a loss of interest to participate. Of the 13 who did provide usable data, 8 were males and 5 were females aged between 29 and 56 years. The mean age of the group was 41.5 ± 9.5 years. The target group only generated 9 shift workers. The other 4 were office workers with fixed schedule.

Totally, over 150 work hours were consumed completing these tests for this study but the contribution was not evenly distributed. 364 valid tests were completed (meaning the full battery) but it ranged from 2 to 100 per participant as illustrated in Figure 14. It shows the different number of measurements at each of the different nodes for the DET task.

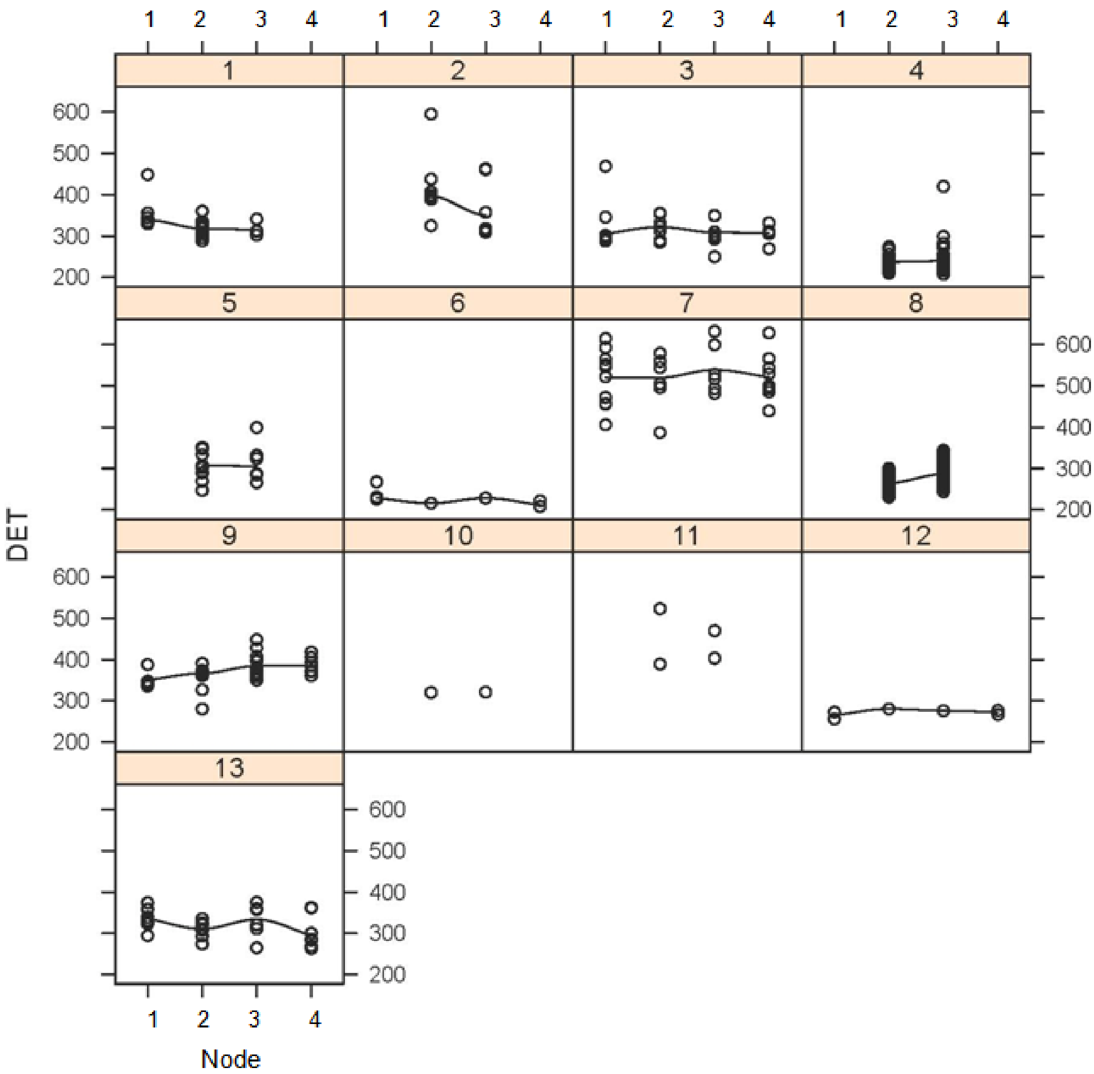

Figure 14 – DET results by node for each participant

The average number completed was 26 (Q1 8; median 20; Q3 30). Full results are presented in Appendix 3. The data was sorted by individual and node with other factors such as sleep and caffeine consumed. Paired data was plotted to visualize the results and look for trends over the testing period per Figure 15.

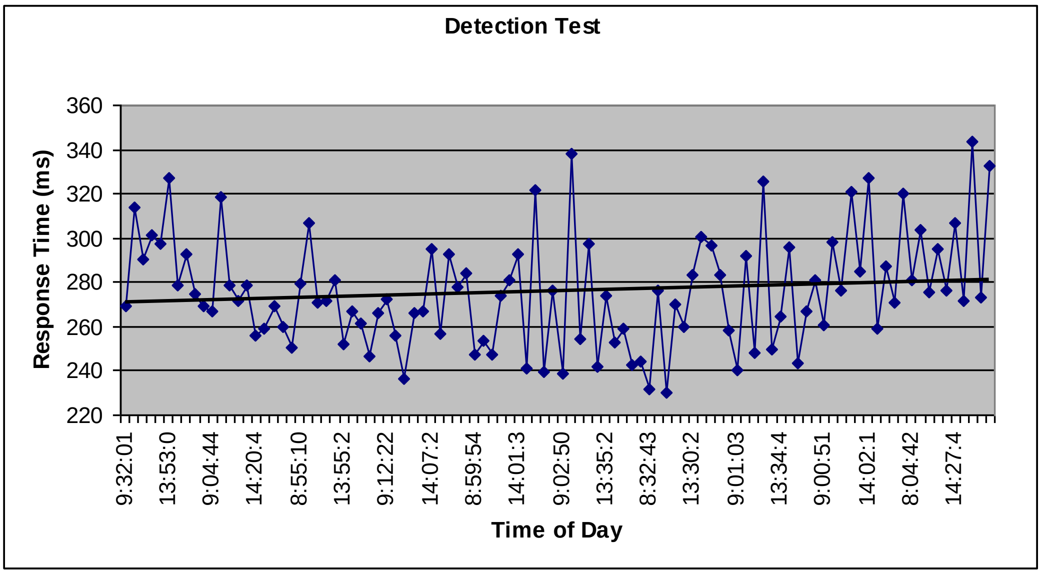

Figure 15 - Detection Test Results for one participant

Figure 15 displays for one task (DET) the variability between tests at paired nodes for mean reaction time for 100 tests. The measured times vary between 230 and 344 ms and generally display an expected diurnal variation with slower times in the afternoon. The trend line indicates that over the 65 days of testing there is a slight increase in average reaction time from 269 ms to 280 ms.

This plotting and trending was done for all data (Appendix 4) to look for patterns and insight on anything unusual such as a learning trend. To assess whether there was any influence of the nodes on the results, a single-sided student t-test was conducted on the paired data. The results of all tests for all valid data are shown in Table 7. Participants are listed in the first column as 'A' through 'M' followed by the paired nodes they tested at; (1 – 2); (2 – 3); or (4 – 1). Blank cells indicate no data was obtained. 14 of the 115 paired testing sessions showed significance. The Groton Maze Learning test was the only task that did not show at least one significant result for a participant. This was also one of the longest tasks that consisted of the most steps. Additionally, one of the steps was completed after all other cognitive tests, meaning that it was not clear how this task might be implemented as a stand-alone test in an operating environment.

	Cognitive Test					
Participant	GMLAvg	IDN	DET	OCL	ONB	TWB
A (1 - 2)	0.28	0.05	0.08	0.50	0.14	0.39
A (2 - 3)	0.20	0.35	0.31	0.21	0.50	0.27
B	0.15	0.43	0.07	0.18	0.14	0.42
C (4 - 1)	0.22	0.04	0.19	0.11	0.48	0.10
C (2 - 3)	0.25	0.41	0.21	0.08	0.49	0.40
D	0.23	0.44	0.10	0.24	0.39	0.27
E	0.13	0.23	0.40	0.20	0.37	0.35
F	0.12	0.27	0.04	0.04	0.22	0.19
G (4 - 1)	0.31	0.13	0.12	0.41	0.25	
G (2 - 3)	0.21	0.05	0.43	0.43	0.03	
H	0.11	1 e-5	1 e-6	0.40	0.03	0.30
I (1 -2)	0.36	0.44	0.47	0.12	0.49	1 e-3
I (2 - 3)	0.25	0.34	0.17	0.43	0.13	0.39
I (3 - 4)	0.24	0.10	0.21	0.32	0.49	0.34
J	0.38	0.39	0.41	0.09	0.11	0.23
K	0.35	0.34	0.39	0.32	0.37	0.06
L (4 - 1)	0.27	0.38	0.27	0.23	0.03	0.49
M (4 -1)	0.34	0.03	0.06	0.01	0.12	0.28
M (2 -3)	0.19	0.47	0.12	0.01	0.08	0.50

Shaded p-values are significant at a 95% confidence level

Table 7 - Comparison of paired t-test p-value results for all participants

Table 8 shows which tests had the most ‘first’ and ‘second’ lowest test results of the t-test for the participants. The Detection (DET) task was the lowest p-value for 4 participants and second lowest for another; the One Card Learning (OCL) task was lowest for 3 participants and second lowest for another 6 participants.

	1sts	2nds
GML	0	0
IDN	3	3
DET	4	1
OCL	3	6
ONB	2	3
TWB	2	1
TOTALS	14	14

Table 8 - Tasks with lowest p-value scores for all participants

Aggregating the results in this way (Table 7 and 8) provided an initial visual on the cognitive tests’ sensitivity to changes in alertness of the participants versus the expectation that alertness levels should naturally vary during the course of the day according to one’s circadian rhythm. It also showed that all except the Groton Maze Learning (GML) task might have some applicability for assessing alertness. The GML task provided no reason to consider it further relative to the other tests. It took the longest to complete and had no significant findings for any participant. It had no significant p-values and was not the lowest or second lowest p-value for any participants.

With the results from the testing, what is sought is a declaration on the aggregate results from the group for the tests and a determination of their strength in measuring any change. Table 7 illustrates varying strength of result with the high-lighted cells showing statistical significance but from individuals who contributed varying amounts of test data. The varying strength of results therefore does not allow any straight forward conclusion for the entire group from this table.

This is not an unusual problem, especially for medical research, where experimental research is often conducted on, and limited to, small groups. The challenge is to properly apportion a representative amount of the extracted results to the final outcome when they have come from a small group with repeated measurements. Mathematically, this is referred to as heteroscedasticity and collinearity. Allison (1999) confirms that the most serious cases of correlated disturbances are likely to come about from sampling the same individuals at different points in time exacerbated by a small sample size. Obtaining large random samples minimizes these factors but in this operational environment that opportunity did not present itself.

To properly assess the effects of time difference tied to nodes at which the paired testing occurred, as well as assess the other variables' influence, linear mixed effects models (Faraway, 2006) was used. The model is termed mixed effects because it consists of fixed effects and random effects required for proper treatment of the measured results. Fixed effects are unknown constants that are estimated parameters to explain (sleep, caffeine, node influence) and the random effect is a variable with estimated parameters that describes the distribution of the random effect (Equation 2).

$y_{ij} \sim X\beta + Zb + \varepsilon_{ij}$ Equation 2

In this case, the dependent variable ‘y’ (cognitive test response results) is shown to be a linear function of the multivariate measurements and estimated parameters where yij denotes the observed response differences measured for individual i at time j. Xβ represents the systemic or fixed components of y and ‘Zb + εij’ is the random and error component. ‘Zb’ and ‘εij’ assume all of the combined effects of all factors influencing measured ‘y’ that are not included or explained in Equation 2 including combined effects of the multivariates as expressed by the collinearity. This random effects term which is assumed to follow a normal distribution with a mean that reflects the overall aggregated baseline response, allows the dependence between measurements from the same participants to be acknowledged. The variance on the random term reflects the heterogeneity in responses between subjects and allows separate intercept terms to be fitted for each participant. At this stage, any factors not measured, which may have contributed to individual alertness levels, will be contained within the random term as will any measurement error associated with the x-values. Several scatter plots were used to look at data relationships per Appendix 4 and two are shown in Figure 16.

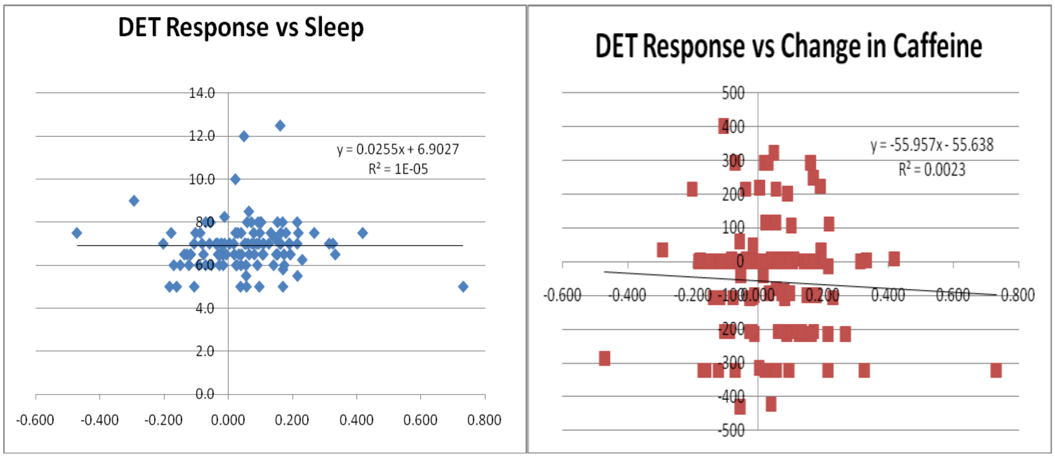

Figure 16 - Scatter Plots to illustrate relationship on Alertness levels.

The analysis was implemented using the lme4 package within the R statistical package (R Development Core Team, 2010) per Appendix 5.

The cognitive screens were analysed to understand the influence of sleep and caffeine and other factors on the results and are shown in Table 9. The effect of node on the outcomes of 5 tasks is shown. There are three levels displayed for each outcome showing (i) a baseline model which has a random intercept term for the participant; (ii) the addition of the effect of node, and; (iii) the additional effect from caffeine, hours awake before test and hours slept prior to test. The third model helps understand the influence of the measured workplace factors on the results and whether their influence might mask the signal from changing alertness due to the circadian cycle. This helps with assessment of how effective the tests might be in an operational setting.

ANOVA Results					
Cognitive Test	**Model**	**Log-Likelihood**	**Chisq**	**df**	**Pr(>Chisq)**
DET	Subject level random effect	-1867.0			
	Effect of Node	-1863.0	7.9620	3	0.04680*
	Additional Effect of Caffeine + Sleep + Hours awake	-1592.8	540.4789		<2e-16***
IDN	Subject level random effect	-1947.3			
	Effect of Node	-1943.7	7.2131	3	0.06541*
	Additional Effect of Caffeine + Sleep + Hours awake	-1626.7	634.0067	3	<2e-16***
OCL	Subject level random effect	-2166.6			
	Effect of Node	-2165.7	1.7942	3	0.6162
	Additional Effect of Caffeine + Sleep + Hours awake	-1849.1	633.3266	3	<2e-16***
ONB	Subject level random effect	-2038.5			
	Effect of Node	-2036.2	4.5849	3	0.20484
	Additional Effect of Caffeine + Sleep + Hours awake	-1712.5	647.5094	3	<2e-16***
TWB	Subject level random effect	-1902.3			
	Effect of Node	-1900.2	4.2601	3	0.23471
	Additional Effect of Caffeine + Sleep + Hours awake	-1594.6	611.0396	3	<2e-16***
GML	Subject level random effect	-2368.8			
	Effect of Node	-2367.0	3.7424	3	0.29066
	Additional Effect of Caffeine + Sleep + Hours awake	-2025.3	683.3685	3.0	<2e-16***
	Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1				

Table 9 - Results from analysis of variance comparing the effects of node and other variables on each of the five outcomes.

The column showing ‘Pr(>Chisq)’ provides indication of the significance of the models which is significant for DET and IDN. For each test, after allowing for the effects of node, the effects of the other measured alertness variables were highly significant (< 0.001).

By further analysing the mean reaction times at each node as well as the variability, a comparison to that expected from the circadian cycle is provided, as shown in Figure 17. The DET and GML test most closely adhere to expectation with slowest and most variable reaction times at node 1; fastest and least varied reaction times at node 2; second slowest at node 3 and; second fastest at node 4. To a lesser degree IDN and OCL also displayed this pattern, as shown in Figure 17 and Table 10.

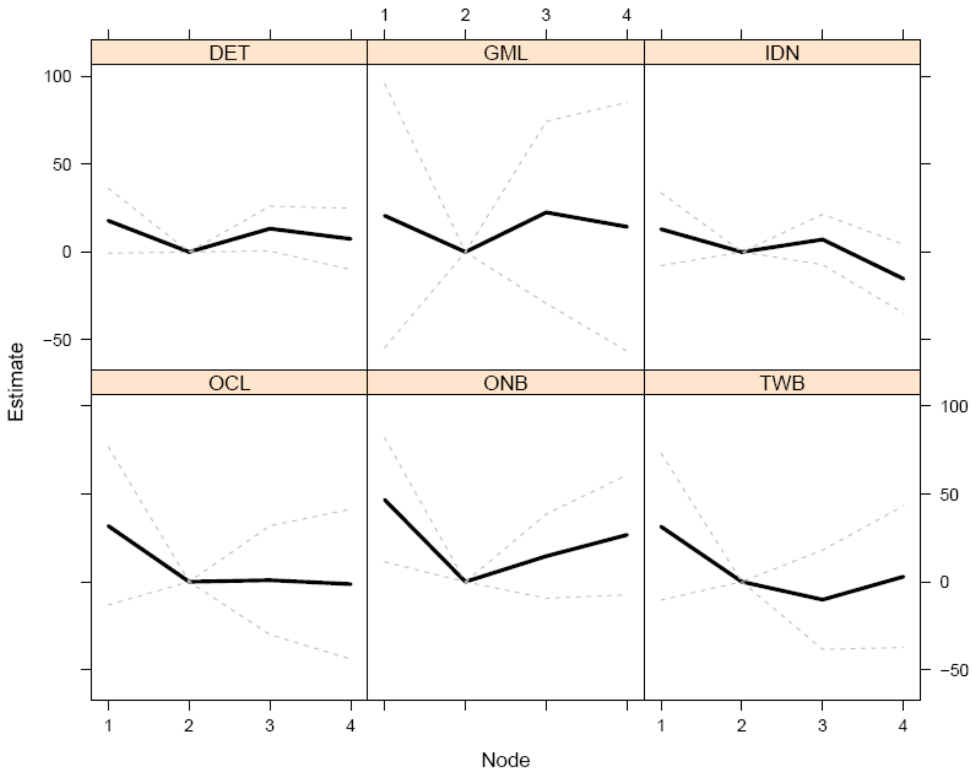

Figure 17 - Effects of node on outcomes from the cognitive tests; estimates and 95% confidence intervals are shown relative to response at node two (given as zero). Effects are estimated allowing for the effects of caffeine, hours awake and hours of sleep taken before test.

For all tests, the first node showed the slowest reaction times coincident with the major low in the circadian cycle however, the pattern for each other node is less clearly tied to the expected profile.

