Investigating sleepiness and distraction in simple and complex tasks

This item was submitted to Loughborough University’s Institutional Repository by the/an author.

Additional Information:

- A Doctoral Thesis. Submitted in partial fulfilment of the requirements for the award of Doctor of Philosophy of Loughborough University.

Metadata Record: https://dspace.lboro.ac.uk/2134/19123

Publisher: © Alan wales

Rights: This work is made available according to the conditions of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) licence. Full details of this licence are available at: https://creativecommons.org/licenses/by-nc-nd/4.0/

Please cite the published version.
University Library

Author/Filing Title .............................................. WALES ...............................................................
........................................................................
Class Mark ............................................................ 7
........................................................................

Please note that fines are charged on ALL overdue items.
Investigating Sleepiness And Distraction In
Simple And Complex Tasks

A doctoral thesis
April, 2009

© Alan Wales

Department of Human Sciences
Loughborough University

Submitted in partial fulfilment of the requirements for the award of Doctor of
Philosophy of Loughborough University.

Key words: Sleepiness, Distraction, Vigilance, Psychomotor, Multi-Attribute
Task, Luggage Search
Abstract

The cost of sleepiness-related accidents runs into tens of billions of dollars per year in America alone (Leger, 1994), and can play a contributing role in motor vehicle accidents and large-scale industrial disasters (Reason, 1990). Likewise, the effects of an ill-timed distraction or otherwise lack of attention to a main task can be the difference between elevated risk, or simply a lack of productivity. The interaction between sleepiness and distraction is poorly researched, and little is known about the mechanisms and scale of the problems associated by this interaction. Therefore, we sought to determine the effects of sleepiness and distraction using overnight and daytime sleepiness with various levels of distraction on three tasks ranging from a simple vigilance task to a challenging luggage x-ray inspection task.

The first and second studies examined overnight sleepiness (7pm to 7am) for twenty-four healthy participants (m = 23.2yrs old – same for both studies) using a psychomotor task compared to a systems monitoring task, while also manipulating peripheral distraction through a television playing a comedy series. The results showed significant effects of sleepiness on the psychomotor task and evidence for interactive effects of distraction, whereas the systems monitoring task showed no changes with either sleepiness or distraction. Subjects were far more prone to distraction when sleepy for both tasks, and EEG findings suggest that the alpha frequency (8-13Hz) power increases reflect impairments of performance. There is a decaying exponential relationship between the probability of a subject's eyes being open as the response time increases, such that longer responses above three seconds are 95% likely to have occurred with the eyes closed.

The third study used a sample of twelve young (m = 20.8yrs) and twelve older (m = 60.0yrs) participants, and examined the effects of sleep restriction (< 5hrs vs normal sleep) with three levels of distraction (no distraction, peripheral in the form of television and cognitive distraction as a simulated conversation by means of verbal fluency task). The task used was an x-ray luggage search simulator that is functionally similar to the task used for airport security screening. The practice day showed that speed and accuracy on the task improved with successive sessions, but that the older group were markedly slower and less accurate than the younger group even before the experimental manipulations. There was no effect of daytime sleep restriction for either the younger or older groups between the two experimental days. However, distraction was found to impair the performance of both young and old, with the cognitive distraction proving to be the most difficult condition.

Overall, it is concluded that overnight sleepiness impairs performance in monotonous tasks, but these risks can be diminished by making tasks more engaging. Distractions can affect performance, but may be difficult to quantify as subjects create strategies that allow themselves to attend to distractions during the undemanding moments of a task. Continuous cognitive distraction does affect performance, particularly in older subjects, who are less able to manage concurrent demands effectively. Humans appear capable of coping
with a 40% loss of their usual sleep quota or 24-hours of sleep restriction on complex tasks, but performance degrades markedly on monotonous tasks. Performances for simple and complex tasks are impaired by distracters when the effect of distraction is large enough, but the magnitude of impairment depends on how challenging the task is or how well the subject is able to cope with the distractions.
Acknowledgements

Firstly, I owe a huge debt of gratitude to Dr Clare Anderson, who encouraged me to pursue a Doctorate degree and has brought me from being a novice undergraduate into a productive academic. I am also indebted to Prof. Jim Horne, who has helped me fulfil my potential as a researcher and has provided a remarkable example for me to follow. Without the hard work these two have put into my research and career I would not be the person I am today and for that I am truly grateful.

I would also like to acknowledge my parents for the love and support they have given me throughout my life, and who have taught me to value education and hard work. I thank Katherine Jones for the work she put into these projects and help with data collection. I would like to extend my gratitude to my colleagues at the Sleep Research Centre and my Director of Research, Prof. Kevin Morgan, who have made my studies enjoyable. I'd like to thank Prof. Adrian Schwaninger for his assistance and inviting me to work for his group in Switzerland.

Finally, I would like to thank my girlfriend Pamela Nicholson for her love and companionship, and taking the risk of moving abroad with me.
Title Page
Certificate Of Originality
Abstract
Acknowledgements
Table of Contents
List Of Tables
List Of Figures
CHAPTER ONE
Literature Review
1.1.1 Research Agenda
The Dangers Of Sleepiness
1.1.2 The Costs Of Sleep-Related Accidents
1.1.3 Losing Sleep
1.1.4 Occupational Vigilance
1.1.5 Shift Work
1.1.6 Fluctuations In Alertness
1.1.7 Assessing Sleepiness
Distraction
1.2.1 Irrelevant And Confounding Variables
1.2.2 Distraction And Divided Attention
1.2.3 Measuring Distraction
1.2.4 Driver Distraction
Applied Visual Inspection
1.3.1 Visual Search Theory
1.3.2 From A Radiological Background
1.3.3 X-Ray Luggage Search
Investigating Sleep In A Controlled Environment
1.4.1 Task Durations
1.4.2 Lapse Hypothesis
Performance Changes With Age
1.5.1 Sleepiness Affects Older Subjects
1.5.2 Elderly Reactions
1.5.3 Elderly Distraction
1.5.4 Visual Search Studies Using Older Subjects .......................................................... 21

Neurobiology And The Electroencephalogram .......................................................... 22
1.6.1 The Pre-Frontal Cortex ......................................................................................... 22
1.6.2 Electroencephalogram ......................................................................................... 23
1.6.3 Alpha And Theta Bandwidths ............................................................................... 23
1.6.4 Neurobiology In Vigilance Research ................................................................... 23

Dimensions Of Personality ....................................................................................... 24
1.7.1 Personality Scales ................................................................................................. 24
1.7.2 Introverts And Extraverts ................................................................................... 25
1.7.3 Anxiety ................................................................................................................ 26
1.7.4 Attention, Arousal And Alertness ........................................................................... 26

CHAPTER TWO ........................................................................................................... 28
Methodology ................................................................................................................ 28
Tasks Utilised ................................................................................................................. 29
2.1.1 The Psychomotor Vigilance Task ........................................................................ 29
2.1.2 The Multi-Attribute Task Battery ........................................................................ 30
2.1.3 The X-Ray Object Recognition Task ................................................................... 31

Measurements ................................................................................................................. 32
2.2.1 The PVT ................................................................................................................. 32
2.2.2 The MATB .............................................................................................................. 33
2.2.3 The X-Ray ORT .................................................................................................... 33
2.2.4 Video Ratings Of Behaviour ................................................................................. 34
2.2.5 The Electroencephalogram ................................................................................... 35
2.2.6 Actigraphy ............................................................................................................. 36

Building The Laboratory ............................................................................................... 36
2.3.1 Planning The Distraction Laboratory ................................................................... 36
2.3.2 Piloting The Distraction Laboratory .................................................................... 40
2.3.3 Participant Screening ............................................................................................ 41
2.3.4 From The Laboratory To The Workplace ........................................................... 42

Preamble To Chapters ..................................................................................................... 43
2.4.1 From The PVT To The ORT ................................................................................ 43

CHAPTER THREE ......................................................................................................... 44
Overnight Sustained Vigilance And Distraction ............................................................ 44
Abstract ........................................................................................................................... 44
3.1.1 Monotonous Work ................................................................................................. 45
3.1.2 Risks Associated With Loss Of Sleep ................................................................... 46
3.1.3 Psychomotor Vigilance ................................................................. 47
3.1.4 Distractions During Office Work ............................................... 48
3.1.5 The Psychomotor Vigilance Task ................................................ 49
3.1.6 The Lapse Construct .................................................................. 50
3.1.7 A Sleep-Restricted Pilot Study .................................................. 51
3.1.8 Summary .................................................................................... 52
Methods ............................................................................................. 53
3.2.1 Screening And Subjects .............................................................. 53
3.2.2 Protocol ....................................................................................... 54
3.2.3 Laboratory Settings ...................................................................... 55
3.2.4 Scoring Distractions ................................................................. 56
3.2.5 Subjective Scales ..................................................................... 56
Results ............................................................................................... 56
3.3.1 Descriptives ................................................................................ 56
3.3.2 Sleepiness And Fatigue ............................................................... 57
3.3.3 Distraction .................................................................................. 61
3.3.4 Sleepiness x Distraction .............................................................. 64
3.3.5 Lapse Durations ......................................................................... 66
3.3.6 Personality .................................................................................. 72
3.3.7 EEG Analysis ............................................................................. 74
Discussion .......................................................................................... 80
CHAPTER FOUR ..................................................................................... 85
Multi-Tasking And Distraction Under Sleep Pressure .......................... 85
Abstract ........................................................................................... 85
4.1.1 Resourceful Work ....................................................................... 86
4.1.2 Military Operations ...................................................................... 86
4.1.3 Pilot Studies ................................................................................ 87
4.1.4 Multi-Tasking And Divided Attention .......................................... 88
4.1.5 Multiple Component Tasks ....................................................... 89
4.1.6 Psychophysiology Of Sleepiness And Task Load ...................... 90
4.1.7 Summary .................................................................................... 90
Methods ............................................................................................ 91
4.2.1 Study Protocol ............................................................................ 91
4.2.2 MATB Settings ........................................................................... 91
Results ............................................................................................... 93
4.3.1 Data Collection .......................................................................... 93
6.1.3 Simple Vs Complex Tasks ............................................................... 150
6.1.4 Speed Vs Accuracy ................................................................. 151
6.1.5 Personality Dimensions ......................................................... 151
6.1.6 Transport Security ................................................................. 152

Experimental Design: A Discussion ........................................... 152
6.2.1 Laboratory And Study Design ........................................... 152
6.2.2 Electroencephalography ..................................................... 153

Future Research ................................................................. 153
6.3.1 Chronic Sleep Loss ............................................................ 153
6.3.2 Modelling Lapse Behaviour ............................................... 154

Concluding Remarks ........................................................... 155
6.4.1 Conclusion ................................................................. 155

References ................................................................. 156

Appendix ................................................................. 195
Appendix A: Glossary .............................................................. 195
Appendix B: PVT and MATB Distraction Scoring Card .............. 196
Appendix C: The International 10-20 Electrode Placement System 197
Appendix D: Fast-Fourier EEG Analyses ...................................... 198
Appendix E: Study Questionnaires ............................................... 208
List Of Tables

Table 1: Definitions of distraction in the literature (Roper and Juneja, 2006) ........ 4
Table 2: Comparison between screening guidelines for young and old subjects. ........................................................................................................42
Table 3: Counterbalancing of the conditions meant that there was no bias towards sleepiness or distraction. This Latin-squares design was repeated for alert and sleepy conditions for the same participants.............55
Table 4: Lapse duration and quartile breakdown for the sample of 24 subjects. ....................................................................................................69
Table 5: Lapse duration and quartile breakdown when the major outlier, subject no. 6 is removed. The largest change is seen at the 95th percentile. ........................................................................................................69
Table 6: The decaying exponential model for all subjects is applied to individuals and shown for predictability. The 95% confidence limits were highly predictive of EC and EO lapses, and the 50% confidence limits performed well. The final column shows the midpoint in milliseconds where the probability of an EC lapse exceeds that of EO for each subject. Subject six is removed from the average of the individual midpoints.......71
Table 7: The decaying exponential model for all subjects is applied to individuals and shown for predictability similar to Table 6, but with the new cutoffs derived from Table 5. There is a marginal increase in predictability of the model for these new cutoffs. ................................................72
Table 8: Split-half confirmation of group differences. ......................................73
Table 9: Comparing personality dimensions to distraction and performance.73
Table 10: Baseline alpha levels for C4-A1 recordings for a subsample of sixteen subjects. .....................................................................................77
Table 11: Ranked performance comparisons between PVT mean reaction time in each condition and the ordered ranks of alpha and theta levels during the task for both hemispheres. “1” means perfect agreement, “-” means disagreement, and “*” represents those trials that were eliminated from the analysis because of excessive artifact or data corruption. Theta results are similar to levels of chance, whereas alpha results show higher accordance with the performance data. NN = Night, No Distraction, ND = Night Distraction, MN = Morning, No Distraction, MD = Morning, Distraction..........................................................79
Table 12: Chi-square cross tabulations for the two recording sites and associated significances. Alpha rankings show highly significant results compared to theta, which are a poor indicator of performance.............78
Table 13: Grand interhemispheric correlations per condition for all subjects. A matrix of bins per subject for the C4 cortex region was compared to the matrix derived from the C3 scalp location. R = correlation coefficient, N = number of data points. ...........................................................................78
Table 14: Personality predictors of MATB performance for the split-half analysis whereby participants were median-split into two groups of high and low values.................................................................99
Table 15: Ranked comparisons between MATB lapses (missed dials and lights), alpha and theta increases. If performances coincided in rank, then a “1” is marked for accordance. Any misalignment results in a “-” (counted as zero), whereas damaged data was omitted (“*”). The alpha findings
are related to MATB performance, whereas theta does not exceed chance. .................................................................104

Table 16: Ranked comparisons between MATB resource deviations, alpha and theta increases. The alpha findings are related to MATB performance for both EEG recording sites, whereas theta does not exceed chance. 105

Table 17: Chi-square cross tabulations for the two recording sites and associated significances. Alpha rankings show highly significant results compared to theta, which are a poor indicator of performance. .......106

Table 18: Grand interhemispheric correlations per condition for all subjects. A matrix of bins per subject for the C4 cortex region was compared to the matrix derived from the C3 scalp location. R = correlation coefficient, N = number of data points .................................................................106