Cognitive Screen	Variable	Model (a) Estimate	Std. Error	t	Model (b) Estimate	Std. Error	t
DET	Intercept (Node 2)	327.28162	23.07	14.19	345.8	31.13	11.11
	Node 1	15.87	8.13	1.96	17.8	9.44	1.88
	Node 3	10.9	4.53	2.41	13.29	6.51	2.04
	Node 4	4.61	8.6	0.54	7.44	8.91	0.83
	Caffeine				-0.005	0.02	-0.26
	Hours sleep				-3.84	1.82	-2.11
	Hours awake				0.016	0.97	0.02
IDN	Intercept (Node 2)	519.93	22.92	22.69	527.45	29.52	17.87
	Node 1	13.7	10.23	1.34	12.99	10.6	1.23
	Node 3	6.75	5.71	1.18	7.1	7.31	0.97
	Node 4	-14.64	10.83	-1.35	-15.27	10	-1.53
	Caffeine				-0.02	0.02	-0.75
	Hours sleep				-3.52	2.05	-1.72
	Hours awake				0.14	1.08	0.13
OCL	Intercept (Node 2)	843.62	29.86	28.25	930.54	44.57	20.88
	Node 1	5.71	19.84	0.29	31.8	22.97	1.38
	Node 3	-11.75	11.08	-1.06	0.88	15.8	0.06
	Node 4	-12.83	21.16	-0.61	-1.35	21.78	-0.06
	Caffeine				-0.02	0.05	-0.5
	Hours sleep				-10.73	4.47	-2.4
	Hours awake				-3.68	2.34	-1.57
ONB	Intercept (Node 2)	668.44	30.39	22	689.67	44.02	15.67
	Node 1	35.18	16.66	2.11	46.69	18.05	2.59
	Node 3	5.19	9.04	0.57	14.56	12.24	1.19
	Node 4	17.28	18.05	0.96	26.71	17.43	1.53
	Caffeine				-0.05	0.04	-1.47
	Hours sleep				-5.49	3.43	-1.6
	Hours awake				-0.6	1.87	-0.32
TWB	Intercept (Node 2)	744.28	34.04	21.86	796.81	48.75	16.35
	Node 1	20.57	19.54	1.05	31.42	21.34	1.47
	Node 3	-13.99	10.45	-1.34	-10.21	14.42	-0.71
	Node 4	-11.01	20.58	-0.54	2.89	20.58	0.14
	Caffeine				-0.05	0.04	-1.14
	Hours sleep				-9.51	4.35	-2.19
	Hours awake				-0.64	2.17	-0.3
GML	Intercept (Node 2)	626.86	69.23	9.05	698.79	88.18	7.92
	Node 1	26.21	33	0.79	20.66	38.36	0.54
	Node 3	34.69	18.43	1.88	22.54	26.46	0.85
	Node 4	11.26	34.93	0.32	14.35	36.17	0.4
	Caffeine				0.004	0.08	0.05
	Hours sleep				-6.2	7.41	-0.84
	Hours awake				-0.51	3.92	-0.13

Model (a) includes a random effect for each subject and the effects of node (baseline is node 2)
Model (b) additionally includes effects of caffeine, hours sleep and hours awake

Table 10 - Summary of models with nodes and outcome of tests with and without allowance for other alertness-related variables.

In addition to the influence of caffeine, hours of sleep before the test and hours awake before the test is shown. Hours of sleep before the test showed a negative effect for all tasks meaning that the less sleep one had before the test, the worse was their result and this was significant for the DET, OCL and TWB tests.

The number of errors per round was the metric of interest for the Groton Maze Learning task. Some tasks also determined and carried an accuracy rate which was the percentage of correct responses versus the number of tries. In no case, was this factor found to be significant as a measure tied to node per Appendix 6.

Both in terms of statistical significance as well as the pattern displayed relative to the nodes of the circadian cycle, the results show that the cognitive screening tests can detect changes in alertness. The tests displayed an ability to detect changes associated with natural changes from the circadian cycle in the presence of confounding elements. Some factors were identified and built into the assessment and some were not and make up the error profile. The results also support others' findings that reaction time is first affected by fatigue and those tasks most reliant on processing reaction times displayed most sensitivity to changes in alertness driven by the circadian. This was most evident in the results from the DET task and IDN task. The results from the linear mixed effects analysis also shows that, in particular the DET task aligns with expectations of changes in alertness associated with the circadian. Participants showed that they performed best and with less variability when they were meant to be alert and worst when they were supposed to be sleeping. Finally, the results are consistent with results found with these same cognitive tests by Falleti's (Falleti et al., 2003) controlled experimental results showing most deterioration in reaction time responses than other performance measures. The OCL results support Turner's (Turner et al., 2007) findings regarding memory loss when lacking adequate sleep.

The findings associated with caffeine, hours awake before testing and hours slept before testing support others' findings as well (Van Dongen, 2004; Lim and Dinges, 2010). These factors can have a significant effect on test performance suggesting follow-on work embrace this as part of the study as well.

7.8 Summary

This study attempted to understand if cognitive tests could be utilized in an operating environment to detect fatigue impairment based on an alertness detection associated with the circadian cycle as possible surrogate measures of worker performance. An important aspect of this study was that it was conducted in a real workplace. Therefore, success from this study had real world relevance in assessing the acceptance and difficulty with taking the tests as well as understanding if there was enough test sensitivity to changes in alertness to be a possible tool in assessing fatigue in the face of normal confounding factors. These results suggest there is significant sensitivity. Values for response speeds varied on average 3.4% between nodes (e.g. 2-3; 10 ms slower) suggesting a range of sensitivity for future studies to account for circadian influence. Additionally, the simplicity of the tests and brevity allowed a full battery to be completed by most participants twice a day. This would not be an accepted or permissible long-term circumstance and it was never contemplated that a full battery would be required. Participants also quickly learned the tests. Only 7 of 447 completed tests failed the integrity check and most participants (11 of 13) passed all integrity checks.

There was a demonstrated failure with this study which adds insight into real work place applications of testing technology. Of the volunteers who started this phase of the research, 11 were unable to provide any usable data. Feedback provided indicated most was due to conflicts with their regular work preventing them from completing the required 2 tests in a shift. While this is not an unreasonable failing, it does help to establish boundaries on expectations of what is possible in work place setting for any type of fatigue testing. The analysis also showed there were no systematic failures associated with any kind of learning effects from repeated testing. Thus there does appear to be preliminary support from this research for stable results with minimal practice effects in a busy work place operating around the clock, consistent with findings from controlled experimental results (Falleti et al., 2006).

This study took workers as they presented themselves, with all individual variability they brought to measures of alertness over the course of their shifts. The results indicate that this is a dominant factor in discussions on fatigue. The size of the population that was available for this study clearly limited the strength of the findings of other specific factors influencing results. Indeed, within the individual-specific factors, there is a lack of visibility as to what might further explain the range of differences in response of participants. Some worker comments may give insight. In a work environment, additional factors can be obtained over time but due to the sensitive nature of collecting personal information from workers, extracting this data will always be a challenge.

The results from this part of the research with the cognitive tests will be used to take some forward for further validation with actigraph-based fatigue models in Chapter 9.

CHAPTER 8 – MEASURING SLEEP TO VALIDATE FATIGUE MEASUREMENTS

8.1 Introduction

Sleep is a key component to all discussions on fatigue. It is the only naturally occurring antidote to fatigue. Some tools used to manage fatigue in the workplace have been introduced that only focus on sleep. Actigraphs are one device now extensively used in research and commercial applications, to assist in fatigue management and to aid sleep diagnostics and repair. Originally actigraphs were rather bulky instruments worn to gather data for interpretation of rest/wake cycles. The interpreted sleep/wake cycles were in turn interpreted against norms for adequacy of sleep. The norms evolved into rather complex and sophisticated models. Now with the evolution of computers, memory chips and software, much of what was done in sleep labs with large desk-top computers can be embedded into a package with the actigraph, to give more immediate displays and feedback associated with individuals' sleep and effectiveness levels. This chapter focuses on establishing grounds for using actigraphs and sleep models as a basis for validating cognitive screening results. The selected technology is explained and reviewed with focus on the technology developed by Hursh, as it was used extensively in this research. The Hursh patent (Hursh et al., 2008) for this technology is heavily drawn upon for this analysis.

8.2 Actigraphy

The first actigraph device was described by Krupke in 1978 (Lee-Chiong, 2008). Early versions intended to measure motion of a patient and correlate to restlessness and wakefulness. Actigraphs are built around accelerometers which provide initial reading on the movement of the limb they are worn on. Combinations of electronics and software separate and classify the type of movement as either rest/sleep or wake. To facilitate the determination, modern units also have time monitoring built in and can double as a watch as the unit collects data. To obtain accurate data on individuals, the units must be worn continuously. To allow this, modern units are light-weight, water-proof in packaging for continuous wear even while bathing or swimming. The units carry significant battery life and memory to collect data for long periods. During this research units were targeted for 30 days of continuous use. There are several commercial operations which offer packages and services built around these technologies for a user fee. Fatigue Science (Fatigue Science, 2012) is one such entity, whose units were used in this research, that offer integrated services including full reports based on data analysis. Figure 18 shows one of the actigraphs worn during this research.

Figure 18 - Wrist-worn actigraph from Fatigue Science

In recent years, development has led to an actigraph unit that is robust for continuous wear and has significant embedded analyses capability. Data can be stored and easily offloaded to companion programmes for full analysis. Balkin and Belenky worked on developing such actigraph-based devices to assist managing sleep in military applications over a decade ago (Balkin et al., 2002). Concurrently, significant effort was being put into modelling effectiveness based on sleep. Such units were also worn by truckers as part of the programme led by Dinges in 2005 and referenced in Chapter 6. They have also been worn by construction workers in 2007 (Powell and Copping, 2010). The units used in 2007 have been superseded numerous times by new models. Significant improvement has been made in robustness, battery life and memory capacity. In all cases, they proved valuable in gathering insight into the work/rest cycles of workers and were stepping stones to our current technology.

8.3 Sleep-based Fatigue Models

A significant advantage of the actigraphs is that they provide highly reliable data from the sleep/wake cycles without themselves interfering with the sleep. Unlike electroencephalographic measurements of brain states, actigraphy is non-invasive and otherwise does not require wearers to move from their regular schedules or surroundings. Thus data is collected as close to actual circumstances as possible, but it is also accurate to results obtained from full electroencephalographic data (Morgenthaler et al., 2007). Additionally, it is able to collect data over multiple sleep/wake cycles to more accurately represent fatigue cycles based on any inadequate rest as well as reflect performance relative to natural oscillations of the circadian rhythm.

In all cases, the data is interpreted for further declarations about the wearer and for the purpose the data was collected. As actigraphs were evolving in the last decade so were sleep models which are fed estimates of individuals' sleep/wake cycles. In 2002, at a workshop focused on modelling fatigue and performance, six separate teams comprised of leading sleep scientists from around the world developed models to predict fatigue and performance based on five different scenarios. Predicted results were compared with experimental data and using mixed effects regression the models were compared. A conclusion reached was that all of the models were capable of predicting performance responses and did so quite well for three scenarios. The variance between the six models was reported to be quite small. The best performing was one put forward by Professor Hursh (Van Dongen, 2004).

8.3.1 Hursh's SAFTE Model

Through the 1980s and 1990s the Walter Reed Army Institute of Research (WRAIR) funded development of a bio-mathematical model to provide guidance to countermeasures as part of a fatigue management initiative with military personnel (Eddy and Hursh, 2001). This led to the Hursh team having the top performance predictions at the Seattle workshop in 2002. This was extended into the first funding to develop an actigraph with an embedded sleep model, which was done by Hursh while at Scientific Applications International Corporation (SAIC). Hursh went on to expand and develop the model and ready it for more practical real world applications. The Hursh model was initially accepted by the US Department of Defence as their warfighter fatigue model (Hursh et al., 2004). Hursh's model, introduced as the Sleep, Activity, Fatigue, and Task Effectiveness (SAFTE) Model, was used as the core for the Fatigue Avoidance Scheduling Tool (FAST). A patent was finally granted for the Hursh model in 2008 (Hursh et al., 2008).

The Hursh SAFTE Model shown in Figure 19 has at its heart a sleep reservoir that gets emptied when awake and filled when sleeping.

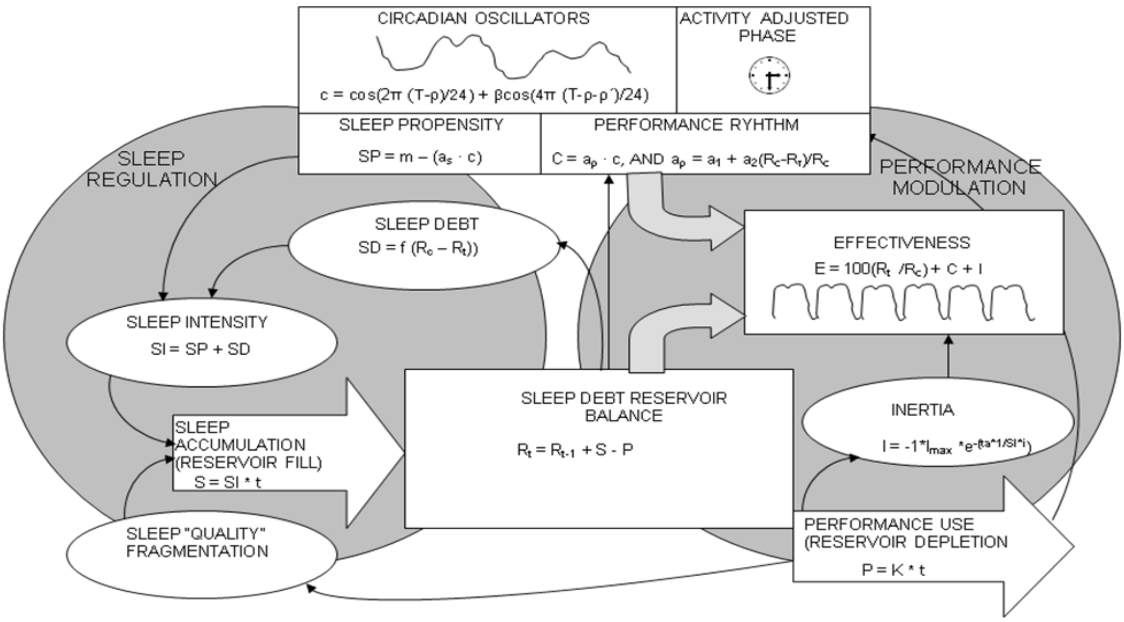

Figure 19 - Hursh's SAFTE Model reconstructed from (Hursh et al., 2008)

8.3.1.1 Modelling the Homeostatic Process

The two large circles in Figure 19 represent a sleep regulation cycle and performance modulation cycle, between which is a constant interplay of sleep factors on the left and performance factors on the right. The reservoir is one of two bridges between these two; the circadian influence is the other. This relationship is tied to the homeostatic (Chapter 3) process of sleep defining cognitive output. Depending on the level in the sleep debt reservoir, one's effectiveness, and capacity to perform is driven but limited and governed by other factors, especially the circadian cycle. This reservoir level is therefore constantly changing, either filling or draining in a kind of oscillation referred to as the sleep/wake cycle. Hursh uses minutes of effective sleep as the measure of the reservoir's currency. It is either being spent or saved according to $R_t = R_{t-1} + S - P$ where 'R_t' represents the measure in increments of minutes at time 't'. The model is capable of updating and reporting on a minute-by-minute basis, making it close to real-time in iteration. This generates a significant amount of information. As an example, this research project collected data on a minute-by-minute basis per wearer for a month and therefore generated close to 4 million lines of data. The measure of 'S' comes from the 'Sleep' side and the measure of 'P' comes from the usage or 'Performance' side.

8.3.1.2 Sleep Debt

The model also builds in a modulating loop associated with sleep debt. In this case the assumption is that longer periods of sub-optimal sleep feeds back as debt but at the same time increases or strengthens the intensity of sleep according to Sleep Intensity (SI) = Sleep Propensity (SP) + Sleep Debt (SD). This defines the rate of fill of the reservoir. The underlying assumption is that there is a dynamic interplay between sleep debt, sleep intensity and rate of fill such that a new, albeit a lower equilibrium is approached with inadequate sleep. This is opposed to a continuous fall off a cliff as would occur if sleep debt increased uncontrolled. In cases of inadequate sleep, every hour less than 8 hours adds to an accumulating sleep debt that cannot be repaid as long as an individual stays on a sleep schedule short of 8 hours. This relates to debate (Chapter 3) on how much debt can be accumulated. The model sets limits to explain the effects of sleep loss to prevent complete performance collapse. As reported by Breus (2012), a US Center for Disease Control (CDC) study with over 15,000 responses to a National Health Interview indicated 30% of workers were not getting more than 6 hours of sleep per day, which is significantly below the stipulated 8 hours. If this apparent short fall was not somehow accounted for physiologically, one third of the population would grind to a halt over time, due to a burden of debt and inability to repay. Additionally, the level of performance must be explained if 8 hours is required for full performance and major portions of our population have on-going performance at 6 hours sleep per day.

8.3.1.3 Sleep Accumulation

To account for this requirement, an innovative aspect of the model is that it effectively sets boundaries around sleep loss. As the model uses 8 hours as the sleep requirement for each 24-hour period, all time short of the 8 hours per 24-hour not slept is added as a debt. If 8 hours is obtained from the sleep accumulation and it is entirely consumed in performance use, the sleep-debt loop is 0. The model starts in balance at 1 hours of sleep equalling 2 hours of performance. Short-term loss of sleep is simply accounted for as a debt. However, per 8.3.1.2 above, Sleep Intensity (SI) increases by the amount of the debt. 'SI' feeds forward as the key component of Sleep Accumulation (S). If an individual's schedule prevents full sleep, for instance 6 hours per night, the sleep accumulation, 'S' increases by the sleep debt such that eventually 6 hours of sleep will balance 18 hours awake with no further debt realized. The model maintains a lower over-all performance level than with full sleep but the model will converge to the new equilibrium with continuous inadequate long term sleep. This is consistent with others' findings (Monk, 1991, Akerstedt and Folkard, 1995) and helps align with reality. It neither suggests they are fully functioning but rather performing at new and lower performance equilibrium. The model does set a minimum boundary however at 4 hours of sleep requirement after which the model does not reach an equilibrium state meaning performance capacity will approach zero.

This feedback loop from the sleep reservoir defines sleep debt which is a daily moving variable. There is a limitation in the model here. As the model specifies a fixed amount of sleep required for all of 8 hours per day, it will systematically increase debt for everyone by the amount short of 8 hours per day, even if their physiological requirement is short of 8 hours sleep per 24 hours. In a similar fashion, anyone who needs more than 8 hours of sleep but only takes 8 hours will not show any degradation in performance due to inadequate sleep. This is a challenging aspect of the model as it is built on population norms without knowing individual requirements, which is all it can do. A clear enhancement would be to have an adaptive model to reflect individual needs. When used over a longer period of time as was done in this research, an adaptive algorithm that started to score the individual relative to their emerging normal data over time would be a significant feature. At least, the algorithm should default to a required 8 hours of sleep per 24-hour period but allow this to be a tuning variable that could be adjusted within limits based on information provided by participants.

8.3.1.4 Compensating for Quality of Sleep

The concept used in the SAFTE model parallels earlier work, such as Monk's (1991) three process model of sleep and performance which was further developed by Folkard and Akerstedt (1991) and adapted by Fletcher (Fletcher et al., 2003). On the sleep side, the sleep debt reservoir is filled with sleep accumulation 'S' which is determined by sleep intensity (SI) and time slept 't' according to S = SI * t. It is adjusted according to sleep quality measured as a percentage of time slept in the allotted time for sleeping. Or conversely, it provides a measure of how much wake time there was during the period of sleep. Anything short of 100%-quality sleep subtracts from sleep accumulation. The model refers to this as 'fragmentation' and it is actually a product of the number of waking episodes during sleep and their duration. This accounts for how much sleep is lost from the total time spent trying to sleep. This is an impossible measurement for individuals to know unless monitored. Without monitoring, estimating time awake during a sleep session or time spent not sleeping is extremely difficult and for most people it is an unknown. It is a very important aspect of sleep however, as it ultimately defines the sleep taken and the debt owed. Actigraphs help this area immensely and can even identify individuals with sleep disorders such as sleep apnoea. The number of waking episodes and duration of the waking episodes can and has been normalized over healthy populations. As such, when individuals' wake episodes fall outside of normal boundaries, an abnormal condition can be flagged.

The Hursh model also builds in a couple of additional refinements on sleep quality. It assumes it takes 5 minutes to fall asleep so each time there is a waking episode, there is a an additional 5 minutes applied to not sleeping. This is an arbitrary assignment which could add a significant penalty for a restless night. Sleep apnoea sufferers for instance, would only have to wake 20 times during a night to lose an hour in the model. In the analysis from the model, up to 16 wake episodes per night is considered normal.

8.3.1.5 The Circadian Influence on Sleep and Performance

The circadian cycle has a significant influence on sleep, based on the timing of sleep. In the model this is the Sleep Propensity 'SP'; ($SP = m - (a_s * c)$). The circadian influence 'c' modulates the MESOR, 'm'. If sleep is taken according to one's natural circadian rhythm at a time matching natural sleep periods, the sleep propensity is high. As presented in Figure 8 the circadian is not a simple repeating wave. This has been developed from other previous foundational research (Folkard, 1975; Monk and Embry, 1981; Folkard and Akerstedt, 1987) by principally mapping out the endogenous temperature profile. It is recognized to have two major alertness points and two minor alertness points in a 24-hour period. Hursh modelled this with two additive oscillating out-of-phase cosine waves of different periods; one 12 hours and one 24 hours ($c = \cos(2\pi (T-\rho)/24) + \beta\cos(4\pi (T-\rho-\rho')/24)$). In this way, Hursh believed he was better able to align changes in alertness profile with body temperature patterns to estimate the circadian cycle. Hursh uses the output to explain both changes in performance and propensity to sleep. Just as one's propensity to sleep increases if one is trying to sleep in harmony with their natural circadian rhythm, one's cognitive performance will peak if aligned with peak periods of alertness from the circadian. Or conversely, if one is trying to work when their natural rhythm is trying to shut them down for sleep, performance will be lower. Eddy and Hursh (2001) reference several reports that confirm the bimodal pattern in the circadian cycle's influence on performance. It is also consistent with findings of Dinges (Dinges, 1995). It is most evident from Meta data studies on time-of-day of accident rates where accidents peak at the low points of alertness suggested by the circadian. In this way, the circadian influences both propensity to sleep and performance.

For performance, it is given by the value of C in ($C = a_p * c$, AND $a_p = a_1 + a_2 * (R_c-R_T)/R_c$). R_c is an arbitrary term for total reservoir capacity which is set to 2880 units. Hursh uses the logic that is derived from the standard 8 hours of sleep required per 24-hour period for healthy adult populations. This leaves 16 hours per day for 'performance' meaning that if the homeostatic process is in balance, wake hours draw down the reservoir at half the rate per hour that sleep fills it at per hour. He further cites (Hursh et al., 2008, p8) that total sleep deprivation studies have shown that 'cognitive capacity depletes at a rate of about 25% per day' for complete depletion in 4 days. There is a clear error in the patent stating this as a rate of depletion. Based on his calculation, he intended that it depletes at a rate such that all cognitive functionality is gone in 4 days or at a linear rate of 25 percentage points per day from 100%. Therefore, he calculates 4 days X 24 hours per day X 60 min/hour X 0.5 = 2880 units of sleep. This is equivalent to the cognitive depletion rate of full capacity being lost without sleep in 4 days. No reference is given for this claim however, Eddy and Hursh (2001, p13) suggested 'the sleep homeostat is not infinitely elastic:…' and '…performance capacity will gradually deplete to zero...'. The model also seeks a zero equilibrium if less than 4 hours of sleep is taken per night.

These boundaries seem to be reasonable assumptions for the model, given their distance from normal sleep levels taken by healthy adult populations.

8.3.1.6 Performance Draw on Sleep

On the 'performance' side of the SAFTE model, the reservoir is depleted according to a Performance use factor: $P = K * t$. As previously explained the 'K' factor is set at 0.5 sleep units per hour awake versus the 1 unit per hour sleeping added from sleep. The important output from the modelling exercise is a declaration on the effectiveness at which an individual is performing.

8.3.1.7 Individual Effectiveness

Effectiveness is governed by the amount of currency in the reservoir bank that can be spent and the alignment of the performance versus the circadian cycle, both previously discussed according to: $E = 100*(R_t/R_c) + C + I$. This is just a pure percentage number versus a theoretical 100% available.

8.3.1.8 Sleep Inertia

The SAFTE model has another modulating factor between performance and individual effectiveness, which is inertia; the 'I' – term in 8.1.3.7. This modulator effectively reduces performance within the first 2 hours after waking based on a belief that a kind of groggy state exists from point of waking which decays exponentially according to $I = -1*I_{max}*e^{-(ta*1/SI*i)}$ over the 2-hour waking period. This is a potential area for improvement based on real-world circumstances. A limitation of the model throughout as applied to real-world situations is that it is purely sleep based. It does not account for any other elements associated with alertness. A common practice for workers is that while they are in their 'wake-up' phase, they will consume food and beverages which can have a dramatic effect on alertness. Caffeine in particular is a stimulant deliberately taken to promote or facilitate performance after waking. A suggested improvement for the model would be to allow for modulation of the 'Inertia' impact by factoring in caffeine. This could be done by recording cups consumed of the type of beverage and recording it as a kind of tuning factor, as was done with this research. Another concern for this factor would be that depending on the time of waking relative to one's natural sleep cycle, individuals will experience different degrees of inertia. Someone waking naturally out of REM sleep would not be expected to suffer as much inertia as someone woken by alarm clock out of deep sleep. The improvement would be complex as it would require understanding of an individual's sleep phase on waking. To build in caffeine effects would also be challenging as they would require personal responses to caffeine.

8.3.1.9 Shift Changes and Time Zone Changes

There is one other feature of the SAFTE model called the 'activity adjusted phase'. It is associated with modelling circadian rhythm shifts such as when travelling across time zones or when moving to a new work-shift pattern. This feature is clearly useful for aviation applications to facilitate predictions of performance and for shift-workers with desynchronized circadian influences. Time zone changes were not a factor for this research but there were several shift-workers monitored. As discussed in Section 7.5, the approach taken for short term shift changes (< 3 days) was that circadian shifts were not significant. This is supported by others' findings as a valid assumption and with the few participants of phase 1 of this research. Chapter 9's phase II of this research includes many more shift workers and their circadian alignment was determined with this model which effectively follows others' research findings. As the model iterates every minute for sleep/rest activity, it looks for alignment with the model's default normal cycle. It does this by continuously calculating a running average of the 'awake' hours and checking for alignment with the acrophase in the model. The acrophase is advanced or delayed according to finding and does so according to an understanding that the circadian cycle shifts slowly in humans. The model works with the assumption that the average person will shift their circadian about 1.5 days per hour of desynchronization when advancing (as in crossing time zones in an eastward direction) or about 1 day per hour of shift if delaying (travelling west or moving shift forward in a clockwise movement) (Dinges et al., 1996). The net effect of desychronosis is a performance penalty coming from the circadian influence. This adaptive functionality is a strong feature of this model.

8.4 Model and Measurement

The extensive review of the Hursh SAFTE model was undertaken as it is the model which not only has demonstrated strongest predictability of sleep-based performance; it is also exclusively used by Fatigue Science as the underpinning of their performance analytics with actigraphy. The extensions supplied with the Fatigue Science analytical package, includes a presentation of predicted performance displayed as degradation from 100% and also with an associated performance equivalency to a Blood Alcohol Concentration (BAC) level (Appendix 7). BAC is the prevailing measurement used to declare impairment from alcohol consumption (Figure 20).

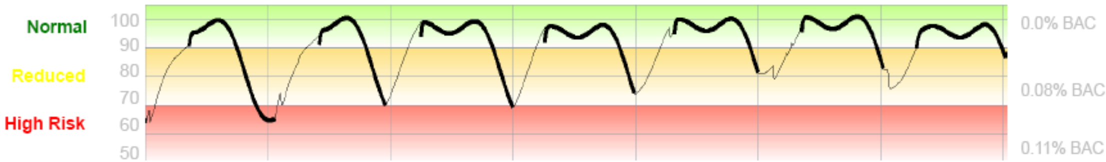

Figure 20 - Typical graph of predicted effectiveness levels from Fatigue Science report showing Blood Alcohol Concentration equivalencies

This builds on the original work of Dawson and Reid (1997) and puts the fatigue-related impairment into new perspective for the average user. The unpublished equivalency used in conversion is shown in Table 11 which does not reflect refinements in BAC equivalencies which have been advanced by Falleti (Falleti et al., 2003).

	Mental Effectiveness Range	Percentage of Time	Blood Alcohol Equivalence (BAC%)	Risk of Accident / Serious Error
Normal	90 - 100%	72%	0.00%	Very Low
Reduced	80 - 90%	20%		Low
	70 - 80%	5%	0.05%	Elevated
High Risk	60 - 70%	3%	>0.08%	High
	0 - 60%	0%	>0.11%	Very High

Table 11 - Fatigue Science conversion of Effectiveness Levels to BAC

The net effect is that it perhaps over-states the BAC equivalency but it is not a core metric and rather one used for effect and better understanding of the impairment level attained.

Typical reports from the actigraphy-collected data consists of a day-by-day display of sleep and wake data as shown in Figure 21. The striations in the blue sleep area are wake periods during sleep.

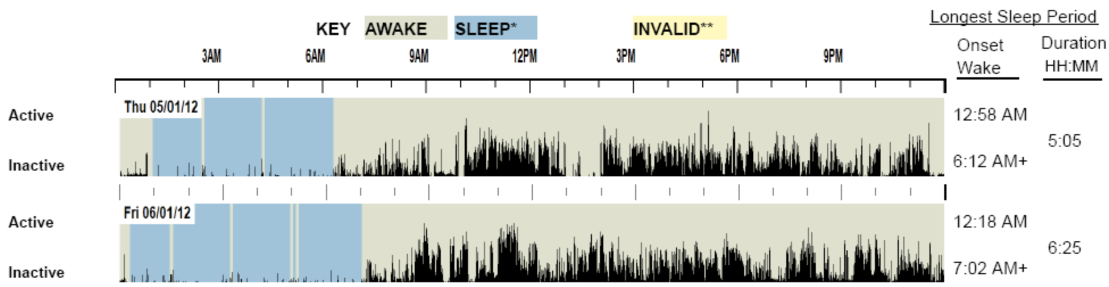

Figure 21 - Actigraph-measured sleep/wake cycle

The collected data is analysed with sleep statistics reported for the period of data collection (Figure 22) and finally presented as companion day-by-day effectiveness profiles, after flowing through the SAFTE model.

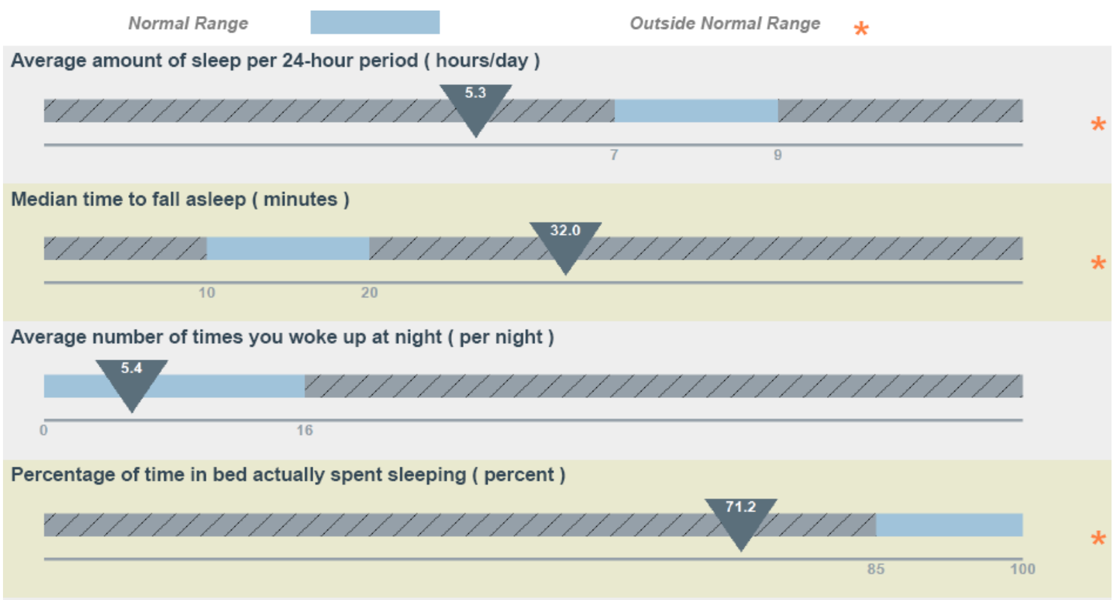

Figure 22 - Sleep statistics derived from actigraphy

The statistics shown in Figure 22 are derived from the period of time an actigraph is worn. For most participants in this research it was for about a month, making the statistics quite accurate for each individual. There are two distinct aspects to the report; the actual measurements of the sleep/wake cycle with derived statistics associated with it, and the information derived by passing it through the SAFTE model.

8.5 Summary

The patented SAFTE model was introduced and reviewed in detail in this chapter as it played an important foundational role in this research. Accepted globally by industrial and other commercial operations, as well as the US military, as a good tool for shift scheduling, the Fatigue Science actigraph data feed of the SAFTE model provides a reasonable standard for assessing fatigue-impairment in the workplace. The validation and testing that this model has undergone (Van Dongen, 2004,Hursh et al., 2006) makes it an excellent candidate as the standard to use for comparison to the capability of cognitive testing in an operational setting which is what was done and reported on in Chapter 9.

CHAPTER 9 – CORRELATING REAL-TIME MEASUREMENTS WITH SLEEP

9.1 Introduction

This chapter presents the final phase of this research to test whether any of the selected cognitive tests, which correlated to alertness changes, have any correlation to predicted individual effectiveness as determined by a proven and elaborate actigraph-fed fatigue model. Results are taken from actual workers on different shifts who had their sleep/wake patterns measured 24 hours per day for a month. In conjunction with the use of actigraphy and the SAFTE sleep-based fatigue model, effectiveness levels were assessed and compared to results from random cognitive testing. Results for correlation between a screening tool reading and fatigue-impairment are established and analysed using linear mixed effects models. The models show that indeed there is significant correlation between the results from the cognitive testing and the sleep-based effectiveness modelling.

9.2 Background

Chapter 7 presented the phase I research and results which confirmed that simple cognitive tests were sensitive enough and effective enough to detect small changes in alertness associated with the circadian cycle, in the workplace. This was the first known research of this type to assess use of cognitive tests as screens for fatigue-impairment in the workplace with all confounding elements of alertness uncontrolled. Selection of appropriate cognitive screening tests, based on sensitivity to fatigue-related impairment and resistance to other masking agents such as caffeine, identified three tests from the battery available as having significant potential for use in testing individual effectiveness. They were the Detection (DET) task; the Identification (IDN) task and the One Card Learning (OCL) task. The attractiveness for developing and using these tests includes their simplicity, quick completion time, and potential low cost to assess workplace impairment. As impairment influences so many elements of an accident chain from decision making, recall of learned responses, alertness, learning and reaction time, screening it out remains a key mitigation in any safety management system. The use of a brief battery set to identify fatigue-related cognitive impairment has been demonstrated (Falleti et al., 2003) in a controlled environment and now demonstrated in the uncontrolled work environment. To determine if the cognitive tests might be possible tools for fatigue-related impairment management as a practical, real-time objective test in the workplace, a comparison to other proven tools was required. The purpose of this final phase of the research was to compare and assess output from the selected cognitive tests as surrogate markers for effectiveness versus the effectiveness of the SAFTE model fed from actual actigraph-monitored sleep input.

9.3 Methods

9.3.1 Funding

In preparation for this portion of the research, funding support was necessary. The key pieces of the research were the usage of the cognitive tests, the actigraphs and the SAFTE-based assessment software. All are commercially available. The first part of the research shown in Chapter 7 was supported without fee by CogState Research of Australia. They agreed to extend access to their battery of tests and database for this phase as well, without fee. CogState Research also provided a named resource to support with any arising research issues.

Fatigue Science, of Honolulu, Hawaii, supply for a fee, actigraph-based analysis of sleep and fatigue based on the SAFTE model. They were approached for support with this research and made some accommodation to their regular services in support of the research but still required a significant fee to use their technology. An application was made to the local Worker’s Compensation Board (WorkSafeBC) for the project to obtain financial support to cover the costs associated with the actigraph portion and it succeeded, in exchange for a full report and presentation on findings back to the Board. Fatigue Science normally temporarily provides the actigraphs and produce reports back based on the recovered data from the units for a report fee. For this research, given the research intent, length, and participant numbers planned, they agreed to sell the actigraphs for fee and lease all software associated with extracting the actigraphs' data and analysing it. Prior to unlocking the software which was also required to activate the actigraphs for each session, a one-day webinar was provided for training by Fatigue Science scientists who were identified as on-going contacts and support for any issues/questions arising during the research. Funding allowed for 25 actigraphs to be purchased and for leasing of the software for 6 months which defined the window for project completion. The software also provided an output to Crystal Reports which allowed generation of a personal sleep report for each participant.

9.3.2 Participants

To conduct this research, a broadcast went out for volunteers willing to wear an actigraph for a month to collect information on their personal sleep/wake cycles and activity whilst periodically randomly completing the pre-selected cognitive tests. In exchange, each would receive a personalized report assessing their sleep. With the 25 actigraphs purchased, the intent was to complete four full rotations of one month each requiring totally 100 volunteer workers to participate. Participants were taken on a first come first serve basis and participated only with the expectation of receiving a personalized report on their sleep after completion of a Memorandum of Understanding shown in Appendix 8. No worker was screened out from participating and the only requirement was that they had access to a computer, a willingness to conduct a cognitive screening test regularly and keep a sleep log for a month. Unavailability of a personal computer did not impact this study. There was no objective to obtain any proportional representation from subgroups other than to collect data around the clock using shift workers and have males and females represented to test gender effects. Testing expectations and instructions are shown in Appendix 9.

The workers were followed through different shifts and shifts of different lengths over the course of the month they wore the actigraph during the period September 2011 to January 2012. All participants in this research came from normal full-time working roles in metropolitan Vancouver. Participants included office workers, field-based workers, police, teachers, professional and trades personnel. The majority was employed by a subsidiary of Canada's largest engineering and construction company that operates and maintains an automated commuter rail system in a continuous operating environment of 24 hours per day, seven days per week. To secure 100 participants, others from the broader pool who were interested were accepted. The actigraphs were worn continuously by the volunteer workers for one month each.

9.3.3 Cognitive Testing

Each participant worker was given access to the selected cognitive tests developed by CogState Research (CogState, Melbourne Australia) with instruction per Appendix 9 to try and complete at least one round of tests per day at a time convenient to them.

The tests used were selected from the tests discussed in Chapter 7 which confirmed significant sensitivity to changes in alertness. The tests are as described in Chapter 7. Detection (DET) was a simple reaction-time based test requiring participants to monitor a card presented face down on the computer monitor and hit the 'K' key on the computer key board to indicate that "YES" it had turned over. The Identification (IDN) task was the choice reaction-time test requiring participants to monitor a card presented face downed and hit the 'K' key for "YES" it is red or the 'D' key for "NO" it is not red. The third task the One Card Learning (OCL) test required participants to decide whether they had seen the card before or not. The 'K' key was hit for "YES" it had been seen before, or the 'D' key was hit for "NO" it had not been seen before. This was the longest of the three tests and in addition to processing reaction time recorded errors to determine accuracy of memory.

There was no stipulated time for conducting the tests. With prior knowledge that the actigraphs were capable of producing a minute-by-minute measure of effectiveness, flexibility on timing of test taking was offered to suit individual schedules. Unlike the first testing with a battery of six cognitive tests, the total time to test with the three tests chosen was less than one fifth of the time as they did not need to be done twice a day. However, they had a month to contribute test results. Participants were given a USB memory stick which gave them access to the cognitive tests but it required a personal computer or access to one to conduct their cognitive tests.

Each participant went through a comprehensive tutorial on the testing and was left with the instructions of Appendix 9. The target was at least one test per day per participant. After being shown how to complete the 3-test battery (DET, IDN and OCL), participants were left to do as many practice rounds as they felt they needed to do before submitting 'base line' test results. When participants finished tests they transferred the data to the CogState database on-line, along with any test comments at which point they were visible for review and analysis. Part of the comments to be attached included a record of the number, type, size and time of consumption of any caffeinated drinks taken. Participants were also invited to provide any other information to the researcher which they felt was relevant in explaining their test results.

Participants were also given a blank template for the month they were wearing the actigraph to keep a log of their sleep. The log was to note sleep times and naps. The purpose of the logs was to verify results from participants, especially those who were working abnormal shifts and sleeping abnormal times.

9.3.4 Actigraphy-based effectiveness modelling

The model used as a sleep-based effectiveness model was heavily tied to accurate sleep data for accurate predictions. A commonly cited issue is inaccuracy of sleep data when it is left solely to individuals to estimate. In part this is due to individuals loosely correlating sleep with time in bed; in part this is a real issue associated with individuals simply not understanding the quality of their sleep and how much wake time is accumulated in a normal night whilst attempting to sleep. Additionally, tracking this data accurately over any extended time like a month will accumulate error in the estimated results at any given time. This is particularly true if identifying all wake periods and accurately measuring their duration. Also, naps and rest periods may be easily missed. An advantage of the technology used in this research is that the sleep model is actually fed by actigraph-measurements of the sleep-wake cycle, strengthening the validity of all sleep values versus personal estimates. Especially over extended study time-frames such as this study, which spanned holidays and multiple work/rest schedules solely relying on sleep logs was expected to carry substantial error and therefore was only used as a guide and was most useful for shift changes.

The actigraph designs were those shown in Figure 18 and stored the data they collected in the unit's memory for the full duration of the wearing. They are a packaged wrist-worn device like a wrist watch with user interface features. The units were light-weight and attractive, which facilitated acceptance by the participants who were signing up to wear them for at least a month. The units doubled as a watch for those who normally wore a wrist watch, which also lessened the burden of wearing it for an extended period. Most importantly, the units were advertised as being water-proof for continuous wear and as having battery life good for 60 days.

The actigraph units were activated via a wireless connection to a computer and thereafter immediately fit to participants. During activation, the actigraphs were programmed with session information including coordinates for the geographic location and the name of the participant for verification of data when the unit was returned. Individuals were instructed to wear their 'sleep bracelet' continuously and only take it off if there was risk of shock to the unit. This meant the units were to be worn while showering, bathing, and swimming but came off for contact sports. In case they were ever taken off, participants were instructed to put them back on within 4 hours, which was the established time limit for removal by the supplier. There was some possibility of the sleep lab in Honolulu ‘stitching’ data together if the units were removed and participants forgot to put them back on, but for integrity of results the lab would not do so if removal periods occurred during sleep periods. For full integrity and continuity of the sleep units based on duration and quality of sleep, visibility is needed to confirm sleep actually occurred. This visibility is not possible if the actigraph is removed. Wake periods on the other hand do not have a quality concern so it was not as critical to see wearers were awake when the unit was removed. If removed during a wake period the data could be stitched together as ‘awake’ hours.