Table 19: Counterbalancing table for the young and old groups. This protocol ensured that distraction and sleepiness were not favoured in the experimental design for treatment order effects ........................................121

Table 20: Model fit parameters for Drury’s model. Mean value and standard deviation are given for young and old, untrained (u) and trained (t).....126

Table 21: Inferential tests on Non-Search Time (NST) and Search Time (ST) for hits (h) and false alarms (fa). Only NSTh and NSTfa improved significantly with practice, showing no age effects ........ ..........126

Table 22: Correlations between detection performance (A') and confidence. .............................................................................................................. 133

Table 23: Personality differences between the young and old group. Only extraversion differed between the groups significantly. ...........134

Table 24: Split-half personality dimensions ........................................134

Table 25: Predictors of performance when nonsleepy and nondistracted for detection measures (A'). Several scattered predictors achieve significance such as arousal, anxiety and neuroticism while psychoticism and extraversion do not. .......................................................................................135

Table 26: Sleepiness changes across the conditions for the two brain regions collapsed across the age groups. There is no strong evidence that either brain region is activated with sleep restriction, which follows the main results ........................................................................................................... 137

Table 27: Alpha changes with sleepiness for the two brain regions, collapsed across the age groups. There are no strong results that might suggest that alpha activity is heightened significantly. ..........................137

Table 28: Tabulation of overall power bin (0.1-13Hz = 13 bins) changes comparing all conditions. NS = Normal Sleep, SR = Sleep Restricted, ND = No Distraction, D = Peripheral Distraction, DC = Cognitive Distraction. .................................................................................................................. 137

Table 29: Tabulation of alpha power (8-13Hz) changes for all conditions. There is an alpha power increase for sleep restriction in young subjects, and for the more challenging distraction conditions. Older subjects show the opposite results for alpha sleepiness ........................................138

Table 30: Tabulation of theta power (4-8Hz) changes for all conditions. There is no clear link between theta power and distraction condition with the exception of frontal D-ND for young participants ..........138
List Of Figures

Figure 1: Unintentional sleep episodes by time (Carskadon, Littell and Dement, 1985). A count of unintended sleep episodes aggregated from two studies show the two major peaks in sleepiness...........................................5

Figure 2: Z-Scores of performance efficiency during a full day (Folkard and Tucker, 2003, p. 96). Z-scores of performance efficiency, as measured by reading errors and excessive time to manage simple tasks, during the 24-hour cycle. .........................................................................................7

Figure 3: Publications in visual search by year (Yang et al, 2002). The number of papers returned from a Science Citation database search for different keywords. 2001 data only represents three-quarters of the year. ....................................................................................................................................16

Figure 4: The Psychomotor Vigilance Task (Dinges and Powell, 1985). A screenshot of the PVT in action..........................................................29

Figure 5: The Multi-Attribute Task Battery (Comstock and Arnegard, 1992). A screenshot of the MATB near the beginning of the task. Dial F1 is out of place (it is greater than one notch from centre), light F6 requires action and pump 3 is broken.................................................................30

Figure 6: The X-Ray Object Recognition Task (Schwaninger, Michel and Bolting, 2007). Subjects only have two choices: ok, or not ok, and then rate their confidence. No initial position is set for confidence. In this example a gun is clearly visible and so the bag is “not ok.” .......................32

Figure 7: The final layout of the laboratory. This schematic shows the laboratory setup that was used in both studies. Separated by a large curtain (dotted blue line), subjects sat at each terminal with the television at 90 degrees in the periphery. Subjects could neither see nor hear the experimenters.........................................................37

Figure 8: Photograph of one of the experimental-side participant areas. The keyboard has the controls for the MATB highlighted, and there are two more cameras on the ceiling that are out of view.............................................38

Figure 9: The control side of the laboratory, which is split by dividers and a curtain. The beige blackout-blinds can be seen at the top-left of the picture, and the main computers are shown in the foreground. All experimental requirements, including EEG, were controlled from this vantage point.................................................................................39

Figure 10: The quad-screen monitor setting meant that the task, EEG recordings and behavioural measures were combined into one DVD recording. The experimental controls allowed each of these recordings and tasks to be started in synchronisation, and allowed later analysis of behavioural and task determinants of performance.........................................................40

Figure 11: Top: Average head turns towards the peripheral distraction, and below: the average number of lapses in each condition. From Anderson and Horne’s 2006 study.................................................................52

Figure 12: A diagrammatic representation of the study protocol........................................54

Figure 13: PVT mean reaction time and lapses across sessions. Sessions 1 and 2 took place between 10pm-12am, while sessions 3 and 4 took place between 4-6am.................................................................57
Figure 14: PVT mean reaction time for fastest and slowest reactions. Variability in results increases as fatigue and sleepiness accrue, although increases in the slowest reaction times becomes most noticeable.

Figure 15: Subjective sleepiness across trials from 10am-6am. Subjective sleepiness increases across the sessions, showing support that perceived sleepiness is the factor responsible for performance degradation.

Figure 16: Correlation between lapses and subjective sleepiness. A high correlation ($r^2 = 0.449$) is found between transformed lapses that reduce individual variance and worst performances (lapses).

Figure 17: Percentage difference in scores between alert and sleepy. The biggest differences with sleepiness are found in the slowest reaction times and lapses. The smallest is in the fastest reaction times and the median.

Figure 18: PVT head turn bins with stacked head turn frequencies. There is a larger preponderance of short distractions in the distraction condition than any other.

Figure 19: Percentage differences between the distraction levels. There is only a negligible change that does not exceed chance.

Figure 20: State instability for distraction and head-turn lapses. Peaks in distraction should coincide with peaks in head turn lapses occurring rhythmically to support state instability, and not arise from random variation.

Figure 21: Stacked column graph of PVT lapses in each of the four conditions.

Figure 22: Boxplots of the transformed lapses per condition. The interaction effect of sleepiness and distraction has a low effect size.

Figure 23: Lapse magnitude split by eyes open, eyes closed and head turns.

Figure 24: Graphical plot of lapse magnitude against percentage chance that eyes are open. The formula can be used to predict a percentage chance that eyes are open at any given reaction time.

Figure 25: The same graph as.

Figure 26: Comparison of global alpha and theta differences between distraction or sleepiness conditions. If a bandwidth increased in potential difference it was scored as “plus”, otherwise it was scored as “minus.”

Figure 27: Each individual frequency bin is shown as either increasing (“Plus”) or decreasing (“Minus”) for all subjects for the two EEG recording sites. An increase represents higher voltage in distraction conditions, whereas a decrease represents lower voltage brain activity recorded by the electrodes in distraction conditions compared to no distraction.

Figure 28: Each individual frequency bin is shown as either increasing (“Plus”) or decreasing (“Minus”) for all subjects for the two EEG recording sites. An increase represents higher potential in sleepy (morning) conditions, whereas a decrease represents lower potential brain activity recorded by the electrodes in sleepy conditions compared to non-sleepy conditions.

Figure 29: Global alpha and theta levels for the 3-minute baseline recordings for the C4-A1 and C3-A2 recording sites.
Figure 30: The dependent variables used in the MATB. Tank deviations from 2500 units (root-mean square) are combined for A and B. Dials are used for lapses (response not within 15s of deviation, e.g. F1 is out of synch in the figure) and reaction time, while the red and green lights are used for lapses and reaction time also.