The units collected data continuously and recorded it by the minute. A low-power wireless interface was used to activate and off-load data from their memory storage. Once the data was off-loaded to a computer, it was analysed to establish the activity levels of the individuals consistent with their sleep/rest and wake cycles. All statistics associated with their sleep was then derived, graphed and reported. This data also fed the SAFTE effectiveness model to give minute-by-minute estimates of individual effectiveness over the course of the time the actigraph was worn.

9.3.5 Caffeine

As verified in Chapter 7, caffeine was found to have significant influence on results of alertness. This final phase of research implemented the same process for tracking consumption of drinks containing caffeine. Estimates of caffeine in the participants at the time of testing were then made according to the same formula previously used:

$$C = C_0 - C_0 * (t - 1)/10 \qquad \textbf{Equation 1}$$

C is the estimate of residual caffeine affecting alertness in mg;

C_0 is the estimated amount of caffeine consumed with each particular drink in mg

t is the number of hours the drink was consumed before the test in hours.

9.4 Results

During the period September, 2011 to January, 2012, 4 groups of 25 participants provided data and wore an actigraph for a combined 2320 days or about 55,680 hours. The workers ranged in age from 23 to 63 and had an average age of 40.2 years. 46 females and 54 males participated; 68 were shift workers. Of the 100 participants, 62 held field-based jobs. There were 5 participants who did not have valid data from a full month wearing the actigraph due to removing it for more than 4 hours, causing a session termination. Most participants kept logs of their sleep; 14 did not. Wake periods and identified short sleeps were missed on logs.

All participants except one wore an actigraph for their designated month. The range of days worn was 19 to 42 with average 31 and median 31 days. Several participants commented they liked their look and liked it as a watch. There were no operating issues during the research project with the units. In total, by-the-minute, 4.3 million records of results and associated data was derived. Complete files are in Appendix 10.

Cognitive test data was extracted from the CogState database and prior to use and analysis underwent integrity and completeness checks utilizing CogState built-in checks as described in Chapter 7. A total of 2321 cognitive test sessions were completed. All participants completed some cognitive tests which ranged from 5 to 81 completed tests in a month by participants. The average was 23, or about 1 per work day for the four months of testing. A session was intended to include a battery of three tests DET, IDN, and OCL but of the 2321 sessions started only 2151 were accepted and it was typically the third and longest test, OCL that was abandoned or failed integrity. Associated with each participant's data were comments they provided about their test session, as well as details on caffeinated drinks consumed including timing prior to test, size and type of beverage. Each result from a cognitive test was time stamped to the minute which allowed a look-up in the 4.3 million lines of the model-determined effectiveness level at the same time per Table 12.

Zulu Time	Local Time	Actigraphy	In Bed (Up=0 Bed=1 Bad=2)	Sleep Wake (Sleep=0 Awake=1)	Mental Effectiveness
13/10/2011 20:47	13/10/2011 13:47	347	0	1	82.63
13/10/2011 20:48	13/10/2011 13:48	212	0	1	82.55
13/10/2011 20:49	13/10/2011 13:49	117	0	1	82.46
13/10/2011 20:50	13/10/2011 13:50	198	0	1	82.38
13/10/2011 20:51	13/10/2011 13:51	63	0	1	82.29
13/10/2011 20:52	13/10/2011 13:52	195	0	1	82.21
13/10/2011 20:53	13/10/2011 13:53	203	0	1	82.12
13/10/2011 20:54	13/10/2011 13:54	200	0	1	82.04

Table 12 - Example of data derived from each participant's effectiveness profile.

Extraction of data for all tests proceeded in this way to create the file used for comparative analysis between the modelled effectiveness and the surrogate effectiveness measures from the cognitive test results per Appendix 11. Data was segregated into different categories for comparison including shift workers and non-shift workers, males and females; shift workers by gender per Table 13.

Ranking	Category	Number in Category	Avg. Sleep per night (hours)
#1	Male Shift Workers	42	6.54
#2	Males	54	6.57
#3	Male Non-Shift Workers	12	6.70
#4	Shift Workers	68	6.90
#5	ALL	100	6.96
#6	Non-Shift Workers	32	7.07
#7	Female Non-Shift Workers	20	7.30
#8	Females	46	7.40
#9	Female Shift Workers	26	7.48

Table 13 - Summary of Sleep Results by Category

The 68 shift workers averaged 6.9 ± 0.95 hours of sleep per night and non-shift workers averaged 7.07 ± 0.86 hours sleep per night. Table 14 summarizes statistics for the group based on gender.

Variable	54 Males			46 Females			ALL 100		
	Avg	**Std Dev**	**Range**	**Avg**	**Std Dev**	**Range**	**Avg**	**Std Dev**	**Range**
Age (years)	43	10	24 – 60	37	11	23 - 63	40	11	23 - 63
Caffeine (mg)	61	94	0 - 800	54	91	0 - 1125	58	92	0 - 1125
Mental Effectiveness (%)	78	12	36- 100	85	11	50.7 - 100	82	12	36- 100
Avg Sleep Full Study (hours)	6.6	0.9	5.1 – 9.7	7.4	0.8	5.9 – 9.2	7.0	0.9	5.1 – 9.7
Awake Before Test (min)	518	340	5 – 2092	487	321	5 - 1716	504	332	5 - 2092
Sleep Before Test (min)	372	102	0 – 835	436	98	100 - 810	401	105	0 - 835
IDN Response	502	88	330 – 1054	490	111	333 - 1370	497	99	330 - 1370
DET Response	315	86	192 – 1834	323	95	200 - 1139	319	90	192 - 1834
OCL Response	862	186	404 - 1877	863	248	382 - 2727	863	220	382 - 2727

Table 14 - Summary Statistics by Gender

The 54 males were slightly older than the females averaging 43 years ± 10 years versus 37 years ± 11 years. The males averaged 6.6 ± 0.9 hours sleep per night and females averaged 7.4 ± 0.8 hours of sleep per night over the full study. The average mental effectiveness of the workers was 82 percent ± 12. Mental effectiveness ranged from 35.8 percent to 100 percent at time of testing. Females recorded higher mental effectiveness levels during testing than males at 85 percent ± 11 versus 78 percent ± 12. Of the 2151 valid tests there were 21 that recorded with a companion effectiveness level of 50 percent or less.

Caffeine was a commonly used stimulant amongst the workers participating. Of 100 participants, only 19 did not record any consumption of caffeinated drinks near their testing time. However, the remainder recorded a total consumption of 120,815 mg of caffeine during the period of this study, an average of 1492 mg per drinker. Coffee was the beverage of choice contributing over 95% of the caffeine consumed and estimated to be in the system at time of test followed by tea at 3%, colas at 1.9% and chocolate at

0.1%. At the time of testing, participants averaged 58 mg of caffeine in their system with males' content slightly more than females' at 61 versus 54 mg.

Unlike the results of Chapter 7 where hours slept and hours awake were both estimated by participants, the hours slept and hours awake were monitored by the minute in this phase and taken from the time-stamped actigraphy results, not estimates from participants. At the time of testing, participants average time awake was 504 minutes ± 332 minutes. Females were awake slightly less time at 487 minutes ± 321 minutes versus 518 minutes ± 340 minutes for males. The females also collected more sleep before testing at 436 minutes ± 98 minutes versus 372 minutes ± 102 minutes for the males.

Of the three cognitive tests used in this phase of the research, the DET task had the lowest mean response time and was the least variant at 309 ± 87 ms. Males displayed slightly faster response times with the DET task than females, at 318 ± 87 versus 323 ± 95 ms. The OCL task displayed the highest mean response time and was the most varied at 861 ± 223 ms. The IDN task fell between DET and OCL, averaging 486 ± 87 ms. These appear consistent with expectations tied to simple, choice and working memory reaction times. The IDN task was the only one in which males averaged a slower time than females, but were less varied at 505 ± 87 versus 490 ± 111 ms.

Some participants complained that the OCL test was too long. The OCL test which took about 4 minutes to complete compared to about a minute and a half for each of the other two cognitive tests had a higher percentage of incomplete tests. 2.5 percent of the OCL attempts were not completed versus 1 percent for both IDN and DET.

Each of the variables monitored during the research is plotted against alertness levels for each participant and shown in Figure 23 – Figure 29. A trend line has been added to each to look for patterns.

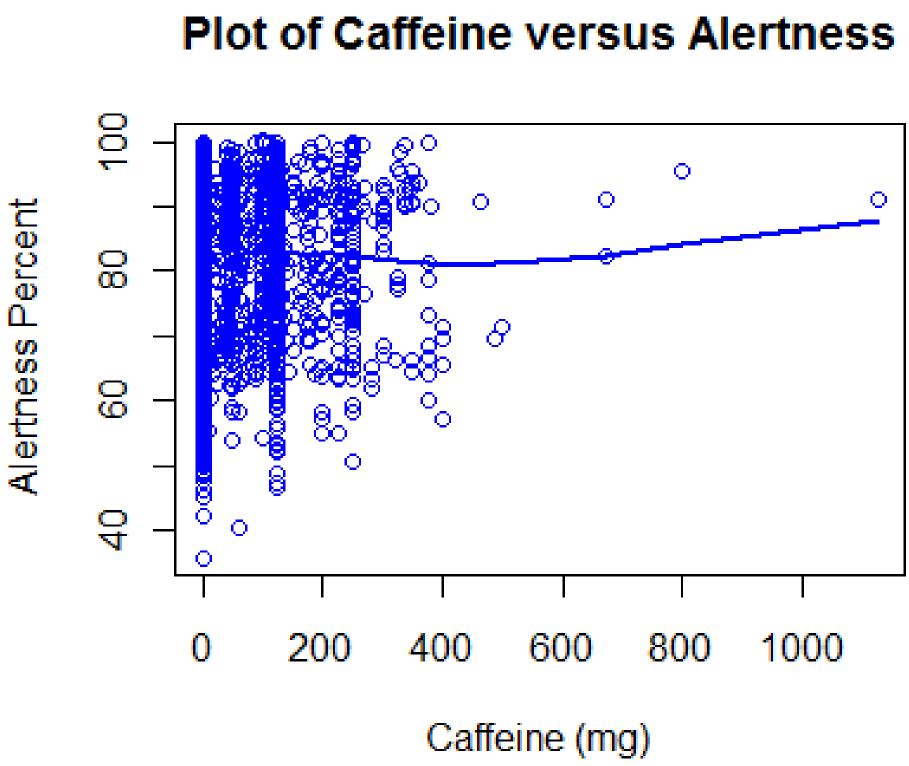

Figure 23 – Caffeine versus Alertness

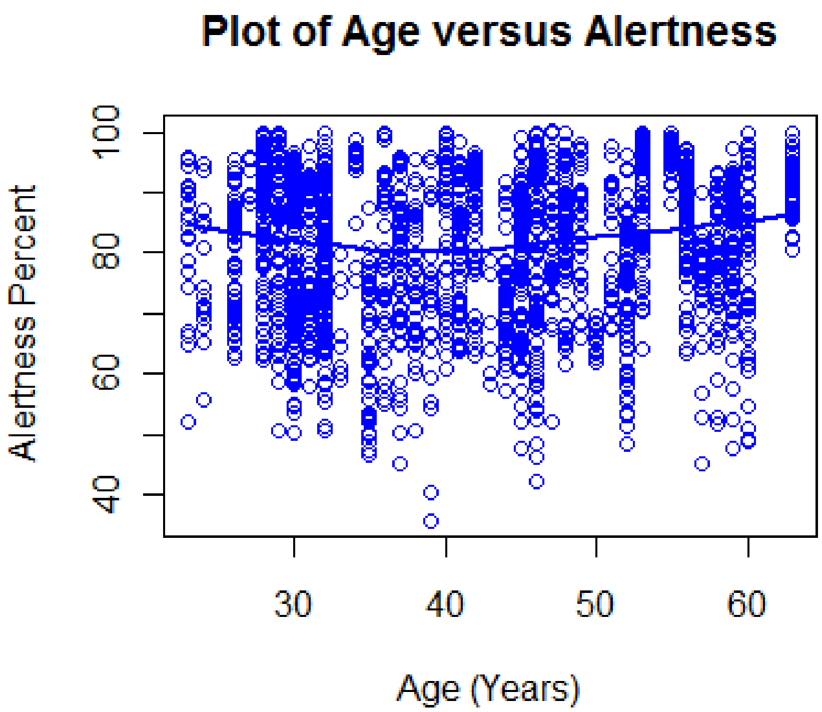

Figure 24 – Age versus Alertness

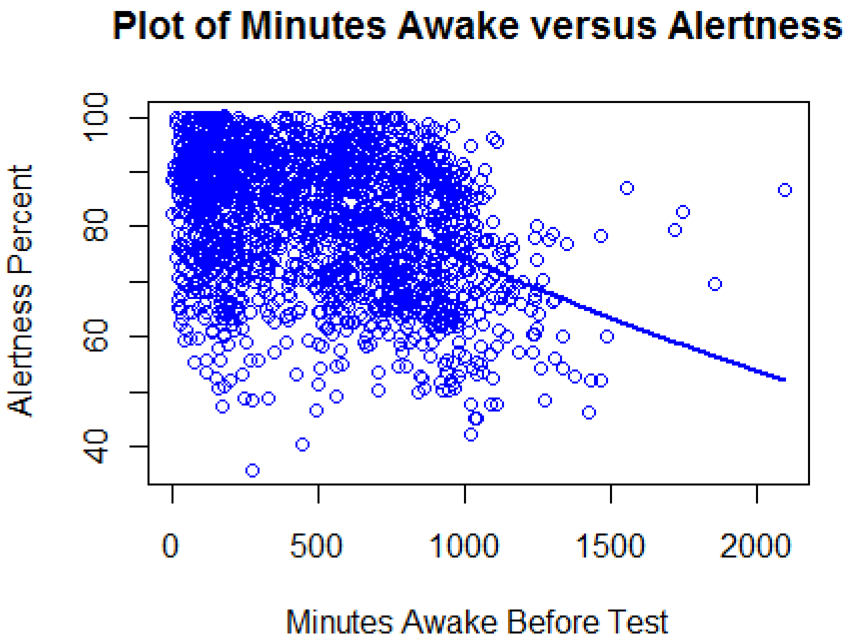

Figure 25 – Minutes Awake versus Alertness

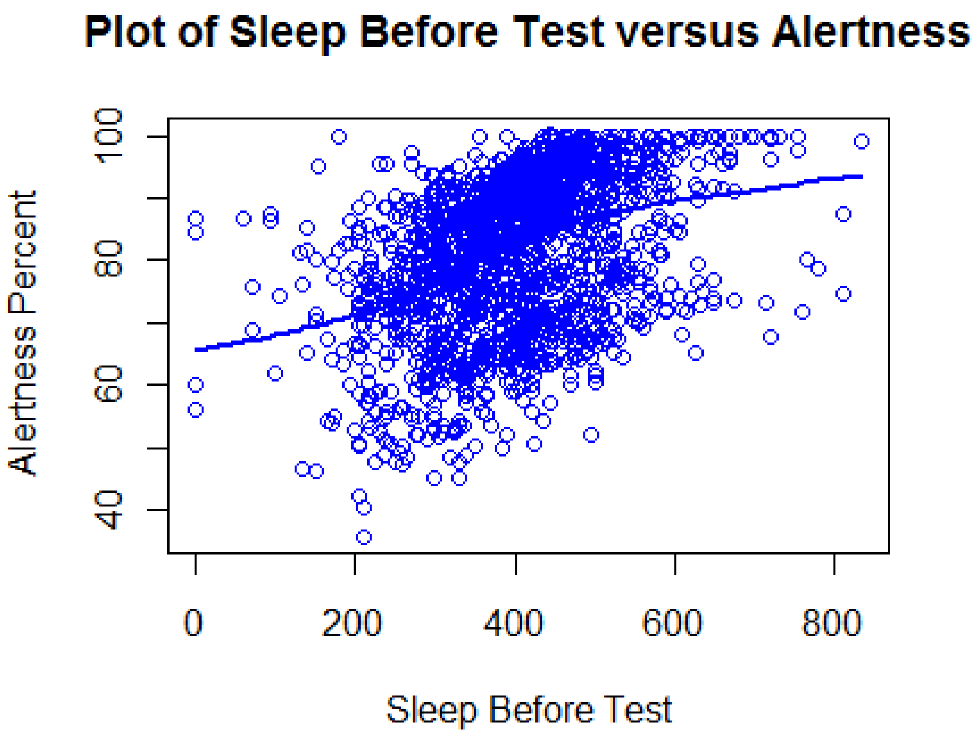

Figure 26 – Sleep before Test versus Alertness

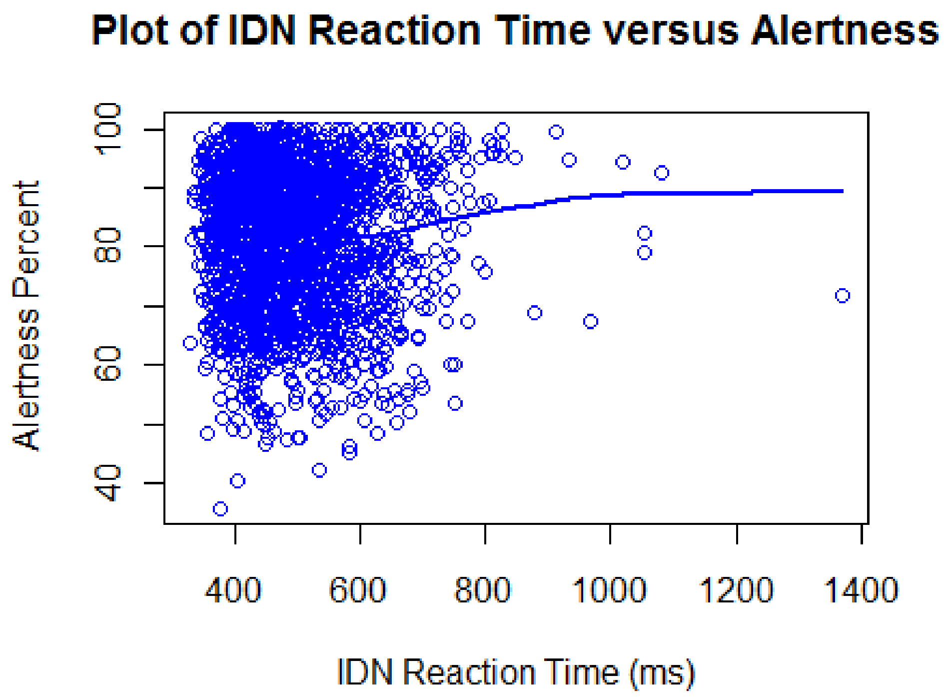

Figure 27 – IDN Reaction Time versus Alertness

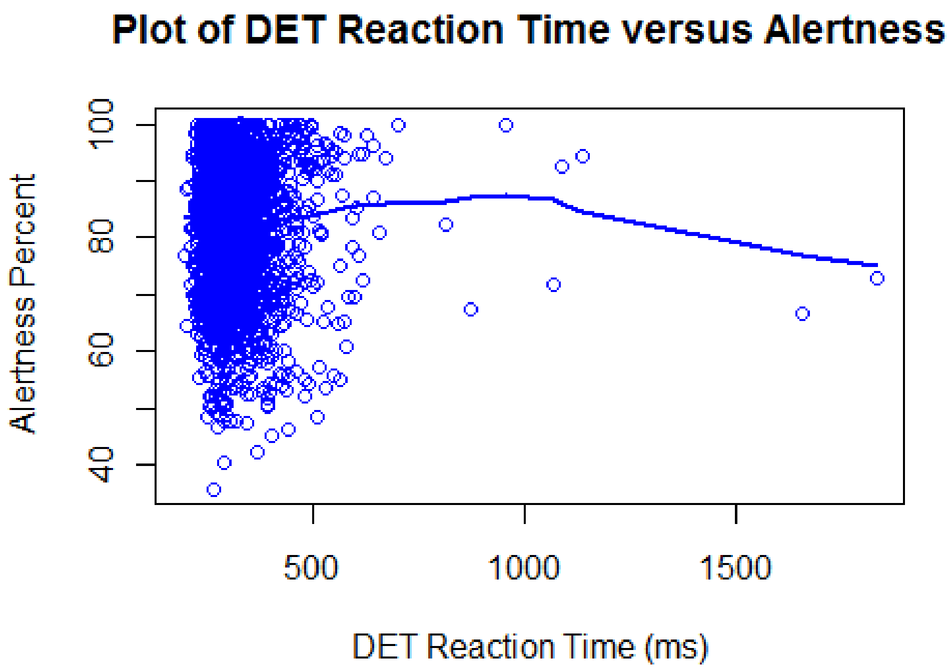

Figure 28 – DET Reaction Time versus Alertness

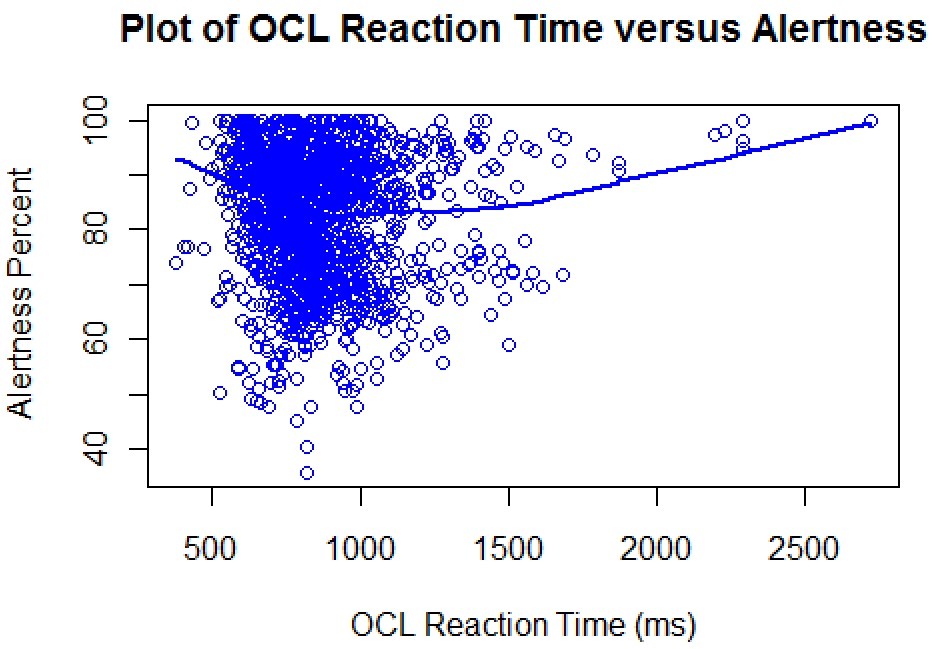

Figure 29 – OCL Response Times versus Alertness

9.5 Analysis and Initial Observations

Figure 23, which is a plot of caffeine versus alertness, shows a slight trend upwards which is consistent with expectations. More caffeine in the system means more alertness. The plot of participant age against alertness displays a broad ‘u-shape’. Alertness was higher amongst the 20 – 30 year age group; drops from 30 – 50 year age group then rises again. There is no obvious explanation for this. Time awake before test shows a negative relationship with alertness. This is as expected; the longer one is awake, the larger is the drop in alertness. Ultimately, this is expected to lead to a loss of consciousness. Sleep before test shows a strong positive relationship with alertness. Again, this is as per expectation; less sleep reduces alertness.

The IDN reaction time plotted against alertness does not exhibit any strong trend. DET reaction time plotted against alertness is also quite flat through the majority of responses. It falls off as it passes through a few outlying points. Finally, the OCL processing times plotted against alertness appear to have a broad ‘u-shape’ similar to age versus alertness. Some of this initial trend is also influenced by a few outlying points. Overall, the grouping of the results is apparent and the range reported in Table 14 is visible. All results exhibit a small number of outlying results and the variability of the data is quite apparent.

9.5.1 Linear Mixed Effects Model

The data from the actigraphy and cognitive testing was assembled into a single file, per Appendix 11, for analysis of correlation and fit. The purpose of the analysis was to understand whether any significant relationship existed between the results of the cognitive tests and the estimated performance effectiveness results from the SAFTE model. The challenge is to properly apportion a representative amount of the extracted results to the final outcome, when they have come from a smallish group of 100 with repeated measurements and with each contributing different numbers of measurements. The sample size, while approaching sizes large enough for classification with general statistical analysis, was analysed with linear mixed effects models as reviewed in Chapter 7 due to the expected correlation of results within each subject from the repeated testing each participant undertook.

To assess the effects as well as assess the other variables' influence, linear mixed effects models (Faraway, 2006) was again used per Chapter 7. Fixed effects are the unknown constants that are estimated parameters to estimate the effect of covariants such as the test results, sleep, and caffeine. The random effect is a variable with estimated parameters that describes the distribution of the random effect (Equation 2).

$$y_{ij} \sim X\beta + Zb + \varepsilon_{ij} \quad \textbf{Equation 2}$$

In this case, the dependent variable 'y' (estimated effectiveness results) is shown to be a linear function of the multivariate measurements and estimated parameters where y_{ij} are the estimated effectiveness results estimated for individual *i* at time *j*. X is the matrix of variables that represents the systemic or fixed effects, and β is its coefficient vector. Z is a matrix of random effects variables with b the random effects. 'ε_{ij}' is the error component.

Therefore the generalized model is:

$$\mathbf{y_{ij} = \alpha + \beta_i + \zeta_{ij} + \varepsilon_{ij}}$$ **Equation 3**

Making up the data, were 2151 results from 100 participants. Thus, $1 \leq i \leq 100$ and $1 \leq j \leq n_i$ with $\sum_{1 \leq i \leq 100} n_i = 2151$, y_{ij} represents the j^{th} observation of effectiveness from the i^{th} subject. The term ζ_{ij} represents the fixed effects terms including the effects of caffeine, time awake before testing, sleep obtained before testing, age and gender.

The random effects term is assumed to follow a normal distribution with a mean that reflects the overall aggregated baseline response that allows the dependence between measurements from the same participants to be acknowledged. The variance on the random term reflects the heterogeneity in responses between subjects and allows separate intercept terms to be fitted for each participant. Any factors not measured which may have contributed to individual effectiveness levels will be contained within the random term, as will any measurement error associated with the measured values.

The analysis was implemented using the lme4 package within the R statistical package (R Development Core Team, 2010) per Appendix 12.
Several linear mixed effects (LME) models were derived to help analyse the relationship between the cognitive test results and the mental effectiveness output from the fatigue-model with actigraph measurements as input, whilst accounting for normal workplace changes in alertness (caffeine and other factors). Table 15 illustrates the models that were initially analysed, comprised of variations of the variables for which data was collected during the research. Entirely fixed effects models were also looked at to compare to the mixed effects results and to verify that the linear mixed effects approach was proper.

Model	Response Variable	Random Effects	Fixed Effects
A	Effectiveness	Subject	Caffeine+Awake+SBT+Age+Gender*
B	Effectiveness	Subject	Caffeine+Awake+SBT+Age+Gender+IDN
C	Effectiveness	Subject	Caffeine+Awake+SBT+Age+Gender+DET
D	Effectiveness	Subject	Caffeine+Awake+SBT+Age+Gender+OCL
E	Effectiveness	Subject	Caffeine+Awake+SBT+Age+Gender+ IDN+DET+OCL
F	Effectiveness	Subject	Caffeine+Awake+SBT+Age+Gender+IDN accuracy
G	Effectiveness	Subject	Caffeine+Awake+SBT+Age+Gender+DET accuracy
H	Effectiveness	Subject	Caffeine+Awake+SBT+Age+Gender+OCL accuracy

Table 15 - Initial Models Analysed with Linear Mixed Effects modelling

*Awake = Time awake before taking the cognitive test
*SBT = Sleep before taking the cognitive test

Each model had a common response variable 'Effectiveness' of the subject at the point of taking the cognitive test as derived from the actigraph-fed fatigue model. Common independent variables of interest, including the subject, caffeine content when testing, time awake before testing, amount of sleep prior to testing, age and gender, were included in the model to help explain the measured 'effectiveness' results. Determining whether there was any significant relationship between the results from the cognitive tests for DET, IDN and OCL and the sleep-based fatigue modelled result was of primary interest. Both the processing speeds as well as the accuracy measurements from the tests were considered when analysing results.

Each of the common variables was added as a fixed effect, with 'subject' being the random effect. Each subject gave multiple data points from tests but effectiveness results from each subject are expected to be highly correlated compared to between subject effectiveness. Therefore 'subject' was added as a random variable with this anticipated collinearity of subject effectiveness results. For the gender term, female was the base gender such that the intercept term covers females and an extra variable was added for males. The processing speed translated into a reaction time in milliseconds (ms) and accuracy was the total number of errors in responses. The data also underwent integrity checks built into the CogState system, explained in Chapter 7. To ensure reasonable efforts were made and that the data fit expected norms, integrity checks included that DET accuracy be 90% or greater; IDN accuracy be 80% or greater; and OCL accuracy be 53% or greater.

9.6 Analysis of Models

The three cognitive test results were analysed to understand the influence of all factors on the effectiveness results as shown in Table 16. The effect of the cognitive test on the outcomes of 3 tests, DET, OCL, and IDN is shown.

ANOVA Results					
Test	**Model**	**Log-Likelihood**	**Chisq**	**df**	**Pr(>Chisq)**
DET	Subject level random effect	-7846.3			
	Additional effect of DET	-7765.6	161.39	1	<2.2e-16***
	Additional Effect of Caffeine + Sleep + Hours awake + Age+ Gender (no DET)	-7209.7	1111.73	4	<2.2e-16***
	Additional Effect of Caffeine + Sleep + Hours awake + Age+ Gender + DET	-7138.4	142.71	1	<2.2e-16***
	Additional effect + OCL+IDN	-5111.4	4053.96	2	<2.2e-16***
OCL	Subject level random effect	-7846.3			
	Additional effect of OCL	-5574.8	-5543	1	<2.2e-16***
	Additional Effect of Caffeine + Sleep + Hours awake+Age+Gender (no OCL)	-7209.7	0	4	1
	Additional Effect of Caffeine + Sleep + Hours awake + Age + Gender + OCL	-5193.2	4033.05	1	<2.2e-16***
	Additional effect + DET+IDN	-5111.4	163.63	2	<2.2e-16***
IDN	Subject level random effect	-7846.3			
	Effect of IDN	-7774.6	143.46	1	<2.2e-16***
	Additional Effect of Caffeine + Sleep + Hours awake + Age + Gender	-7209.7	1129.7	4	<2.2e-16***
	Additional Effect of Caffeine + Sleep + Hours awake + Age + Gender + IDN	-7148.8	121.9	1	<2.2e-16***
	Additional effect + DET+OCL	-5111.4	4074.8	2	<2.2e-16***
Significance codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1					

Table 16 - Correlation of Cognitive test on estimated Effectiveness

There are five levels displayed for each outcome that build on a baseline model showing (i) subject effect which has a random intercept term for the participant; (ii) the addition of the effect of a test result, and; (iii) the additional effect from caffeine, hours slept prior to test, hours awake before test, participant age and participant gender; (iv) the combined effect of (ii) and (iii); (v) the effect of adding the results from the other two tests. In all cases, the results are considered significant at the 5% significance level which from the table is illustrated by an (*). All but one result in Table 16 far exceed this and are highly significant at the 2.2e-16 level as shown in the Pr(>Chisq) column.

The results shown illustrate that adding the effects of any of the three cognitive tests makes a statistically significant difference to the model. When modelling to try and explain the ‘effectiveness’, OCL would be the strongest of the three tests followed by DET then IDN based on the ‘Chisq’ values shown (Model D followed by C followed by B). Additionally the influence of caffeine, hours of sleep before the test, hours awake before the test, age and gender are shown. Hours of sleep before the test showed most effect for all tasks meaning that the less sleep one had before the test, the worse was their result, classified as ‘effectiveness’. This was followed by time awake before the task was done. Caffeine had the least effect of the three as shown in the results of Appendix 13.

9.7 Accuracy Results

Table 17 shows the ANOVA results with the influence of accuracy from each test included. Each cognitive test maintained a count of the number of errors made, as well as the processing time which was the measurement used for processing accuracy. As a factor, accuracy was a much weaker indicator of performance than processing time for each test. However, it was a significant factor in each case; weakest for the IDN test and strongest for OCL test per the ‘Pr(>Chisq)‘ values of Table 17.

ANOVA Results with Accuracy Data					
Test	Model	Log-Likelihood	Chisq	df	Pr(>Chisq)
DET	Base = Effect of Caffeine + Sleep + Hours awake + Age+ Gender	-7209.7			
	Plus effect of DET	-7138.4	142.7	1	< 2.2e-16 ***
	Plus effect of DET_err	-7134.3	8.2	1	0.004274**
OCL	Base = Effect of Caffeine + Sleep + Hours awake + Age + Gender	-7209.7			
	Plus effect of OCL	-5193.2	4033	1	< 2.2e-16 ***
	Plus effect of OCL_err	-5177.1	32.3	1	<1.33e-08***
IDN	Base = Effect of Caffeine + Sleep + Hours awake + Age + Gender	-7209.7			
	Plus effect of IDN	-7148.8	121.9	1	< 2e-16 ***
	Effect of IDN_err	-7147.2	3.2	1	0.0724*
Significance codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1					

Table 17 - ANOVA Results from Accuracy Measurements

Thus far, results indicate that several parameters are significant in explaining the ‘effectiveness’ results. Analysis now will focus on defining a good model based on these parameters.

9.8 Model Selection

The statistical significance relative to the results shows that the cognitive screening tests match with changes in predicted individual mental effectiveness. The cognitive tests displayed a strong correlation to the effectiveness determined by the SAFTE model in the presence of confounding elements, some of which were identified and built into the assessment and some which were not and make up the error profile.

Selecting which model best describes the measured results of ‘effectiveness’ requires additional consideration to simply getting a predicted result close to the actual measured value. If all influences on one’s mental effectiveness were known and measured, it might allow a perfect fit but is not practical or the best approach. Two opposing factors require consideration; simplicity of model with broader applicability or scope versus more parameters to explain the specific results. Alternately stated, this leads to consideration of adding parameters to minimize residual sum of squares or largest maximum likelihood versus removal of unnecessary parameters to simplify the model. There are a couple of similar techniques and tools to assist with this.

One approach introduced by Akaike in 1974 is known as the Akaike Information Criterion (AIC): AIC = $-2\ln(f) + 2p$

f represents the maximum log likelihood and p represents the number of parameters in the model. This technique seeks a balance to adding parameters for maximum log likelihood by considering the strength of what they add. What is sought is a minimum AIC value. In similar fashion, the Bayesian Information Criterion (BIC) also penalises for adding parameters that have diminishing value. Additional parameters have to greatly improve the likelihood of the model to reduce these values. The penalty term is larger with BIC than AIC and shown as: BIC = $-2\ln(L) + k\ln(n)$ where n is the number of data points, k is the number of free parameters to be estimated and L is the maximized value of the likelihood function estimated for the model (Cavanaugh, 2009). Results of this analysis are shown in Table 18.

Cognitive Test	Model	AIC	BIC	Log-Likelihood	Dev.	REMLdev
All other parameters	**A**	14469	14514	-7226	14419	14453
A + IDN	**B**	14358	14409	-7170	14298	14340
A + DET	**C**	14338	14388	-7160	14277	14320
A + OCL	**D**	10446	10494	-5214	10386	10428
A + IDN + DET+ OCL	**E**	10304	10362	-5141	10223	10282
A + IDN_err	**F**	14370	14420	-7176	14314	14352
A + DET_err	**G**	14407	14458	-7195	14351	14389
A + OCL_err	**H**	10421	10468	-5201	10366	10403

Table 18 – AIC & BIC Values for Model fit

The results shown in Table 18 suggest that the best performing models are going to be or are going to be derived from elements of D, E and/or H.

Four new models were reviewed, combining accuracy and processing speed responses. If adding another parameter from the same test strengthens a particular model, compared to combined test results, it would be considered a better model as it reduces the time of test taking and handling of additional parameters from other tests. Table 19 shows three new model results.

Cognitive Test	Model	AIC	BIC	Log-Likelihood	Dev.	REMLdev
Base +IDN + IDN_err	I	14361	14417	-7171	14294	14341
Base + DET + DET_err	J	14336	14391	-7158	14269	14316
Base + OCL+ OCL_err	K	10421	10474	-5201	10354	10401
Base + OCL+ OCL_err+DET+IDN	L	10279	10342	-5128	10191	10255

Table 19 – Results from combined processing time and accuracy

The results suggest the fourth model just based on AIC/BIC scores, is a better model. All lowest BIC/AIC scores include OCL. To complete the review of the models additional analysis of variance was conducted to illustrate the effects of each parameter. The results are shown in decreasing significance based on the t values shown in Table 20.

Model	AIC	BIC	Log-Likelihood	Dev.	REML dev	Var.	Std. Dev.	Est	Std. Error	t value
Effectiveness ~ All Parameters	10279	10342	-5128	10191	10255					
Random effects Subject						45.21	6.8			
Random effects Resid.						70.88	8.42			
Fixed effects intercept								84.37	4.04	20.86
Fixed effects Awake								-0.01	0.00	-12.6
Fixed effects Sleep								0.02	0.00	9.6
Fixed effects OCL_err								-0.19	0.03	-5.65
Fixed effects OCL								-0.01	0.00	-3.98
Fixed effects Gender								-2.84	1.81	-1.57
Fixed effects IDN								0.01	0.00	1.18
Fixed effects DET								0.01	0.01	0.94
Fixed effects Caffeine								-0.00	0.00	-0.39
Fixed effects Age								-0.00	0.08	-0.01

Table 20 – Model Coefficients

Time awake before the test showed most significant effect for all tasks confirming that the longer one was awake prior to testing the lower their mental effectiveness. From Table 20, an additional minute awake reduced mental effectiveness 0.01%. Similarly, the more one slept, the better their effectiveness. Each additional minute sleeping improved effectiveness 0.02%. This was followed by errors made on the OCL task followed by time to complete the OCL task. Each additional error reduced effectiveness almost 0.2% and each additional millisecond taken to complete the task reduced effectiveness by 0.01%. As females were the base gender, the negative value for gender suggests less mental effectiveness for males but it is not strongly significant. Based on low t-value scores for this model, Age had almost no effect followed by Caffeine, the DET results and then the IDN results. All however, contributed to minimizing AIC and BIC values. For a final run, these latter five were dropped from the model and the new analysis of parameters was produced per Table 21.

Model	AIC	BIC	Log-Likelihood	Dev.	REMLdev	Var.	Std. Dev.	Est	Std. Error	t value
Effectiveness ~ Most significant parameters	10541	10578	-5264	10490	10527					
Random effects Subject						46.46	6.82			
Random effects Resid.						70.85	8.42			
Fixed effects intercept								85.67	2.19	39.03
Fixed effects Awake								-0.01	0.00	-13.12
Fixed effects Sleep								0.02	0.00	10.24
Fixed effects OCL_err								-0.17	0.03	-5.57
Fixed effects OCL								-0.01	0.00	-3.60

Table 21 – Final Model Coefficients

The results, with coefficients shown in Table 21, support the findings in strength of models. The influences of Table 20 are unchanged. The negative value for effects of time awake (-0.01) suggest that for every minute people are awake, their mental effectiveness drops by 0.01%. However, for every minute of sleep obtained, individuals' effectiveness increased by 0.02%. Finally, on the principal measurements, OCL_err now has slightly less influence indicating a 0.17% drop in effectiveness for each additional error and OCL indicates a drop in effectiveness of 0.01% for each additional millisecond of response time. With model coefficients as shown in Table 21, the findings suggest the following simple model would be an effective and significant one for estimating individual performance effectiveness:

Effectiveness ~ Subject + OCL + OCL_err + Sleep + Awake **Equation 4**

Two other aspects were reviewed for consideration of model fit. When OCL is added to any model, the significance of the other parameters changed. Of particular note was gender and age, both of which drop substantially. A possible explanation for this is that whatever influence age and gender have on the results were somehow accounted for in the OCL task. To further check this possibility, a final model was run with interaction parameters to see their significance. Table 22 shows findings with additional interaction terms of OCL and age as well as OCL and gender and one additional term to test interaction between caffeine and sleep.

	Estimate	Std. Error	t value
(Intercept)	106.40	6.82	15.6
Caffeine	0.00	0.01	0.03
Awake	0.01	0.00	-12.62
Sleep	0.02	0.00	8.75
Age	-0.04	0.16	-2.82
OCL	-0.03	0.01	-4.18
Gender	-3.07	3.76	-0.82
OCL_err	-0.18	0.03	-5.75
Age:OCL	0.00	0.00	3.39
OCL:Gender	0.00	0.00	0.13
Caffeine:Sleep	0.00	0.00	-0.12

Table 22 – Interaction effects

The results show an apparent interaction between OCL and Age but the magnitude was minimal as shown at 0 (to 3 decimal places). Gender showed no effect and was not significant with OCL and finally, there was no apparent interaction between caffeine use and sleep. As a final check, OCL was plotted against age in Figure 30. It illustrates another u-shape trend observed on the original data of section 9.4.