Figure 31: Comparing dependent variables of the MATB for sleepiness. Lapses for the red light show the largest effect, but it is still a statistically insignificant effect of sleepiness. 50% represents no change between the alert (night) and sleepy (morning) conditions. FA = False Alarm.

Figure 32: Reaction times for dials & lights, and resource deviation by session e.g. session 1 would be the first MATB, session 2 is the second etc. The lack of change with increasing fatigue and sleepiness is shown for the dials and lights reaction times.

Figure 33: Time on task for the resource deviations for all four MATB trials, between alert and sleepy sessions. Resource deviations do not fluctuate significantly either during the task or with successive sessions.

Figure 34: MATB head turn bins with stacked head turn frequencies. Short distractions (<3sec) dominate during the task in distraction conditions, followed by long head turns (>3 sec) in distraction conditions.

Figure 35: Percentage changes with distraction in MATB components.

Figure 36: Absolute values for the three components in all conditions. The resource task is shown on the primary (left) axis whereas the lights and dials are shown on the secondary axis.

Figure 37: Collapsed alpha and theta bandwidths across all subjects, expressed as a simple non-parametric increase or decrease in activity from baseline. There is no case that can be made against the null hypothesis from these data.

Figure 38: Each frequency bin is expressed as either increasing (“Plus”) or decreasing (“Minus”) in value from baseline compared between distraction and no distraction conditions. The top figure shows the data from the C4-A1 derivation, whereas the bottom figure shows the data from C3-A2. More subjects experienced increased in activity in the lower frequencies (0-5Hz), as well as upper alpha (11-13Hz) during distraction conditions.

Figure 39: Count of increases or decreases in fast-Fourier transformed EEG results per individual frequency bin across the alert and sleepy sessions. Increases represent higher brain activity during sleepy conditions, whereas decreases indicate the opposite. C3-A2 shows trends towards increased alpha similar to the distraction findings, although the same cannot be said of C4-A1.

Figure 40: Global alpha and theta levels using only the 3-minute baseline recordings for C4-A1 and C3-A2. There are no perceptible differences between the levels when averaged for all participants.

Figure 41: A parametric scatterplot of alpha activity to the number of lights lapses. There is no clear conclusion that can be drawn between alpha magnitude and number of errors on the task.

Figure 42: The same data from Figure 41 but with a logarithmic transformation of the voltage.

Figure 43: Subjective sleepiness by performance lapses for PVT and MATB.
Figure 44: Head turns between the PVT and MATB conditions. Apart from long head turns in the distraction condition, the PVT shows much higher levels of distraction during the task than the MATB. .................................109

Figure 45: Effects of viewpoint (Schwaninger, 2006). A gun (a), switchblade (b) and shuriken (c) can look very different in a 2D x-ray image. .............................115

Figure 46: Study design for continuous wakefulness paradigm (Basner et al., 2008). T1-T7 are training bouts designed to accustomise subjects to the task. .................................................................117

Figure 47: Improvements in A' and RT for young and old groups. There is no loss in performance when subjects respond quicker to the task. In fact, substantial improvements in detection ability come with training. .............127

Figure 48: Speed-Accuracy Operating Curve for young and old groups. ..............................128

Figure 49: Stopping time policy for young and old, trained and untrained. ..........................129

Figure 50: Detection performance and speed differences in the young group. 
..................................................................................................................................................130

Figure 51: Detection performance and speed differences in the old group. ..........................131

Figure 52: Balakrishnan's d values for each condition and age group. 
Dramatic differences between the young and old groups are seen when reaction time and accuracy are joined, as well as for distraction. .................131

Figure 53: Stopping policies for no distraction and distraction by age. A marginal shift is evidenced between no distraction reactions and cognitive distraction, but the relationship between threat items and non threat items remains similar as shown by the likeness in slopes. .......................132

Figure 54: Predictions of performance ranks across the conditions to alpha level rank. The black line represents chance, with any results below it supporting the null hypothesis. None of the conditions are sufficiently above chance to suggest a link between alpha levels and performance. 
..................................................................................................................................................139

Figure 55: Predictions of performance ranks across the conditions to theta level rank. The black line represents chance, with any results below it supporting the null hypothesis. None of the conditions are sufficiently above chance to suggest a link between alpha levels and performance. 
..................................................................................................................................................140

Figure 56: Brain power at different regions during normal sleep. Almost across the board the older age group are showing heightened brain region activity than the younger group. .................................................................141

Figure 57: Brain power at different regions during sleep restriction. A similar result is found to that shown in normal sleep, where older subjects have higher brain activity than younger subjects. .................................................................141
CHAPTER ONE

Literature Review
1.1.1 Research Agenda

Sleepiness throughout the diurnal cycle has been studied exhaustively using a wide mixture of methods and apparatus to measure performance, as evidenced in many major reviews in this area (e.g. Dijk, Duffy and Czeisler, 1992). How an individual views their performance under sleepy conditions is poorly related to the reality of their abilities (Murphy, Richard, Masaki and Segalowitz, 2006), often resulting in latent danger to themselves and others that sometimes can only be known retrospectively in the aftermath of a serious performance failure. Nowhere is this more obvious than where humans are the operators of technical machinery, automobiles, or in monotonous observation; all tasks that are impacted by sleepiness, and roles which humans are poorly suited for.

Studying this area of research involves a recognition that in between large periods of safe activity, it may only take a momentary lapse in concentration to result in increased threat or fatalities. Other times, performance gradually deteriorates with time on task; accelerated by prior sleepiness and demand to work at the circadian nadir. Alongside the study of sleepiness, there exists various irrelevant variables impacting on performance that can be explored independently of the main effect of sleepiness, which can often explain some variability in results. In this instance I have focused on the poorly-researched irrelevant variable distraction. It is routine practice for researchers to try to eliminate distraction effects in their experimental designs, but rarely in life are human operators situated in completely sterile environments. Most of what is known about distraction in an applied research setting comes from the driver sleepiness literature, which is rich with examples of performance being affected by unpredictable task-irrelevant events. Indeed, performance can sometimes be enhanced by seemingly distracting actions such as putting the radio on, or chatting to a colleague; but the dangers arise when attention is diverted to a deleterious effect, as is seen all too often in police accident files, driving studies and control plant accidents.

Ballard (1996), in a review of the sleep literature, isolated three factors that affected performance in vigilance tasks: task parameters (e.g. novelty), environmental factors (e.g. distraction), and subject characteristics (e.g. personality). Throughout this document these three themes will reoccur frequently and in many guises, forming a thorough dissection of Ballard's summation. Distraction on its own hasn't been satisfactorily defined (Pettit, Burnett, and Stevens, 2005) although the best available definition arises from Stedman's Medical Dictionary (1995), where distraction is "a condition or state of mind in which the attention is diverted from an original focus or interest."
Other definitions have been offered by various researchers (Table 1), but they fail to capture the term in comparable parlance to each other. Sleepiness hereafter is operationally defined as behavioural change characteristic of drowsiness resultant from neurobiological processes regulating circadian rhythms and the drive to sleep. All other technical terms and abbreviations that are not expressly defined in the text can be found in the Glossary, where the first usage of the defined term is emphasised with italics in the main text.
The two-tailed hypotheses examined in this thesis can be summarised in a short list:

- Does sleepiness interact with distraction?
- Does distraction alter performance in monotonous or repetitive tasks?
- Is sleepiness or distraction differentially affected by task complexity?
- Is there an age effect in ecological (job simulated) tasks, and if so does it vary with sleepiness or distraction?