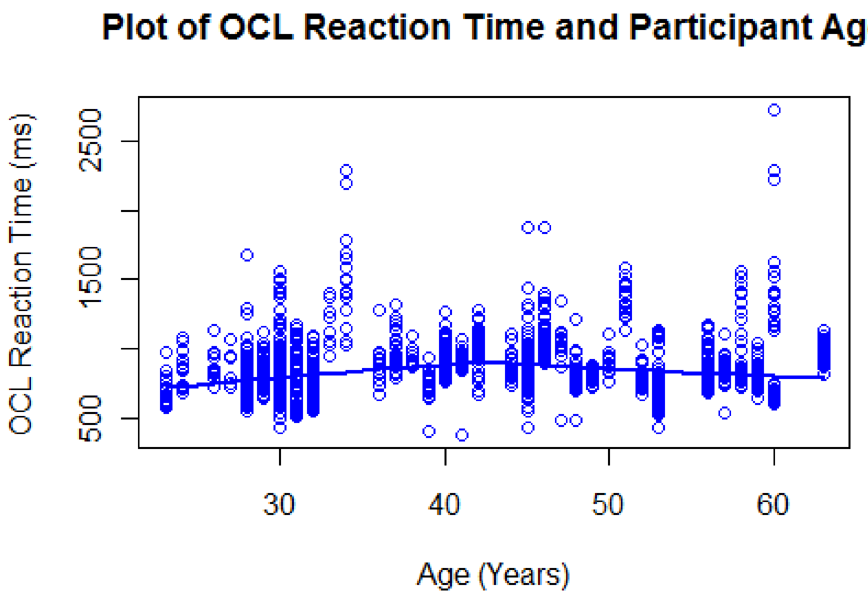

Figure 30 – OCL Reaction Time versus Participant Age

There is no explanation offered for the apparent u-shape trend. Finally, the models were redone, after removing the subject random effects, to create linear all-fixed-effects models. Table 23 shows that in all cases tested, the AIC/BIC values were decreased by adding the random subject effect thus validating the LME approach.

Linear Model Followed by Linear Mixed Effects Model	**AIC**	**BIC**
lm(Alertness ~ IDN, data=rawdata)	16374	16392
lmer(Alertness ~ (1|Subject.ID) + IDN, data=rawdata)	15565	15588
lm(Alertness ~ IDN+Awake+Sleep, data=rawdata)	15052	15080
lmer(Alertness ~ (1|Subject.ID) + IDN+Awake+Sleep, data=rawdata)	14491	14524
lm(Alertness ~ IDN+Awake+Sleep+Caffeine+OCL+DET, data=rawdata)	10633	10675
lmer(Alertness ~ (1|Subject.ID) + IDN+Awake+Sleep+Caffeine+OCL+DET, data=rawdata)	10302	10349
lm(Alertness ~ OCL+OCL_err, data=rawdata)	11643	11664
lmer(Alertness ~ (1|Subject.ID) + OCL+OCL_err, data=rawdata)	11138	11164
lm(Alertness ~ IDN+OCL+DET, data=rawdata)	11456	11482
lmer(Alertness ~ (1|Subject.ID) + IDN+OCL+DET, data=rawdata)	11005	11037
lm(Alertness ~ DET+Awake+Sleep+Caffeine+Age+Gender, data=rawdata)	14816	14860
lmer(Alertness ~ (1|Subject.ID) + DET+Awake+Sleep+Caffeine+Age+Gender, data=rawdata)	14338	14388

Table 23 – Paired Models with and without random effects.

9.9 Goodness of Fit

As a final review of model results, the model estimates of effectiveness were compared to actual measurements. One model was put forward as a best fit and shows some results with actual data shown in Table 24.

OCL Model						
	Estimate	Worker B2	Worker D4	Worker B22	Worker C4	Worker D9
Intercept	85.67					
Awake	-0.01	97	864	297	540	945
Sleep	0.026	425	410	260	402	545
OCL_err	-0.17	22	29	21	17	38
OCL	-0.006	845	991	823	588	628
Effectiveness		86.94	76.81	80.95	84.30	80.16
Actual		92.65	67.56	75.53	81	77.9
Error		6.2%	13.7%	7.2%	4.1%	2.9%

Table 24 – Estimated Effectiveness from OCL Model

As an alternative, the simplest model, the fastest test to take and also statistically significant would be based on the DET task with the same actual data as Table 24 is shown in Table 25.

Simplest Model DET						
	Estimate	Worker B2	Worker D4	Worker B22	Worker C4	Worker D9
Intercept	76.21					
Awake	-0.01	97	864	297	540	945
Sleep	0.026	425	410	260	402	545
DET	-0.002	305	279	275	281	246
Effectiveness		85.71	77.70	79.48	80.73	80.47
Actual		92.65	67.56	75.53	81	77.9
Error		7.5%	15.0%	5.2%	0.3%	3.3%

Table 25 – Estimated Effectiveness from Simplest Model DET

The scale shown associated with the effectiveness estimates from Fatigue Science is based on the Blood Alcohol Concentration (BAC) scale shown in Table 26.

Effectiveness Percent	BAC
100	0%
90	
80	
70	0.08%
60	
50	0.11%

Table 26 - BAC Equivalents

A quick comparison of the estimated results from the two models shows some similarity in results but workers 'B2' and 'D4' were off as fits with both models. In both cases, the concerning error is associated with declarations on mental effectiveness below 70%. Both models carried most error for individuals in the 'Red' zone where more certainty would be needed, as these are levels of mental effectiveness where policies might suggest some action should be taken. Such error could lead to improper declarations about the impairment and therefore lead to improper actions.

To address this, the modelling was re-done to fit to data with effectiveness levels below 80%. This resulted in a best fit model as per equation 4 but with different coefficients per Table 27.

Effectiveness ~ Subject + OCL + OCL_err + Sleep + Awake **Equation 5**

Model	AIC	BIC	Log-Likelihood	Dev.	REML dev	Var.	Std. Dev.	Est	Std. Error	t value
Effectiveness ~ Most significant parameters	3781	3811	-1883	3729	3767					
Random effects Subject						7.01	2.65			
Random effects Resid.						44.14	6.64			
Fixed effects intercept								69.57	2.22	31.35
Fixed effects Awake								-0.005	0.001	-5.25
Fixed effects Sleep								0.014	0.003	5.02
Fixed effects OCL_err								-0.029	0.031	-0.944
Fixed effects OCL								-0.001	0.002	-0.804

Table 27 - Revised Coefficients for Equation 4 from best fit in 'Red Zone'

The resulting equation provides the following comparative estimates for 'red zone' results including for Worker D4 previously estimated.

Simplified Model Fit Best Fit to Red Zone						
	Estimate	Worker D4	Worker B12	Worker B22	Worker B19	Worker C21
Intercept	69.57					
Awake	-0.01	864	951	511	898	1144
Sleep	0.014	410	310	360	385	195
OCL_err	-0.029	29	31	27	25	24
OCL	0.001	991	953	761	690	954
Effectiveness		69.16	67.3	70.5	69.1	64.93
Actual		67.56	61.9	69	66.2	60.2
Error		2.4%	8.70%	2.1%	4.3%	7.9%

Table 28 - Revised Estimates with Red Zone Fit

The resulting estimates are now much tighter on all 'red zone' effectiveness levels or those falling at > 0.08% BAC equivalencies.

9.10 Discussion

The original aim of this research was to determine if there was an effective means of screening for fatigue-impairment and declaring worker fitness-for-duty to reduce risk of accidents and improve productivity in all places of work. The first step of this was verifying that one's mental effectiveness can be measured in real time. The results show that there are several cognitive tests that could be used for correlation to mental effectiveness and that there are multiple models that can be developed which might provide a good result to meet this aim.

The results from the analysis supports previous findings (<u>Chapter 7</u>) that suggested fatigue is very impactful on reaction times and memory and it can be detected in the workplace. A common metric of interest in all three tests was the reaction time in processing the task. However, both IDN and OCL had added dimensions to their processing times. Whereas the DET task was a simple reaction time task, IDN was a choice reaction time task and OCL brought working memory into play. That dimension moved OCL to the front in predicting task effectiveness levels as estimated by the SAFTE model. The results also suggest using both the processing time and memory errors as key parameters. All three tasks showed a level of brevity (< 5 minutes to complete), were easily learned and did demonstrate a level of strong correlation to the SAFTE sleep-based model. The results suggest the OCL task or tests of similar kind would be best candidates to employ as a cognitive test screening tool.

As an individual test, OCL demonstrated most or second most sensitivity for 9 of 13 participants from the first phase of this research and was a clear winner in the second phase in significance relative to the SAFTE results. The best results, based on lowest AIC and BIC scores, came from using all three tests. However consideration must be given to the time required to conduct three tests. A significant time difference exists between OCL and DET or IDN to complete the test. If a single test were given, the results suggest that OCL is the one to give. However, it comes with a small time penalty per test versus the other tasks to complete. On the positive side it provided an additional metric in error count that is significant in predicting effectiveness. This issue is relevant for operational settings as part of a programme conducting tests over time, as more tests could amount to a considerable time investment in testing. There is value in minimizing the time of testing but not at the expense of poor prediction.

Hours awake before testing and hours slept before testing showed strong significance in explaining the results. With this finding on test performance future work should continue to embrace this as part of the study as well. Surprisingly, caffeine, age and gender had no significant effects in some models. With respect to applying findings in an operational setting, this is a good finding. It simplifies the data collection process for applying the model especially for caffeine, which is difficult to measure and keep track of consumption.

The results further suggest that cognitive tests must be considered possible tools and surrogate measures for actigraph-fed performance models for predicting mental effectiveness. The strong matching of estimated effectiveness, whilst accounting for workplace confounding elements, suggests further development could make them very useful tools in a Fatigue Risk Management programme.

The analysis showed there were no systematic failures associated with any kind of learning effects from repeated testing, just as was experienced in the first phase of this research with a smaller population. Thus, there is additional support to the preliminary findings (Chapter 7) from this research for stable results with minimal practice effects in a busy work place, even when operating around the clock. These tests were all conducted on a voluntary basis. As part of a Fatigue Management Programme, the tests could be mandated to drive higher response rates. This would have to be tempered with the effects mandating may have. If individuals are hostile towards testing as opposed to willing to support the management of safety, results could be tainted. However, at least with the Cogstate database, boundaries have been established to determine if a reasonable effort was made on a task. With these results significantly correlating to mental effectiveness, the surrogate real-time cognitive tests could now be extended into forecasting equivalent Blood Alcohol Concentrations as a direct link to the effectiveness readings.

This study took workers as they presented themselves with all of the individual variability they otherwise brought to the workplace. The results indicate that use of caffeine in the workplace need not be a significant factor in discussions on fatigue and coping mechanisms. A possible explanation for this is that the effect of caffeine is most impactful on sleep quality. Perhaps if it was showing influence it was already accounted for in the measurement for sleep, which included sleep quality, if it was in the system at time of sleeping. However, interaction between caffeine and sleep was checked and there was no significant interaction detected.

As a final point, during the course of this study, the estimated mental effectiveness dipped into the red zone at least once during testing (BAC equivalency > 0.08%) for 80 of 100 workers. This number alone provides strong evidence that workers and organisations are not focused enough on addressing fatigue impairment.

9.11 Summary

This chapter presented the first known findings associated with a correlation between cognitive tests and fatigue modelling from actigraph measurements of sleep, derived from real workers in a workplace setting. The results are conclusive in establishing that there are several parameters used in this study that correlate strongly with the effectiveness results modelled from sleep measurements. These parameters themselves can be used to develop models based on the measurements from cognitive tests. Cognitive tests have the advantage of deployment into work environments in real time and if done as part of a fatigue management strategy for roles carrying most risk, they could be an invaluable tool in improving safety.

CHAPTER 10 – RECOMMENDATIONS, LIMITATIONS AND CONCLUSIONS

10.1 Introduction

The results of this research are the first findings from real workplaces that show that simple cognitive tests can predict fatigue-related impairment with high confidence levels. This comes from high correlation to estimated performance effectiveness levels obtained from continuous measurements with actigraphy and sleep-based modelling. Unlike use of actigraphy which has practical limitations for continuous use in the work place, including challenges to overcome privacy concerns, use of cognitive tests as screening techniques could be limited to use only at work and in conjunction with work schedules. Unlike actigraphs which must be worn during sleep to be effective, the unacceptable intrusion into workers' personal and private space would not be needed for on-going cognitive testing. However, it must be recognized that this is just one missing tool for use in a broad systematic approach, focused on managing fatigue and alertness in a hazardous workplace. This chapter reviews the research findings and outlines the recommended actions to take as next steps and lists the limitations associated with the research.

10.2 Findings

Several objectives were set for this research and each objective was successfully met. Following is a summary.

1. To detect and provide actual objective measurements and assessment of fatigue-impairment to confirm fatigue-impairment exists.

Chapter 9 reported on 100 workers who provided access to their work/rest cycles for a month and produced evidence that fatigue-impairment does exist in the workplace, as well as providing quantitative levels for it. Averaged over a one-month period, workers got less than 7 hours of sleep per night, and operated at just over 80 percent mental effectiveness. As fatigue-impairment has been shown to be a workplace hazard, the more concerning findings were that out of the 100 workers, 80 tested at a BAC equivalency of 0.05% or greater at least once during their test period; or legally drunk if on alcohol. Additionally, Chapter 3 reviewed other findings and evidence of fatigue in the workplace. Evidence was also presented in Chapter 4 to help understand worker risk due to fatigue-related impairment and Chapter 5 presented results of a survey that indicated that 69% of 115 respondents were aware of an occupational accident in the past 12 months due to this impairment.

2. To provide an understanding of the impact of fatigue-impairment on construction workers' risk, safety and productivity.

This research has illustrated that there are numerous cases of accidents being directly caused by fatigue-impairment or where fatigue was cited as a major contributor to the accident (Chapter 3). The literature review of Chapter 3 also established an understanding of what drives fatigue-related impairment as well as the impact a lack of sleep can have on health. This research has further illustrated how fatigue-impairment manifests in the workplace as a new hazard. This is an additional hazard to all the other hazards that contribute to making the construction sector one of the poorest safety performers globally (Chapter 4). As both an active and passive element in the chain of causation to accident, fatigue-impairment itself is properly managed in a Fatigue Risk Management System but there is no evidence this is being done is this sector.

Chapter 5 showed that fatigue-impairment is viewed as the most serious impairment factor in construction but also confirmed it is one which receives little recognition or mitigation putting construction workers' at risk. Without even basic polices and training programs in place to address fatigue when planning work, risk is perpetuated.

3. To provide understanding on issues and impediments associated with dealing with fatigue-impairment in the workplace.

Chapter 5 exposed views from four different continents regarding fatigue as an impairing agent and what is and what is not being done in the construction industry about it. The survey results allow us to conclude not much is being done about fatigue-impairment in the sector even though it is seen as the number one impairment problem. This appears to be tied to erroneous views about the safety and health aspects of fatigue and a belief that it is acceptable to be impaired from fatigue. Our views support that how we became impaired is more important than the level of our impairment. The lack of action regarding fatigue-impairment is primarily tied to this whereas it should be viewed as a workplace hazard and managed. It was concluded that heavy lifting will be required to change attitudes towards this factor as a necessary pre-condition to change.

4. To review and critique means of assessing organisational and individual fatigue.

Chapter 6 reviewed technologies and results from studies which have attempted to measure fatigue or its symptoms. It was concluded that most of the techniques available have drawbacks for application in construction. However, there are some existing technologies that will come into the sector such as camera-based sleep detectors for mobile equipment operators. As a final line of defence where they are suitable, they should be used. However, for general purpose operations, it was concluded simple cognitive tests have shown an ability to detect fatigue and may hold promise for work site applications as they are simple, inexpensive and can detect cognitive decline not just the onset of somnolence. This research has shown that fatigue-impairment is detectable in the workplace with several different types of technology.

5. To test possible fatigue-detection solutions in an operational setting and assist with development to allow introduction into the construction workplace as a tool to screen out fatigue.

Chapter 7 reported on testing applicability of a battery of possible cognitive tests for screening for fatigue-related impairment. A pilot program was completed to test their sensitivity in measuring expected normal changes in individual alertness correlated to the circadian rhythm. When considering possible candidates for detection of fatigue-related impairment, as guiding criteria to select from the battery tested, simplicity of test-taking, brevity and robustness of results were shown to be key factors. To some degree all six tests studied showed a level of brevity (< 5 minutes to complete), were easily learned but did demonstrate different levels of robustness. The validity of measurement found and robustness of inference from the results assist selection for further development.

With the limitations of this study and based on results found, a priority can be placed on providing strong guidance for tasks requiring cognitive alertness. Loss of reaction time is expected to show up first and as a part of fatigue management planning, a priority should be placed on identifying tasks requiring fast reaction times in safety critical activities. These activities may extend beyond the scope of the task itself and may be the shoulder activities of commuting to and from work. The cognitive tests most reliant on reaction time for processing appear to be the best performers for further development.

6. To validate the effectiveness of tested fatigue-detection tools with proven means of assessing fatigue-impairment.

It was concluded from Chapter 8 that there is solid ground for using actigraphs and the Hursh sleep model as a basis for validating cognitive screening results. The selected technology has been subjected to substantial test and validation and is considered a leading model based on its ability to predict performance effectiveness from normal or abnormal sleep.

Chapter 9 reported on a very large study with 100 workers on different shifts who had their actual sleep/wake patterns measured 24-hours per day for a month with actigraphs to feed the Hursh fatigue model. Performance levels were determined minute-by-minute for the entire month and compared to results from random cognitive testing. It was concluded that all the cognitive test results had an extremely high correlation with output from the predicted Hursh model effectiveness levels. It was further shown that many robust models could be developed for use as a tool to assist real-time assessment of worker mental status due to fatigue. It is concluded this now needs to happen.

7. To contribute to improved productivity through the reduction in work-related injury, illness and working days lost by providing a path forward for the industry in dealing with fatigue-impairment.

The results of this research have validated a new exciting opportunity for real-time screening for fatigue-related impairment in the workplace. It has also been validated that the cognitive tests are easily learned by the workers and there are no operational hurdles to deploying them. The results from the CogState statistics showed little failure by participants and most failure occurred due to the research requirement of doing a multiple-test battery as opposed to a single test. Whether cognitive screening will contribute to improved productivity through the reduction in work-related injury, illness and working days lost is not yet known. However, this research has provided a clear path forward for the industry in dealing with fatigue-impairment to allow this. There is also opportunity to complete a renewed focus on safety with consideration of fatigue management.

Additionally, factors other researchers have found as strong influencers on alertness, such as stimulants and hours awake, were addressed and built into these results. This study took workers as they presented themselves in real workplaces with all individual variability they normally brought into the workplace and gave results as surrogates of alertness over the course of their shifts. The results indicate these tests can also be used as measures of performance effectiveness and can be correlated directly to existing scales of Blood Alcohol Concentration levels for broader understanding of worker condition. Estimating performance effectiveness for workers who either appear impaired, have worked long hours in a day or week or are working an unusual shift with very early start times or very late work hours are candidates for reviews.

10.2.1 Hypothesis Findings

The Hypothesis for this research was that "Fatigue-impairment can be identified in the workplace and the associated risks from individuals' fatigue-impairment can be reduced by knowing about the condition and taking appropriate actions". It is accepted.

10.2.2 Other Findings

While it was not the focus of this research, due to extensive monitoring the research also discovered some interesting sleep facts associated with workers. Female workers obtained more sleep than male workers in this sample and less surprising shift workers get less sleep than non-shift workers. 17% of the workers averaged 6 hours or less of sleep per night, which is the group most at risk in any workplace due to fatigue-impairment.

10.3 Recommendations

While extensive effort was put forward as part of this research to determine an accurate and practical model for assessing mental effectiveness related to fatigue, the primary finding is the fact that it can be determined with cognitive tests. The value of this research is in the presentation of a new tool to construction operations globally. Developing a cognitive test based on this research and deploying the model as a screen for fatigue in a fatigue risk management plan should follow. Where roles demand individuals to work shift or during troughs in their circadian cycle or drive from home to start an early shift or after long hours of work or extended periods of poor sleep, screening with a cognitive test along with some other basic information could quickly estimate one's mental effectiveness level correlated to a Blood Alcohol Concentration level, to guide decisions. Within the limitations of this research and based on results found, a recommendation can be put forward that the cognitive tests selected should be built into safety management systems as screens for fatigue impairment in the construction industry based on roles. This includes the shoulder activities of commuting to and from work for all workers, which can be classified as a safety critical activity. Results from this study also had strong workplace relevance in assessing the acceptance and the difficulty in taking the tests. The simplicity and brevity of the tests, especially the Detection (DET) task, allowed them to be easily completed by the participants each work day whilst maintaining regular and varied work functions.

The construction industry should develop and implement a training programme to help create awareness around fatigue in the workplace due to inadequate sleep. Anything which promotes better and adequate sleep is expected to have a positive outcome on worker wellness and performance. Workers need to understand the impact of poor sleep and what they can do in their personal lives to improve their sleep. Workers getting inadequate sleep not only become more risky workers, they can increase risk for co-workers too.

The construction industry should profile and classify its jobs and ensure that those jobs putting workers at highest risk receive a higher priority of mandatory training regarding this subject. Mobile equipment operators, shift workers, night workers and workers putting in long hours at the workplace are examples of workers who should be targeted for training, monitoring and following rules. Schedules for shifts allowing for commute times should be reviewed to ensure the shifts do not start too early or end too late. Shift workers should be a continued target for education and awareness regarding sleep, as more and more research surfaces, linking health issues to poor sleep and shift-work. Shift-work puts the workers at elevated risk of poor sleep and the derived issues from it. Moore-Ede's (2012) fatigue incident data collection programme is something that could be implemented easily as part of incident investigations to help raise awareness on fatigue and start collecting data.

Finally, there has been substantial work done in other sectors on the subject of fatigue countermeasures, which is an area that deserves some consideration for construction. The data collected regarding worker habits when fatigued and their beverage of choice, illustrates that the consumption of drinks with high caffeine levels is prevalent. All studies have shown that while this might assist alertness it cannot be a substitute for proper sleep.

10.4 Limitations

This project was fortunate to have a home base supported by a large number of interested workers for the study. Interest in sleep is strong at the worker grass roots level as it impacts everyone's life for better or worse. Randomness and large samples were a goal but even with strong interest, this study was limited by funding and had to reach out to multiple organizations and sites to find 100 individuals willing to wear an actigraph for a month. The value of more data is immense but it takes considerable effort to obtain it. The profile of personnel participating including age, gender and type of work varied but it cannot be said it is a random representative sample of any larger group. While no statistical inference is made, to the degree that the samples presented are not reflective of another or larger group, more error is expected in the application of the derived models and the validity of claims made associated with application to a different and larger population. The size of the sample population that the funding for this study allowed was not large enough for standard statistical inference and necessitated the use of linear mixed effects models but is conclusive in result. A residual concern would be associated with the full scalability of these results to general populations.

This research has declared a degree of success in being able to correlate a simple real-time test measurement with a comprehensive and validated sleep-measured model. It must be remembered however, that the sleep model also has error and only produces an estimate of individual effectiveness. Both approaches are therefore just estimates and still not the penultimate measure of an individual's state of mind.

Two of the variables found to be significantly tied to the results were the amount of sleep individuals took before testing and time they were awake before testing. Both of these measurements were derived from the actigraphs, which are accurate in measuring sleep-wake cycles and thus in determining precisely the amount of time individuals were awake before testing. Both the measurement of sleep, which factored in an individual's sleep quality, and their time awake were accurate measurements. In a real world modelling without actigraphs, these measurements would be suspect and impact the models with additional error. As sleep in particular, accounting for sleep quality, is poorly estimated, this factor would need to be reconsidered in a working model. Without measurement, any dishonest reporting associated with time awake or time slept prior to test would result in significant model error due to the strength of these factors on predicting mental effectiveness. Perhaps a model should be selected which does not contain these variables.

Finally, while this research was guided by significant other research and expertise associated with the choice of cognitive tests to use, the selection put forward for the final phase of this research was based on sensitivities found in a small population. It is most probable that other tests exist, perhaps with better sensitivities and effectiveness.

10.5 Implementing the Recommendations and Future Work

To take full advantage of the findings of this research, the development and implementation of the resulting model can only be considered a part of a working solution. It provides one missing piece of a strategy for addressing fatigue in the industry. It will require commitment to several other changes including implementing a fatigue risk management system. It gives invaluable protection by monitoring for alertness management per Chapter 4. From Chapter 4, Table 1 is reproduced below.

<table>
<tr><td>Workload- staffing balance</td><td rowspan="3">Sleep Management</td></tr>
<tr><td>Shift scheduling</td></tr>
<tr><td>Employee fatigue training and sleep disorder management</td></tr>
<tr><td>Workplace environment design</td><td rowspan="2">Alertness Management</td></tr>
<tr><td>Fatigue monitoring and alertness for duty</td></tr>
</table>

Table 1 – From Section 4.4

All factors associated with managing fatigue are required to be in place to derive benefit from any tool assisting in monitoring fatigue in the workplace. If an organization does not care about managing worker fatigue for any reason, this tool will not be useful. Similarly, there would need to be a degree of cooperation from participants to make these models work. For those most concerned with their own health and safety, cooperation would not be an issue.

While confidence can now be given that a cognitive test can be developed into a surrogate measurement for fatigue, continued work is required to select and finalize a test for commercial applications. The resulting model would be capable of predicting a mental effectiveness level correlated to a Blood Alcohol Concentration equivalency, for ease of decision making per Table 26. A framework for implementing a Fatigue Management Programme is shown in Table 29.

Fatigue Management Programme Framework		
1. List purpose, principles & scope 2. Management of the Fatigue Safety Programme 3. Define Roles and Responsibilities 4. List relevant Acts, Regulations and Standards covering work; consider policies to address; 5. Identify workplace hazards, assess and list controls; 6. Establish safe boundaries for work;	7. Provide training and awareness on fatigue and use of tool and its value; establish work policies and guidelines tied to fatigue and results of tool; 8. Start fatigue management, investigation and reporting; tools such as Circadian fatigue incident data collection might be considered; 9. Use tool to requirement and maintain results dbase;	10. Audit roles and hours; maintain a corrective action system 11. Apply to contractors and implement; 12. Manage documents and records; 13. Review incidents and audit findings with site management to a schedule; 14. Maintain and grow results from monitoring;

Table 29 – Framework for Fatigue Management Plan

This research validates that more effort is still required in the construction sector for management of fatigue which is summarised in Figure 31.

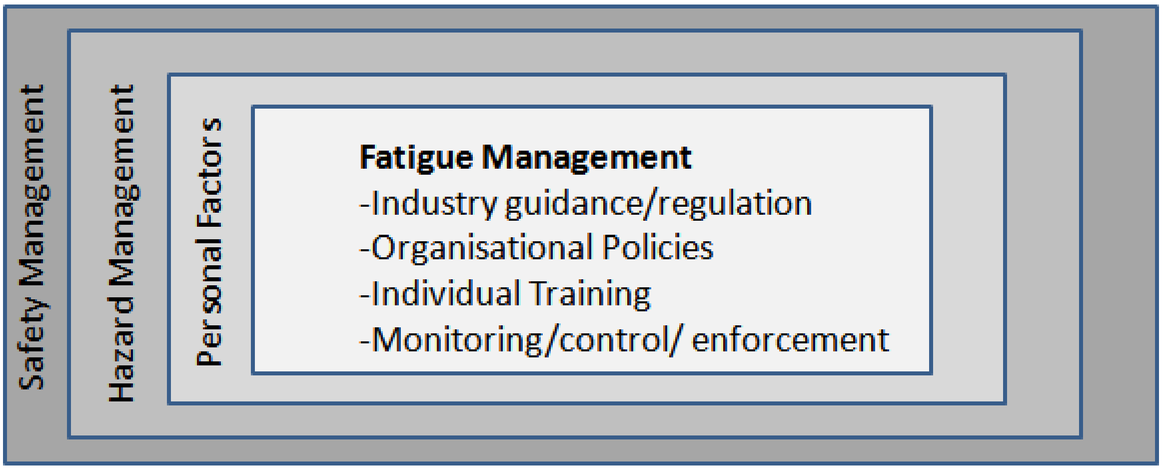

Figure 31 – Support required from multiple levels for Improvement

Required at the highest level:

- Industry guidelines and regulation for fatigue management, based on existing evidence of the impact of fatigue-impairment and successful models from other sectors;
    - Affects resource planning and staffing schedules;
    - Requires industrial standards for change to maintain competitive balances when bidding on jobs;

Required at the sector level:

- Classifications of safety-critical roles;
- Safe work hours for safety-critical roles;
    - Requires discipline, enforcement and consequences for violation;
    - Establish guidelines for setting shift patterns and schedules.
- Awareness programs tied to understanding fatigue

Required at the organizational level:

- A Fatigue Management Programme (Table 29);
    - Use the tools that already exist and develop those missing;

At the individual level:

- Awareness and understanding of the importance of proper sleep as a part of personal health and wellness;
    - Seek medical help when sleep falls outside of expected norms.

There are multiple levels of support required in these recommendations as it would take a coordinated effort between government and industry, organisations and individuals to accomplish this. Continued effort is encouraged to implement fatigue management as part of safety management systems. As all jobs carry some risk, managers may be inclined to simply accept that fact and not work to reduce fatigue in the work place. However, understanding the economic opportunity cost associated with each decrement in performance due to fatigue might lead to a different course of action.

Finally, as much evidence from other research suggests the real issue of increased risk from inadequate sleep is due to sleep disorders, a simple screening mechanism for common disorders would be useful to allow identification of those workers at greatest risk and in greatest need of help. Those problems can be fixed inexpensively.

10.6 Conclusions and Future Research

This research has delivered on all stated aims and objectives and validated a new opportunity for assessing worker fatigue levels in real-time. Key factors were identified as having significant correlation to changes in worker mental effectiveness levels. Techniques and models based on these factors such as the cognitive test OCL are most suited for on-going use. As an emerging alternative technology, cognitive tests, which correlated strongly with the output of the actigraph-fed fatigue models, can now lead to surrogate measures of individual performance effectiveness in the workplace.

The literature review established an understanding of what drives fatigue-related impairment; an impairment that nothing has been done about in the construction sector. Sleep was found to be the only naturally occurring cure for fatigue-impairment making sleep an important topic of discussion for all training programs, and to ensure adequate sleep is allowed for and promoted when planning work. Finding sleep disorders should also be a priority.

These findings extracted from research in excess of four years have revealed that fatigue-related impairment is detectible in real time in the workplace. Final development will allow a new simple tool to be used for integration into safety management programmes to improve safety globally. Without it, and in preparation for it, the industry needs to increase focus on fatigue and its dangers, especially for the most safety-critical roles. Further work, focused on understanding the performance decrement in real dollar terms of fatigue-related impairment, is needed to facilitate development.

Future research should build on these findings and provide a basis for commercial implementation. Any data that could be collected to expand on the first 100 participants would assist in strengthening models and help validate the selection of the cognitive tests chosen in this research for modelling.

REFERENCES

Akerstedt, T. & Folkard, S. 1995. Validation of the S and C components of the three-process model of alertness regulation. *Sleep,* 1**,** p1-6.

Allison, P. D. 1999. *Multiple Regression: A Primer* Thousand Oaks, California, Sage Publications Inc.

Amano. 2011. *Chocolate Does Have Caffeine* [Online]. Available from: http://www.amanochocolate.com/articles/caffeineinchocolate.html. [Accessed 12 August 2012].

Balkin, T. J., Belenky, G. L., Hall, S. W., Kamimori, G. H., Redmond, D. P., Sing, H. C., Thomas, M. L., Thorne, D. R. & Wesensten, N. J. 2002. *System and method for predicting human cognitive performance using data from an actigraph* USA patent application 20020005784. 2002-01-17.

Baumeister, R. F. & Vohs, K. D. 2007. Self-Regulation, Ego Depletion, and Motivation. *Social and Personality Psychology Compass,* 1**,** p14.

BBC NEWS. 2012. The science of sleep. *Science: Human Body & Mind* [Online]. Available from: http://www.bbc.co.uk/science/humanbody/sleep/articles/whatissleep.shtml [Accessed 12 August 2012].

Beaulieu, J. 2005. The Issues of Fatigue and Working Time in the road Transport Sector. Geneva: International Labour Organization (92-2-117720-3).

Belenky, G. 2012. *Curriculum Vitae* [Online]. Spokane: Washington State University. Available from: http://spokane.wsu.edu/researchoutreach/Sleep/belenky.html [Accessed 12 August 2012].

Berrios, G. E. 1990. Feelings of fatigue and psychopathology: a conceptual history. *Comprehensive Psychiatry,* Mar - Apr**,** pp140 -51.

Berson D.M., Dunn F.A. & Takao M. 2002. Phototransduction by retinal ganglion cells that set the circadian clock. *Science,* 295, pp1070-3.

Biosource, 2012. *UBC Practice Leader* [Online]. United BioSphere Corporation. Available from: http://www.unitedbiosource.com/bios/wesnes.aspx [Accessed 12 August 2012].

Boutrel, B. & Koob, G. F. 2003. What Keeps us Awake. *Sleep,* 27.

Bowles-Langley. 2012. *Web Page* [Online]. Available from: http://www.bowles-langley.com/ [Accessed 12 August 2012].

Brace, C., Gibb, A., Pendlebury, M. & Bust, P. 2009. Phase 2 Report: Health and safety in the construction industry: Underlying causes of construction fatal accidents - External research. Secretary of State for Work and Pensions. London: HMSO (978 0 10 176572 5).

Brain, M., Bryant, C. W. & Cunningham, M. 2012. How Caffeine Works. *howstuffworks* [Online]. Available from: http://science.howstuffworks.com/caffeine4.htm [Accessed 12 August 2012].

Brennan, C. 2012. *Are you getting enough sleep?* [Online]. netdoctor. Available from: http://www.netdoctor.co.uk/womenshealth/features/gettingenoughsleep.htm [Accessed 12 August 2012].

Breus, M. J. 2012. *Working a lot; Not Sleeping Enough* [Onine] Available from: http://blogs.webmd.com/sleep-disorders/2012/05/working-a-lot-not-sleeping-enough.html? [Accessed 12 August 2012].

Cauter, E. V., Knutson, K., Leproult, R. & Spiegel, K. 2005. The Impact of Sleep Deprivation on Hormones and Metabolism. *Medscape Neurology.7(1) p* 6.

Cavanaugh, J. E. 2009. Model Selection Lecture VI: The Bayesian Information Criterion. [Online] Department of Biostatistics, University of Iowa. Available from: http://myweb.uiowa.edu/cavaaugh/ms_lec_6_ho.pdf [Accessed 12 August 2012].

Cogstate. 2012. *Homepage* [Online]. Available from: http://www.cogstate.com/ [Accessed 12 August 2012].

Collie, A., Darekar, A., Weissgerber, G., Toh, M. K., Snyder, P. J., Maruff, P. & Huggins, J. P. 2007. Cognitive testing in early-phase clinical trials: development of a rapid computerized test battery and application in a simulated Phase I study. *Contempory Clinical Trials,* 28**,** pp391-400.

Collie, A., Makdissi, M., Maruff, P., Bennell, K. & McCrory P. 2006. Cognition in the days following concussion: comparison of symptomatic versus asymptomatic athletes. *Journal of Neurology Neurosurgery & Psychiatry,* 77**,** pp241-5.

Collie, A., Maruff, P., Makdissi M., McCrory, P., McStephen, M. & Darby, D. G. 2003a. CogSport: Reliability and correlation with conventional cognitive tests used in post-concussion medical examinations. *Clinical Journal of Sport Medicine,* 13**,** 28-32.

Collie, A., Maruff, P., McStephen, M. & Darby, D. G. 2003b. Psychometric issues associated with computerised neuropsychological assessment of concussed athletes. *British Journal of Sports Medicine,* 37**,** pp556-9.

Czeisler, C. A., Duffy, J. F., Shanahan, T. L., Brown, E. N., Mitchell, J. F., Rimmer, D. W., Ronda, J. M., Silva, E. J., Allan, J. S., Emens, J. S., Dijk, D.-J. & Kronauer, R. E. 1999. Stablility, Precision, and Near-24-Hour Period of the Human Circadian Pacemaker. *Science,* 284**,** pp2177-2181.

Darby, D., Maruff, P., Collie, A. & McStephen, M. 2002. Mild cognitive impairment can be detected by multiple assessments in a single day. *Neurology,* 59**,** pp1042-6.

Darby, D. G., Brodtmann, A., Pietrzak, R. H., Fredrickson, J., Woodward, M., Villemagne, V. L., Fredrickson, A., Maruff, P. & Rowe, C. 2011. Episodic Memory Decline Predicts Cortical Amyloid Status in Community-Dwelling Older Adults. *Journal of Alzheimer's Disease,* 27**,** pp627-37.

Dawson, D. & Reid, K. 1997. Fatigue, alcohol and performance impairment. *Nature,* Vol 338.pp235-37.

Dekker, S. 2006. *The Field Guide to Understanding Human Error,* Farnham, Ashgate Publishing Limited.

Dement, D. 2012. *Stanford Sleepiness Scale* [Online]. Dr. Dement Sleep center. Available from: http://www.stanford.edu/~dement/sss.html [Accessed 12 August 2012].

Dinges, D. F. 1995. An overview of sleepiness and accidents. *Journal of Sleep Research,* 4**,** pp4-14.

Dinges, D. F., Ecker, A., Terry, D. & Powell, J. W. 2005. Pilot Test of Fatigue Management Technologies. Washington: FMCSA, US Department of Transportation and Transport Canada, Canadian Ministry of Transportation. (FMCSA-RT-05-002).

Dinges, D. F., Graeber, R. C., Rosekind, M. R., Samel, A. & Wegmann, H. M. 1996. Principles and Guidelines for Duty and Rest Scheduling in Commercial Aviation. *In:* NASA (ed.). Moffett Field, California: National Aeronautics and Space Administration. (NASA Technical Memorandum 110404).

Dinges, D. F., Pack, F., Williams, K., Gillen, K. A., Powell, J. W., Ott, G. E., Aptowicz, C. & Pack, A. I. 1997. Cumulative sleepiness, mood disturbance, and psychomotor vigilance performance decrements during a week of sleep restricted to 4-5 hours per night. *Sleep,* 4**,** pp267-277.

Dong, X. S. 2005. Long workhours, work scheduling and work-related injuries among construction workers in the United States. *Work Environment Health,* 31**,** pp329-335.

Dyken, M. E. & Im, K. B. 2009. Obstructive Sleep Apnoea and Stroke. *Chest,* 2009**,** pp1668-1677.

Eddy, D. R. & Hursh, S. R. 2001. Fatigue Avoidance Scheduling Tool Phase I SBIR Final Report. Brooks: Air Force Research Laboratory.

Ellis, K. A., Bush, A. I., Darby, D., De Fazio, D., Foster, J., Hudson, P., Lautenschlager, N. T., Lenzo, N., Martins, R. N., Maruff, P., Masters, C., Milner, A., Pike, K., Rowe, C., Savage, G., Szoeke, C., Taddei, K., Villemagne, V., Woodward, M. & Ames, D. 2009. The Australian Imaging, Biomarkers and Lifestyle (AIBL) study of aging: methodology and baseline characteristics of 1112 individuals recruited for a longitudinal study of Alzheimer's disease. *International Psychogeriatrics,* 21**,** pp672-87.

Energy Fiend. 2012. Caffeine Content of Drinks [Online]. Available from: http://www.energyfiend.com/ [Accessed 12 August 2012].

Enoka, R. M. & Stuart, D. G. 1992. Neurobiology of muscle fatigue. *Journal of Applied Physiology,* 72**,** pp1631-1648.

Everding, G. 1999. Warning system for fatigued drivers nearing reality with new eye data. *EurekAlert* [Online]. Available from: http://www.eurekalert.org/pub_releases/1999-07/WUiS-Wsff-260799.php# [Accessed 12 August 2012].

EyeDynamics Inc. 2012. *Home Web Page* [Online]. Available from: http://www.eyedynamics.com/products.htm. [Accessed 12 August 2012].

Federal Aviation Administration. 2010. FAA Proposes Safety Management Systems for Airlines. *In:* Federal Avaiation Administration (ed.). Washington: FAA.

Falleti, M. G., Maruff, P., Collie, A. & Darby, D. G. 2006. Practice effects associated with the repeated assessment of cognitive function using the CogState battery at 10-minute, one week and one month test-retest intervals. *Journal of Clinical and Experimental Neuropsychology,* 28**,** pp1095-112.

Falleti, M. G., Maruff, P., Collie, A., Darby, D. G. & McStephen, M. 2003. Qualitative similarities in cognitive impairment associated with 24 h of sustained wakefulness and a blood alcohol concentration of 0.05%. *Journal of Sleep Research,* 12**,** pp265-274.

Faraway J. J. 2006. *Extending the Linear Model with R: Generalized Linear, Mixed Effects and Nonparametric Regression Models,* Boca Raton, Taylor & Francis Group.

Fatigue Science. 2012. *Home page* [Online]. Honolulu. Available from: http://www.fatiguescience.com/home [Accessed 12 August 2012].

Fletcher, A., Lamond, N., Van Den Heuval, C. J. & Dawson, D. 2003. Prediction of Performance during Sleep Deprivation and Alcohol Intoxication using a Quantitative Model of Work-Related Fatigue,. *Sleep Research Online,* 5**,** pp67-75.

Flight Safety Foundation. 1995. Principles and Guidelines for Duty and Rest Scheduling in Corporate and Business Aviation. *Flight Safety Digest,* Vol 14.

Folkard, S. 1975. Diurnal variation in logical reasoning. *British Journal of Psychology*, 66 (1), pp1-8.

Folkard, S. & Akerstedt, T. 1987. Towards a Model for the prediction of alertness and/or fatigue on different sleep/wake schedules in Contemporary Advances in Shiftwork Research (Eds. A. Oginski, J. Pokorski & J. Rutenfranz), 231-240, Medical Academy, Krakow.

Folkard, S. & Akerstedt, T. 1991. A three process model of the regulation of altertness-sleepiness. *In:* Broughton, R. J. & Ogilvie, R. D. (eds.) *Sleep, Arousal and Performance: Problems and Promises.* Boston: Birkhauser.

Fredrickson, J., Maruff, P., Woodward, M., Moore, L., Fredrickson, A., Sach, J. & Darby, D. 2010. Evaluation of the usability of a brief computerized cognitive screening test in older people for epidemiological studies. *Neuroepidemiology,* 34**,** pp65-75.

Garbarino, S. 2008. Sleep disorders and road accidents in truck drivers. *Giornale italiano di medicina del lavoro ed ergonomia,* 30**,** pp291-6.