All of the studies presented throughout this thesis have been vetted to exclude studies that have used populations that are not relevant to the current research e.g. patients suffering from anxiety disorders, elderly subjects who suffer various sleep complaints, child studies etc. Resultantly, this has eliminated a large body of work for the benefit of a coherent and directly relevant presentation of related research. The terms 'subject', 'participant' and 'volunteer' are used throughout this document interchangeably. This document is presented in accordance with the APA Publication Manual, Fifth Edition, although all references are presented in the Journal of Sleep Research (JSR) format e.g. the APA manual states full author listings on first mention, whereas JSR requires shortening (et al.) for all references with more than three authors.

I do not wish to enter the null-hypothesis significance testing debate, other than to echo Frick's (1996) sentiments that "among many it is taken as obvious that it should be abandoned" (p. 379). There are many prominent statisticians who have written excellent overviews of the current state of statistical inference in Psychology (e.g. Berger and Sellke, 1987; Cohen, 1994; Goodman and Royall, 1988) and comprehensive guidelines have been given to researchers on the reporting and interpretation of results (APA, 1994; Wilkinson, 1999). Accordingly, p-values are only given as either supporting ($p > 0.05$) or failing to support ($p < 0.05$) the null hypothesis. Effect sizes (generalised-eta squared) are calculated for each inferential result or given as the effect-size measure of correlation ($r$). Whether a result should be considered highly significant (e.g. $p < 0.01$) or insignificant (e.g. $p = 0.28$) is more appropriately interpreted from the effect size.
Table 1: Definitions of distraction in the literature (Roper and Juneja, 2006).

<table>
<thead>
<tr>
<th>Study</th>
<th>Definition Of Distractions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baron et al (1973); Glass</td>
<td>Distraction is a &quot;manipulation that taxes attentional capacity leading to the organism to make priorities, take cognitive shortcuts, and ignore certain stimuli and tasks&quot;</td>
</tr>
<tr>
<td>and Singer (1972)</td>
<td></td>
</tr>
<tr>
<td>Cohen (1980)</td>
<td>&quot;Uncontrollable, unpredictable stressors that produce information overload&quot;</td>
</tr>
<tr>
<td>Covey (1989)</td>
<td>&quot;Typically require immediate attention and insist on action&quot;</td>
</tr>
<tr>
<td>Corragio (1990)</td>
<td>&quot;Intermittent interruption – externally-generated, randomly occurring, discrete event that breaks continuity of cognitive focus on a primary task&quot;</td>
</tr>
</tbody>
</table>

The Dangers Of Sleepiness

1.1.2 The Costs Of Sleep-Related Accidents

Of the highest profile disasters attributable- in part- to sleep, it is held that errors of omission based on failure to attend to a preventable outcome are chiefly to blame (US Nuclear Regulatory Commission, 1979; Reason, 1990). Sleepiness itself has been proven to be a major factor in serious performance decrement resulting in large-scale industrial accidents (Mitler et al., 1988), as well as everyday accidents caused by increased automation of manual tasks (Dinges, 1995). Human error is accountable for between 30-90% of all serious accidents in industry, and some degree of human error can be traced back as the trigger in about 80% of all accidents (Reason, 1990). Extensive reviews into safety in anaesthetists report a similar figure weighing in around 70-80% of accidents in the operating theatre being caused by human error alone (Wilson, 1993; Williamson et al., 1993). A full review of studies looking at performance in residency training relative to sleep loss can be found in Veasey et al. (2002), who concluded that sleep loss and fatigue result in significant neurobehavioral impairments in healthy young adults.

Evidence towards sleepiness-related accidents can be assumed by examining frequency of accidents at a given time corrected for the probability of the event occurring (e.g. by adjusting for traffic density etc). Composite analyses have shown a clear time of day effect for sleepiness-related accidents that corresponds with nomograms of human alertness (e.g. Åkerstedt, 1995). A German review of 569 automobile accidents related to drowsiness showed two peaks similar to Figure 1 (Prokop and Prokop, 1955), later confirmed by several research groups (Lavie, Wollman, and Pollack, 1986; Pack et al., 1993). This characteristic of sleepiness is found for many biological phenomena in humans, showing differing levels of activity at different circadian time periods (Mitler et al., 1987). Clinical trials have also shown bimodal histograms of unintentional sleep episodes occurring at the zones of
vulnerability (Figure 1) outlined by Lavie (1991) which characterise the type of sleep-related incident associated with overnight shift work.

Accident statistics derived from the various fields of research can be combined to find an overall value, including such fields as commercial driving (Cummings et al., 2001; Lal and Craig, 2001), supervisory control in plant work (Andorre-Gruet et al., 1998), and the medical profession (Jha, Duncan, and Bates, 2001; Weinger, 1999). Analyses of all sleepiness-related accidents estimate a cost between $43.15 to $56.02 billion per annum in America using data accumulated in 1988 (Leger, 1994). This finds accordance with the large body of driver sleep research, where the human cost of sleepiness is estimated at around 1,500 lives per year in the US (National Transportation Safety Board, 1990). However the statistics are derived, it is clear that a loss of performance caused- in part, or wholly- by sleepiness has grave financial and human repercussions. One group who are particularly vulnerable are those whose job descriptions require them to work at unnatural hours or for prolonged periods of time, such as those who work shifts.

**SLEEPINESS**

![Figure 1: Unintentional sleep episodes by time (Carskadon, Littell and Dement, 1985). A count of unintended sleep episodes aggregated from two studies show the two major peaks in sleepiness.](image)

1.1.3 Losing Sleep

Several prominent researchers have made the case that the working populace suffer the effects of sleep loss on a routine basis (e.g. Drake et al., 2001; Roehrs et al., 1996). Epidemiological studies find an average sleep length of between 6-7hrs during the week (Breslau et al., 1997), which is notably less than the majority of people would ideally sleep (Roehrs et al., 1989). In a sample of more than one million Americans, Kripke et al. (2002) found that around a fifth of the sample reported sleeping less than 6.5hrs per evening. Regarding exactly how much sleep the working population needs to function normally, the idea that people now work with a sleep debt is not new. The term "sleep debt" was coined more than forty years ago by Nathaniel
Kleitman (1963) to describe how delayed sleep onset, when coupled with habitual rising time, results in a sleep deficiency. "Sleep debt" as we know it is widely used to describe the effects of sleep loss regardless of the reason for the missing sleep (Van Dongen, Rogers, and Dinges, 2003). When alarm clocks and other modern time cues are removed, subjects will happily sleep up to and beyond eight hours (Bonnet and Arand, 1995). A recent population study of the UK public found that more than half of the respondents would desire more sleep, irrespective of how much it was (Anderson and Horne, 2008).

Yet there remains a wide difference in perceived sleep need that varies from individual to individual. Habitually long and short sleepers can be gleaned from a population as a related taxonomy (Klerman and Dijk, 2005), where habitual long sleepers are hypothesised to have a lower homeostatic sleep pressure than short sleepers (Aeschbach et al., 2001) and receive less benefit from increased time in bed (Rosenthal et al., 1993). Inadequate recovery sleep and build-up of sleep pressure are key features in the outcome of sleep debt (Dement and Vaughan, 1999), but sleep debt can only be reliably expressed as a function of each individual's actual or perceived sleep need. Aggregated data between subjects therefore lacks power to explain habitual sleep patterns.