Gertler, J., Viale, A. & Raslear, T. 2006. Work Schedules and Sleep Patterns of U.S. Railroad Signalmen. Washington: The Federal Railroad Administration Office of Research and Development.

Giannini, A. J. 1992. Chronic Fatigue. *In:* Taylor, R. B. (ed.) *Difficult Diagnosis.* Philadelphia: WB Saunders Co.

Gillingham, F. M. 2012. *Jet Lag - New Approaches to an Old Problem* [Online]. Available from: http://www.hthtravelinsurance.com/travel_center/business/ne_el_01.cfm [Accessed 12 August 2012].

Gillis, W. 2012. Sleepy Air Canada pilot mistook planet for plane, report finds. *The Star*, April 17, 2012.

Gorman, C. 2004. Why We Sleep. *Time* [Online]. Available from: http://www.time.com/time/2006/dowchemical/. [Accessed 12 August 2012].

Gruber, A. J., Hudson, J. I. & Pope, H. G. 1996. The Management of Treatment-Resistant Depression In Disorders on the Interface of Psychiatry and Medicine: Fibromyalgia, Chronic Fatigue Syndrome, Migraine, Irritable Bowel Syndrome, Atypical Facial Pain, and Premenstrual Dysphoric Disorder *Psychiatric Clinics of North America* 19.

Gunzelmann, G., Moore, R. L. J., Gluck, K. A. & Van Dongen, H. P. A. 2008. Individual Differences in Sustained Vigilant Attention: Insights from Computational Cognitive Modelling.

Hassen, F. 2002. Safety: *Sleep Deprivation: Effects on Safety, Health and the Quality of Life*.

Health and Safety Executive (HSE), 2009. *Reducing error and influencing behaviour.* [Online], Available from: http://www.hseni.gov.uk/hsg_48_reducing_error_and_influencing_behaviour.pdf [Accessed 18 October 2012].

HealthCommunities.com. 2012. *Basics of Sleep Disorders* [Online]. HealthCommunities.com. Available from: http://www.healthcommunities.com/sleep-disorders/overview-of-sleep-disorders.shtml [Accessed 12 August 2012].

Hegarty, S. 2012. The myth of the eight-hour sleep. *BBC News Magazine* [Online], February. Available from: http://www.bbc.co.uk/news/magazine-17101370 [Accessed 12 August 2012].

Hokama, Y., Campora, C. E., Hara, C., Kuribayashi, T., Le Huynh, D. & Yabusaki, K. 2009. Anticardiolipin Antibodies in the Sera of Patients with Diagnosed Chronic Fatigue Syndrome. *Journal of Clinical Laboratory Analysis,* 23**,** pp210-212.

Horne, J. 2008. How much sleep do you really need? [Online]. Available from: http://www.dailymail.co.uk/health/article-1078350/How-sleep-really-need-Probably-lot-think-says-expert.html?printingPage=true [Accessed 12 August 2012].

Horne, J. A. 2004. Is there a sleep debt? . *Sleep,* 27**,** pp1047-1049.

Howard, S. K., Gaba, D. M., Rosekind, M. R. & Zarcone, V. P. 2002. The risks and implications of excessive daytime sleepiness in resident physicians. *Academic Medicine,* 77(10)**,** pp1019-25.

Hursh, S., Raslear, T. G., Kaye, A. S. & Fanzone, J. F. 2006. Validation and Calibration of a Fatigue Assessment Tool for Railroad Work Schedules, Summary Report. Washington: Federal Railroad Administration. (DOT/FRA/ORD-06/21).

Hursh, S. R., Elsmore, T. & Eddy, D. R. 2008. Interface for a System and Method for Evaluation Task Effectiveness Based on Sleep Pattern. United States patent application 11/275,304.

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,* 00**,** pp1-10.

Johns, D. M. 2012. The Epworth Sleepiness Scale [Online]. Available from: http://epworthsleepinessscale.com/ [Accessed 12 August 2012].

Kecklund, G., Ekstedt, M., Akerstedt, T., Dahlgren, A. & Samuelson, B. 2001. The effects of double-shifts (15.5 hours) on sleep, fatigue and health. [Online] Available from: http://www.ncbi.nlm.nih.gov/sites/entrez?cmd=Retrieve&db=PubMed&list_uids=14564858&dopt=Citation.[Accessed 12 August 2012].

Lamond, N. & Dawson, D. 1999. Quantifying the Performance Impairment Associated with Fatigue. *Journal of Sleep Research.,* 8**,** pp255-262.

Laughlin, M. 2003. Perchance to Dream. *Aviation Safety Vortex* [Online]. Available from: http://www.tc.gc.ca/eng/civilaviation/publications/tp202-2-03-v041-3339.htm. [Accessed 12 August 2012].

Lee-Chiong, T. L. 2008. Sleep Medicine: Essentials and Review. New York: Oxford University Press.

Lerman, S. E., Eskin, E., Flower, D. J., George, E. C., Gerson, B., Hartenbaum, N., Hursh, S. R. & Moore-Ede, D. M. 2012. Fatigue Risk Management in the Workplace. *American College of Occupational and Environmental Medicine,* 54**,** pp231-258.

Levine, B., Roehrs, T., Zorick, F. & Roth, T. 1988. Daytime sleepiness in young adults. *Sleep,* 11**,** pp39-46.

Lim, J. & Dinges, D. F. 2010. A meta-analysis of the impact of short-term sleep deprivation on cognitive variables. *Psychological Bulletin,* 136(3)**,** pp375-389.

Logan, G. 2012. How to Prevent Jet lag. [Online] Available from: http://traveltips.usatoday.com/long-over-jet-lag-63114.html [Accessed 12 August, 2012].

Loh, S., Lamond, N., Dorrian, J., Roach, G. & Dawson, D. 2004. A system and a method for providing a determination of predicted cognitive performance of an individual preferably based on the time of day and on factors including sleep history based on activity data from an actigraph and the individual's activities. *Behaviour research methods, instruments, & computers: a journal of psychonomic Society Inc.,* 2**,** pp339-346.

MacCollum, D. 2006. *Construction Safety Engineering Principles: Designing and Managing Safer Job Sites.* New York: McGraw-Hill.

McCullogh, K., Baker, A., Ferguson S., Fletcher, A., Dawson, D. 2008. Developing and Implementing a Fatigue Risk Management System. Ottawa: Transport Canada, Canadian Ministry of Transportation. (TP 14575E).

Mainster, M. & Turner, P. 2012. Blue Light's benefits vs blue-blocking intraocular lens chromophores. *Graefes Archive for Clinical Experimental Ophthalmology.* (2012) 250,pp1245-1246.

Makdissi, M., Collie, A., Maruff, P., Darby, D. G., Bush, A., McCrory, P. & Bennell, K. 2001. Computerised cognitive assessment of concussed Australian Rules footballers. *British Journal of Sports Medicine,* 35**,** pp354-60.

Makdissi, M., Darby, D., Maruff, P., Ugoni, A., Brukner, P. & McCrory, P. R. 2010. Natural history of concussion in sport: markers of severity and implications for management. *American Journal of Sports Medicine,* 38**,** pp464-71.

Makeig, S. & Jung, T.-P. 1995. Changes in Alertness Are a Principal Component of Variance in the EEG Spectrum. *Neuroepidemiology,* 7**,** pp213-216.

Maruff, P., Collie, A., Darby, D., Weaver-Cargin, J. & McStephen, M. 2004. Subtle memory decline over 12 months in mild cognitive impairment. *Dementia and Geriatric Cognitive Disorders,* 18**,** pp342-348.

Maruff, P., Falleti, M. G., Collie, A., Darby, D. G. & McStephen, M. 2005. Fatigue-related impairment in speed, accuracy and variability of psychomotor performance: comparison with blood alcohol levels. *Journal of Sleep Research,* 14**,** pp21-27.

Maruff, P., Thomas, E., Cysique, L., Brew, B., Collie, A., Snyder P. & Pietrzak, R. H. 2009. Validity of the CogState brief battery: relationship to standardized tests and sensitivity to cognitive impairment in mild traumatic brain injury, schizophrenia, and AIDS dementia complex. *Archives of Clinical Neuropsychology,* 24**,** pp165-78.

Max, D. T. 2010. The Secrets of Sleep. *National Geographic.* National Geogrpahic Society. May 2010.

Miller, J. C. 1996. Fit for Duty. *Ergonomics in Design*.4(2).

Monk, T. H. 1991. *Sleep, sleepiness and performance,* Oxford, England, John Wiley & Sons.

Monk, T. & Embry, D. 1981. A field study of circadian rhythms in actual and interpolated task performance. *In Night and Shift Work: Biological and Social Aspects (Eds. A. Reinberg, N. Vieux & P. Andlauer)*, 473-480, Pergamon Press, Oxford.

Moore-Ede, D. M. 2012. Fatigue Accident/Incident Causation Testing System: FACTS. [Online]. Available from: http://facts.circadian.com/facts/ [Accessed 12 August 2012].

Morad, Y., Barkana, Y., Zadok, D., Hartstein, M., Pras, E. & Bar-Dayan, Y. 2009. Ocular parameters as an objective tool for the assessment of truck drivers fatigue. *Accident; analysis and prevention,* 41**,** pp856-60.

Morgenthaler, T., Alessi, C., Friedman, L., Owens, J., Kapur, V., Boehlecke, B., Brown, T., Chesson, A., Coleman, J., Lee-Chong, T., Pancer, J. & Swick, T. 2007. Practice Parameters for the Use of Actigraphy in the Assessment of Sleep and Sleep Disorders: An Update for 2007. *Sleep,* Vol. 30.

Moriarity, J., Collie, A., Olson, D., Buchanan, J., Leary, P., McStephen, M. & McCrory, P. 2004. A prospective controlled study of cognitive function during an amateur boxing tournament. *Neurology,* 62**,** pp1497-502.

National Institute For Occupational Safety and Health. 2007. *Construction Safety* [Online]. Available from: http://www.cdc.gov/niosh/topics/constructionsafety/.[Accessed 12 August 2012].

National Institutes of Health. 2003. Sleep, Sleep Disorders, and Biological Rhythms. Colorado Springs, Colorado: BSCS.

National Sleep Foundation. 2012. *How Much Sleep Do We Need?* [Online]. National Sleep Foundation. Available from: http://www.sleepfoundation.org/article/how-sleep-works/how-much-sleep-do-we-really-need [Accessed 12 August 2012].

National Transportation Safety Board. 1990. Marine Accident Report - Grounding of the US Tankship Exxon Valdez on Bligh Reef, Prince William Sound, near Valdez Alaska, March 24, 1989. Washington, DC: National Transportation Safety Board.

National Transportation Safety Board. 1999a. Evaluation of U.S. Department of Transportation Efforts in the 1990s to Address Operator Fatigue. Washington, DC: Department of Transportation.

National Transportation Safety Board. 1999b. Website [Online]. Available from: http://www.ntsb.gov/ [Accessed 12 August 2012].

Nehlig, A. 2010. Is Caffeine a Cognitive Enhancer? *Journal of Alzheimer's Disease,* 20**,** pp85-94.

OHS Insider. 2010. Are Employers Liable For Traffic Accidents Caused by Fatigued Workers After Their Shifts? *OHS Insider.com* [Online]. Available from: http://ohsinsider.com/insider-top-stories/are-employers-liable-for-traffic-accidents-caused-by-fatigued-workers-after-their-shifts [Accessed 12 August 2012].

Page, R. M. 1929. On Supplanting the Industrial Fatigue Concept. *The Journal of Business of the University of Chicago,* 2**,** pp137-150.

Petroleum Industry Training Service (PITS). 2003. *Fatigue Management - CPSC Guide to Safe Work* [Online]. Available from: www.pits.ca [Accessed 12 August 2012].

Pietrzak, R. H., Olver, J., Norman, T., Piskulic, D., Maruff, P. & Snyder, P. J. 2009. A comparison of the CogState Schizophrenia Battery and the Measurement and Treatment Research to Improve Cognition in Schizophrenia (MATRICS) Battery in assessing cognitive impairment in chronic schizophrenia. *Journal of Clinical and Experimental Neuropsychology***,** pp1-12.

Piettrzak, R. H., Snyder, P. J., Jackson, C. E., Olver, J., Norman, T., Piskulic, D. & Maruff, P. 2009. Stability of cognitive impairment in chronic schizophrenia over brief and intermediate re-test intervals. *Human Psychopharmacology,* 24**,** pp113-21.

PMI, I. 2012. *Home Web Page* [Online]. Available from: http://www.pmifit.com/ [Accessed 12 August 2012].

Powell, R. 2009. Fatigue-Impairment in British Columbia's Construction Industry. 27th International System Safety Conference -The Win Win Solution, 3 - 7 August, 2009, Huntsville, Alabama. Huntsville, Alabama: International System Safety Society.

Powell, R. 2010. Impairment in the Workplace ....Fatigue Gets a "Free" Pass. 28th International System Safety Conference - Building Bridges to a Safer Future, August 30 - September 3, 2010, Minneapolis, Minnesota. Minneapolis: International System Safety Society.

Powell, R. & Copping, A. G. A. 2010. Sleep Deprivation and Its Consequences in Construction Workers. *Journal of Construction Engineering and Management,* 136**,** pp1086-1092.

Powell, R. I. 2007. *Fatigue in British Columbia's Construction Industry. Understanding Risk Levels and Opportunities for Improvement.* MSc Dissertation, University of Bath, Bath.

Purdy, M. 2006. First biomarker for human sleepiness identified in fruit flies. *Medical News.* St. Louis: Washington University.

Purves, D., Augustine, G. J., Fitzpatrick, D., Lawrence, C. K., Lamantia, A.-S., McNamara, J. O. & Williams, S. M. (eds.) 2001. *Neuroscience*: Sinauer Associates Ltd.

Quinn, S. 2010. Paramedics confirmed dead in Vancouver Island lake tragedy. *Alberni Valley News*, October 19, 2010.

R Development Core Team. 2010. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Reason, J. 2008. *The Human Contribution : Unsafe Acts, Accidents and Heroic Recoveries,* Surrey, Ashgate Publishing Limited.

Rechtschaffen, A., Gilliland, M. A., Bergmann, B. M. & Winter, J. 1983. Physiological correlates of prolonged sleep deprivation in rats. *Science,* 221**,** pp182-184.

Rhodes, W. & Vincent, A. 2000. Assessment of Fitness-for-Duty Technologies in Transport Operations. Montreal: Transport Canada.

Rosekind, M. R. 2000. Stress, the Business Traveler and Corporate Health: An International Travel Health Symposium. Washington, D.C.:The World Bank.

Rosekind, M. R., Co, E. L., Gregory, K. B. & Miller, D. L. 2000. Crew Factors in Flight Operations. [Online]. Available from: http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20010039028_2001056921.pdf [Accessed 12 August 2012].

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. *Behavioural Medicine,* Vol 21**,** pp157 - 165.

Roth, T. 2006. *Dealing with Excessive Sleepiness: Therapeutics and Public Health Considerations* [Online]. Available from: http://www.medscape.com/viewprogram/4394. [Accessed 12 August 2012].

Rozell, N. 1995. Fly East for Bad Jet Lag. *Alaska Science Forum* [Online]. Available from: http://www.gi.alaska.edu/ScienceForum/ASF12/1261.html [Accessed 12 August 2012].

Russo, M. D. 2005. *Normal Sleep, Sleep Physiology, and Sleep Deprivation: General Principles* [Online]. Available from: http://www.emedicine.com/neuro/topic444.htm. [Accessed 12 August 2012].

Science Daily. 2009. Support Cells, Not Nuerons, Lull the Brain to Sleep. *ScienceDaily*.[Online]. Available from: http://www.sciencedaily.com/releases/2009/01/090128122814.htm [Accessed 12 August 2012].

Siegel, J. M. 2003. *Why We Sleep* [Online]. Available from: http://www.npi.ucla.edu/sleepresearch/sciam2003/sciamsleep.pdf. [Accessed 12 August 2012].

Sleepfoundation. 2007. *How Much Sleep Do We Really Need* [Online]. Available from: www.sleepfoundation.org. [Accessed 12 August 2012].

Smithson, R. 2008. Industry Takes Lead in Workplace Impairment Testing. *HRVoice.org* [Online], Available From: http://www.hrvoice.org/story.aspx?storyid=5384&issueid=879&pagemode=displaystory&regionid=10#postcommentpostcomment [Accessed 10 March 2010].

Snyder, P., Bednar, M., Cromer, J. & Maruff, P. 2005. Reversal of scopolamine-induced deficits with a single dose of donepezil, an acetylcholinesterase inhibitor. *Alzheimer's & Dementia,* 1**,** pp126-135.

Snyder, P., Werth, J., Giordani, B., Caveney, A. F., Feltner, D. & Maruff, P. 2005. A method for determining the magnitude of change across different cognitive functions in clinical trials: the effects of acute administration of two different doses alprazolam. *Human Psychopharmacology,* 20**,** pp263-73.

Straume-Naesheim, T. M., Andersen, T. E., Im, K. H., McIntosh, A. S., Dvorak, J. & Bahr, R. 2009. Do minor head impacts in soccer cause concussive injury? A prospective case-control study. *Neurosurgery,* 64**,** pp719-25.

Strauss, S. 2003. Pilot Fatigue. [Online]. Available from: http://aeromedical.org/Articles/Pilot_Fatigue.html [Accessed 12 August 2012].

The Canadian Institutes of Health Research. 2012. *The Neuronal Switches for Waking and Sleeping* [Online]. Available from: http://thebrain.mcgill.ca/flash/i/i_11/i_11_cl/i_11_cl_cyc/i_11_cl_cyc.html [Accessed 12 August 2012].

Tierney, J. 2011. Do You Suffer From Decision Fatigue. *The New York Times* [Online]. Available from: http://www.nytimes.com/2011/08/21/magazine/do-you-suffer-from-decision-fatigue.html?pagewanted=all [Accessed 12 August 2012].

Transport Canada. 2007. *Aviation Safety Vortex* [Online]. Available from: http://www.tc.gc.ca/CivilAviation/. [Accessed 12 August 2012].

Turner, T. H., Drummond, S. P. A., Salamat, J. S. & Brown, G. G. 2007. Effects of 42 hr of total sleep deprivation on component processes of verbal working memory. *Neuropsychology,* 21**,** pp787-795.

U.S. Chemical Safety and Hazard Investigation Board 2007. Investigation Report Refinery Explosion and Fire (15 killed, 180 Injured).[Online]. Available from: http://www.csb.gov/assets/document/CSBFinalReportBP.pdf [Accessed 12 August 2012].

Van Dijk, F. J. H. & Swaen, G. M. H. 2003. Fatigue at work. *Occupational and Environmental Medicine*.60, pp1-2.

Van Dongen, H., Rogers, N. L. & Dinges, D. F. 2003. Sleep debt: Theoretical and empirical issues. *Sleep and Biological Rhythms,* 1**,**pp 5-13.

Van Dongen, H. P. A. 2004. Comparison of Mathematical Model Predictions to Experimental Data of Fatigue and Performance. *Aviation, Space and Environmental Medicine,* 75**,** A15 - 36.

Van Orden, K. F., Jung, T.-P., & Makeig, S. 2000. Combined eye activity measures accurately estimate changes in sustained visual task performance. *Biological Psychology,* 2**,** pp221-240.

Vice-Chancellor's Office , University of Bath. 2008. Summary notes of the Executive Committee held on 30 January 2008. 17 July, 2008.

Walker, D. 2007. Applying the Human Factors Analysis and Classification System (HFACS) to incidents in the UK Construction industry. Dissertation (MSc), Cranfield University.

Walters, N. & Bolt, H. 2009. Phase 1 Report: Underlying causes of construction fatal accidents – A comprehensive review of recent work to consolidate and summarise existing knowledge. Norwich: Crown.

Waterhouse, J. 1999. Jet-lag and shift work: (1) circadian rhythms. *Journal of the Royal Society of Medicine,* 92**,** pp398-401.

WebMD. 2012. The Causes of Fatigue. *WebMD* [Online]. Available from: http://www.webmd.com/sleep-disorders/ss/slideshow-fatigue-causes-and-remedies [Accessed 12 August 2012].

Wehr, T. A., Moul, D. E., Barbato, G., Giesen, H. A., Seidel, J. A., Barker, C. & Bender, C. 1993. Conservation of photoperiod-responsive mechanisms in humans. *American Journal of Physiology,* Vol 265**,** pp846-857.

Wirz-Justice, A. 207. How to measure circadian rhythms in humans. *Medicographia,* 29**,** pp84-90.

Workrights Institute. 2009. Impairment Testing - Does it Work? [Online]. Available from: http://www.workrights.org/issue_drugtest/dt_impairment_testing.html [Accessed 22 February 2010].

Wylie, C. D. 2005. Sleep, science, and policy change *New England Journal of Medicine,* 352(2)**,** pp196-7.

Wang, Y., Ke, C. & Brown, M. B. 2000. Shape Invariant Modelling of Circadian Rhythms with Random Effects and Smoothing Spline ANOVA Decompositions. *Biometrics*, 59, pp804-812.