1.1.4 Occupational Vigilance

Occupational studies, while more difficult to initiate and to eliminate confounding variables, give an ecological account of a research area that compliments findings in the laboratory. Bjerner et al. discovered in 1955, over a twenty year period in a Swedish gas works, a pronounced increase in reading errors during the night shift and again at what would later be described as the 'post-lunch dip' (Blake, 1967). Complimentary to these findings, subjective sleep quality in the week preceding an accident has been associated with a lowered risk of an accident in a case-control analysis conducted in a recent study by Edmonds and Vincent (2007). Workers with difficulty initiating sleep, insomnia, insufficient sleep and poor sleep quality were found to be more at risk of occupational injury than those without, even after adjusting for multiple confounders (Nakata et al., 2005). Medical interns who regularly work shifts lasting over 24 hours made 35.9% more serious medical errors than their compatriots who worked shorter shifts (Landrigan et al., 2004).

Productivity can be seen as the antithesis of potential danger, reinforcing the idea that nonfatigued workers work better and safer than fatigued workers. As described in Harrington (2001), the classic study relating to occupational vigilance was conducted back in 1920 by Vernon and colleagues on munitions workers during the First World War. Reducing hours of work to between 7-20 per week from 50-55 elicited improvements in quality and quantity of units produced, while ten minute rest breaks in the morning and afternoon increased production between 5-12%. A similar study requiring prolonged vigilance on a monotonous task by Vidacek et al. (1986) produced a figure of around 5% less productivity in the night shift workers than day shift for
producing capacitors. In their review Folkard and Tucker (2003) averaged the results of three typical studies in this field to graphically show the effect that working during certain hours has on work efficiency (Figure 2).

![Figure 2: Z-Scores of performance efficiency during a full day (Folkard and Tucker, 2003, p. 96).](image)

**Figure 2:** Z-Scores of performance efficiency during a full day (Folkard and Tucker, 2003, p. 96). Z-scores of performance efficiency, as measured by reading errors and excessive time to manage simple tasks, during the 24-hour cycle.

### 1.1.5 Shift Work

Shift work compounds operator performance by negatively impacting on sleep quality and quantity (Åkerstedt and Torsvall, 1981), often requiring individuals to work at the circadian nadir where the biological drive for sleep is at its highest (Lamond *et al.*, 2003). It can take days or several weeks of consecutive night shifts to re-entrain the biological rhythm (Sack, Blood and Lewy, 1992) and may never reach adaptation due to rotational shifts or returning to night sleep on days off or weekends. In conjunction with this, daytime sleep is compromised and becomes shorter and of poorer quality (Åkerstedt, 2003), as well as being disrupted due to various rotational methods often put in place to counterbalance shift turns (Knauth, 1995). These shift workers can become the passive victims of difficult occupational demands without adequate training or preparation, and it is not uncommon for workers during the day to seek a second job in the evening to improve income (known as ‘moonlighting’; Sussman, 1998).

Workload, work duration and work intensity are often cited as primary concerns when deciding to implement an eight or twelve hour shift (Duchon and Smith, 1993; Smith and Folkard, 1998). Extending work duration by four hours can increase operational errors, although productivity tends to remain the same (Williamson, Gower, and Clarke, 1994). A study in shift roster change from eight to twelve hours at a power station showed that performance deteriorates rapidly towards the end of a twelve hour shift (Mitchell and Williamson, 2000), with the authors claiming that the six month study was considered ‘short’ for evaluating performance impairments. What is
certain is that performance decrements are routinely observed in overnight shift work (e.g. Rosa and Bonnet, 1993), although the magnitude of these effects are difficult to compare across work places and may not manifest themselves in an easily quantifiable performance impairment.

Long-haul and commercial vehicle drivers are a risk category that have been increasingly regulated due to the risks posed by their lifestyle tendencies and potential for fatal or costly accidents (Lyznicki et al., 1998). Truckers and lorry drivers are more likely to suffer from sleep apnea, chronic restrictive pulmonary disease, diabetes and other disorders commonly associated with obesity (Stoohs et al., 1994). Of particular relevance are the conditions of sleep-disordered breathing (SDB), that create an inability to stay in the deeper, more restorative stages of sleep due to constant awakenings throughout the night as the airways collapse during REM sleep (Marshall, Gander and Neill, 2003). Compounding this are the demands to drive during the night and the unsatisfying sleep that can occur when trying to sleep next to a busy road in a small cabin. Drivers suffering SDB may be unaware of the effect their excessive daytime sleepiness has on their ability to concentrate or may be ignorant to the real possibly of falling asleep at the wheel (George et al., 1987). Reviews of the driver literature demonstrate results that strongly associate SDB and time of night as the two main explanations for sleep-related motorvehicle accidents (Lyznicki et al., 1998); important considering that driving is the leading cause of injury and death in America (National Safety Council, 1997). In actuality, 96% of sleep-related crashes involve drivers of passenger vehicles rather than commercial vehicle drivers (Knipling and Wang, 1994), but this statistic is not corrected for the differences in amount of passenger vehicle drivers to commercial vehicle drivers.

More worryingly, epidemiological data and anecdotal evidence show that shift workers often experience actual sleep episodes on the job (Åkerstedt, Fredlund, and Jansson, 2002; Samel, Wegman, and Vejvoda, 1995). Workers who begin work in the early hours of the morning before 6am generally receive less than six hours of sleep per day (Kecklund and Åkerstedt, 1995), due to multitudinous social factors. On top of all this, shift workers suffer increased psychological and psychophysiological ill effects than those who work routine daytime hours (Harrington, 2001). This results in a demographic who are vulnerable to the effects of impaired performance through depreciated sleep benefit, causing an unquantifiable burden on operational and societal welfare.

1.1.6 Fluctuations In Alertness

Human performance waxes and wanes throughout the diurnal period as circadian factors interact with the sleep regulatory homeostat (Cajochen, Blatter and Wallach, 2004; Rogers, Dorian and Dinges, 2003) elevating sleepiness at certain times of the day and night, also impacted on by a sleep inertia factor in other models (Åkerstedt and Folkard, 1997). When continuously sleep deprived a dose-response relationship is found between performance and hours of lost sleep (Belenky et al, 2003). With these
fluctuations it is shown that human performance is irregular and greatly affected by various processes that impinge on the individual.

Studies utilising a partial sleep deprivation methodology show how repeated loss of sleep over successive days can be more damaging to performance ability and well-being than if the same quantity of sleep is lost in one bout (Drake et al., 2001). In a study of sleep loss over fourteen days, the 4hr sleep period group showed significantly greater deficits than cumulative sleep loss over three days in the total sleep deprivation group (Van Dongen et al., 2003). The authors conclude that “even relatively moderate sleep restriction- if sustained night after night- can seriously impair waking neurobehavioral functions in healthy young adults” (p. 124). Other studies confirm that partial sleep deprivation has considerably larger negative effects than short-term (45hr) or long-term (>45hr) total sleep deprivation (Ferrara and Gennaro, 2001).

An exhaustive review of the literature lists assorted studies showing a deterioration of perceptual and oculomotor function under extended wakefulness protocols (Russo et al., 2005), as well as decreases in visual vigilance, visual acuity, visual detection and visual scanning. Compensatory effort to remain awake and perform at or above baseline increases (Dinges et al., 1992) as participants develop adaptation to chronic sleep restriction over time (Van Dongen et al., 2003). Subjects in this latter study were generally unaware of their increased cognitive deficiencies, posing some difficult questions over human suitability to self-monitor performance. But not all data corroborate in this area, such as Blagrove, Alexander and Horne (1995), who reported in a four-week sleep-reduction study that reducing sleep to 5.2hr per night produced no change in logical reasoning. Similarly, Totterdell et al. (1994) argue that sleep curtailment results in large negative effects on mood, which in turn could impact on motivation and desire to perform at optimum level. In a meta-analysis of the literature Pilcher and Huffcut (1996) came to the conclusion that mood is affected more by sleep deprivation than cognitive function.

Other studies have shown, by contrast, that participants in laboratory trials are capable of knowing when their sleepiness has reached a point where it could impair performance (Horne and Reyner, 1995), but that the interaction between homeostatic and circadian processes is nonlinear in nature and at present impossible to predict to a precise phasic performance impairment (Sturm and Willmes, 2001; Van Dongen and Dinges, 2003). Drivers are seemingly unable to gauge their own fitness to drive while sleepy (Brown, 1997; Hartley, 1995), and may choose to drive under conditions of high sleepiness which results in sleep attacks that many are unaware can occur. While the laboratory studies show that subjects know they are dangerously sleepy prior to a sleep-related incident, in the real world people may choose to ignore their own insight and get behind the wheel irregardless.
1.1.7 Assessing Sleepiness

Assessing sleepiness takes two forms: subjective measures, and objective measures. Certain measures are more practical in an applied setting, such as introspective measures like subjective scales, while others are firmly rooted to the laboratory until advances in technology and reductions in cost are achieved (e.g. EEG, pupillometry etc). Sleepiness scales such as The Stanford Sleepiness Scale (SSS; Hoddes et al., 1973) force the individual to consider whether they are fit to continue any task, although they rely heavily on the individual’s honesty and whether they are compelled to act on that knowledge. Ogilvie et al. (1989) found that increases in the scores found on the SSS correlated highly with the onset of Stage 1 sleep, so when individuals reach a certain level then the appropriate advice should be followed as a countermeasure to this decline in alertness. Changes in SSS self-reports also correlate to actual performance (Rosa et al., 1985), showing that these scales do relate directly to the phenomena they are supposed to represent. There is ample evidence throughout the literature that short naps are an effective countermeasure to fatigue for shiftworkers (Purnell, Feyer and Herbison, 2002), as well as caffeine and other strategies (Reyner and Home, 2000). The problem lies in trying to quantify when humans begin to be at risk from the effects of sleepiness; a conundrum at the core of this Chapter's research.

Complementary to subjective ratings of sleepiness, there are a myriad of objective approaches that use practically every accessible organ to predict sleepiness (galvanic skin response, brain activity, pulmonary respiratory rate, heart rate, eye blink rate...). Of all the available metrics besides those that tap into brain activity changes, it is the eye that provides the truest measure of sleepiness and continues to tantalise even during sleep (e.g. Rapid Eye Movements; Maquet et al., 1996). Slow eye movements (SEMs) occur under the eyelids when sleep is in the early stages (Aserinsky and Kleitman, 1955), and it is these actions that often filtrate throughout wakefulness under periods of sleep loss. Long eye closures during eye blinks are one of the best predictors of sleepiness (Stern, 1980), with Stern reporting that “the oculomotor control system is exquisitely sensitive to fatigue and attentional parameters” (p.1). As humans become increasingly fatigued the frequency of eye blinks begins to slow down (Boelhouwer and Brunia, 1977), as well as the duration of the blink itself slowing down as the brain attempts to forcefully induce sleep. Even when a full eye closure isn’t recorded, the preceding step-percentage eye closure- has been posed as an ‘early-warning system’ for drowsy truck drivers (Hayami et al., 2002).

‘Microsleeps’ are a sleep-related phenomenon that renders an individual unable to gather information or respond to task requirements for a short period of time, typically lasting from a few seconds to fifteen seconds (Tirunahari et al., 2003). Microsleeps typically take place during predominance of the theta electroencephalographic (4-8Hz) rhythm emerging from a background of alpha without eye blinks, lasting at least three seconds (Priest et al., 2001). These episodes of sleep intrusion limits the perception of external stimuli, and often subjects are described as having a glazed
expression much like daydreaming. These short periods of inaction have been attributed to sleep-related motorvehicle accidents, as a study by Corfitsen (1999) dramatically showed when eighteen of his twenty-six person sample experienced microsleeps before a full-blown sleep episode at the wheel. The onset of microsleeps is an indicator of oncoming sleep with epidemiological studies reporting incidences of previous microsleeps being above 40% in normal night-time drivers (Corfitsen, 1999). Driver simulator studies also confirm that performance measures like steering variation and lane drift are worsened during the periods scored as microsleeps (Paul et al., 2005), although this depends on the microsleep coinciding with periods of driving requiring immediate attention.

Most subjective scales and objective measures are tested against the ability to fall asleep as an objective measure of true sleepiness (known as the 'Multiple Sleep Latency Test'; see Johns, 1998). The opposite ability, that of staying awake (the 'Maintenance of Wakefulness Test'; Mitler, Gujavarty and Browman, 1982), is also used and really provides a more ecological account for applied purposes. These tests form a crucial part of the clinical assessment of sleep problems caused by excessive daytime sleepiness and have increased in importance since medicolegal guidelines have been published, where physicians are legally responsible for determining whether a patient is fit to operate motorvehicles (Canadian Medical Association, 2000). When performance failures do occur during a task it could always be caused by a reduction in motivation, or fatigue-effects that are not always related to sleep (Harrison, Bright and Horne, 1996). So while there are a range of methods available to predict sleepiness-related impairments, it is difficult to know how many and which measures cross-correlate with each other, and not just with the MSLT and MWT. Dinges (1989) was clear in his assessment that not all sleepiness measures are equally sensitive to experimental changes, and so these predictors need to be allayed to the task that they were examined with.

Distraction

1.2.1 Irrelevant And Confounding Variables

In typical laboratory setups the effects of distraction or competing stimuli are experimentally controlled to reduce confounding variables, or removed altogether to eliminate the irrelevant variable. The only problem with this approach is that real-life performance is rarely undertaken in a noise-proof or visually sterile environment, meaning that real-life performance invariably differs from the experimental protocol. To investigate distraction, all other variables need to be held as constant as possible and the distraction intervention applied systematically to all participants. While ecological distractions are generally unpredictable and free to vary in signal strength and duration, replicating this would not provide a controlled environment. Studies that manipulate distraction tend to fall into two categories: those that use distractions interspersed with target stimuli (i.e. Hogervorst et al., 1998), and those that use extraneous distracters to simulate typical lifelike scenarios (e.g.
Sleepiness and Distraction 12

Stutts et al., 2001). Both of these methods examine the same basic principles behind shared and divided attention, differing only in locality of presentation.

1.2.2 Distraction And Divided Attention

Forster and Lavie (2007) have recently authored a provocative paper concluding that high distraction-prone subjects (based on the Cognitive Failures Questionnaire; Broadbent et al., 1982) become as distraction-prone as low-distractibility subjects in conditions of high perceptual load. This makes intuitive sense as situations with high attentional load require subjects to remain focused to perform adequately, a concept that we have termed "engagement" that is grounded in attentional narrowing theory. Attentional narrowing theory explains how human perception manages to focus attention while still being attuned to a wider field of vision, presuming that humans have limited attentional resources, and that as perceptual load increases the influence of distracters typically decreases (Lavie, 2005; Madden and Langley, 2003). It is hypothesised that attentional narrowing effectively excludes stimuli outside of the spatial attention window, although when relevant and irrelevant information pertains to the same object attentional narrowing increases distracter processing (Chen, 2003). It is interesting to note that observers are slower, on average, when distracters are spatially close to the target than when they are far apart (Chen, 2003).

However, it should be noted that distracters can often be used to mean irrelevant stimuli that participants are requested to ignore, and not always peripheral events easily distinguishable from the task demands. Often distraction can be interpreted to include any enjoyable peripheral activity that is used to ameliorate sustained performance. For example, fatigued subjects in simulated anaesthetic tasks employ a wide range of sleepiness-reducing strategies such as conversation, music and busy-work to counteract their long and boring duties (Weinger, 1999). Gillberg et al. (2003) found that short breaks during tasks temporarily decreased sleepiness, attributable to social contacts with the experimenters rather than any effect of coffee which is typically short-lived (Reyner and Home, 2000). Anecdotal evidence would appear complimentary to this notion, as people may opt to put on music or television as a peripheral distraction from long driving journeys or during work.

Studies utilising total sleep deprivation methods suggest participants develop an inability to switch attention between two sub tasks (Fisher, 1980), which would specifically refer to an endogenous component prior to the task and an exogenous component afterwards (Rogers and Monsell, 1995 ; Pashler, 2000). The exogenous component relates to stimuli that evokes a tendency to perform an action habitually associated with the task that limits the likelihood of switching attention when such an action is desirable. It may be that a supplementary ability is altered rather than gross ability to switch attention, such as an impaired ability to attend to novel stimuli resultant from the subtask (McCarthy and Waters, 1997). In a review, Krueger (1989) surmised that workers “change their patterns of attention with prolonged work” (p 132) from the literature of skill performance studies, which would indicate a coping strategy to counteract effects of fatigue or sleepiness.
1.2.3 Measuring Distraction

Experimental distraction itself is poorly understood outside of a cognitive paradigm. Related to eye movements, certain researchers conclude that "suppressing reflexive eye movements does not reduce external control over fixation" (Machado and Rafal, 2000, p. 73), meaning involuntary eye-movements are affected but not voluntary eye movements. Initiation of movement, direction of gaze, and target of attention are all voluntarily determined (Russo et al., 2003, p. 734) and impacted by several brain regions during sleep deprivation (Thomas et al., 1998; Thomas et al., 2000). There exists therefore a difficulty in differentiating which eye movements may be reflexive or voluntary, and how sleepiness affects both. It is thought that information is obtained during fixation and the location of the next fixation is based on the information obtained during the previous fixation (Lui, Gale and Song, 2007). These contrast to saccades, which are very fast eye movements that take place in between fixations.

Many studies now use eye-tracking technology to ascertain where attention is focused (Duchowski, 2002; Jacob and Karn, 2003). These can deduce when gaze has been diverted from the screen or area of interest, but they require cumbersome equipment to operate correctly. It is difficult to determine whether a gaze beyond the field of view represents distraction using automated algorithms, so for this reason the focus of our studies uses manual inspection of video recordings to deduce distraction in Chapters Three and Four. In addition to behavioural distraction that formed the focus of the first study, there is a cognitive element to consider too. The effects of cognitive distraction can be seen and measured as a performance decrement (Chapter Five). To this end, neurological studies have attempted to reconcile the brain regions that are responsible for reflexive eye movements.

Neurons in Brodmann's area 8, which include the frontal eye fields, fire selectively to saccade to targets in the visual field, where neural plasticity arises from experience-dependant voluntary shifts of gaze found in primate studies (Bichot and Shall, 1999). Even when target stimuli were used as distracter in subsequent trials the neurons of this area showed a pattern of learning from experience. Similarly, using saccadic eye movements as a behavioural indicator of sleepiness has been suggested (Caldwell et al., 2003), while other eye movements have been verified to correspond to lapses in attention such as slow eye closures and percentage of eye closure (PERCLOS; Dinges et al., 1998). Relating these lapses to behavioural distraction has not been investigated until now, albeit with head turns rather than an impractical voluntary/involuntary eye movement scoring system.
1.2.4 Driver Distraction

Of the applied distraction research studies, by far the most numerous are driver distraction studies. Driver distraction has been implicated in 13.2% of Crashworthy Data System car crashes throughout 1995 (Wang, Knipling and Goodman, 1996), while the broader definition of driver inattention results in anything from a quarter of all traffic accidents (National Highway Traffic Safety Administration, 1997) to between a third or half (Sussman et al., 1985). Component analysis of the NHTSA data indicate several outlets of distraction for the driver e.g. tuning the radio, looking at a billboard, being lost in thought etc that can lead to a crash. Due to the nature of the NHTSA data these figures are probably conservative as drivers do not wish to implicate themselves in an accident, resulting in significant biases (Stutts et al., 2001).

In a naturalistic study examining 121 hours of long-haul lorry drivers Barr, Yang and Ranney (2003) concluded that a total of 4,329 distracting events were logged representing 52% of overall driving time. Driving behaviour has been likened to an attentional field known as the central (or foveal) field of vision, representing the central task of vehicle alignment limited by outward constraints such as lane markers (Janelle and Singer, 1999). When looking purely at eye glance data using protocol outlined by the Society of Automotive Engineers (1999), drivers spent over 80% of the time looking at the road ahead when not engaged in distracting activity. By comparison young persons tend to be most at risk from distraction during driving (Stutts et al., 2001), especially given the ever-growing rate of in-car technologies (Tijerina, 2000).

Bunn et al. (2005) performed a retrospective analysis of accidents in a Kentucky accident report database to determine the relative preponderance of sleep and distraction-related fatal accidents. They found that distraction/inattention and falling asleep or fatigue were strongly associated with fatal commercial motor vehicle collision; findings that mirror Pratt’s study (2003) which found that 15-33% of fatal truck crashes were due to driver fatigue. In a driver simulator study looking at mobile phone use while driving, Aim and Nilsson (1994) reported small effects of using a mobile phone while driving but conceding that using the mobile phone led to increased workload. That the increased workload only affected driver lateral position was due to subjects slowing down to compensate; a strategy predicted by speed-accuracy operating curves (Drury, 1994) that proves problematic in self-paced tasks.

Applied Visual Inspection

1.3.1 Visual Search Theory

The human visual perception system is adept at rapidly directing gaze towards areas of interest; a process that has a long evolutionary history (Treisman and Gelade, 1980). Visual search is commonly viewed as being a two-stage process (Hoffman, 1979; Swensson, 1980), where the first stage is rapid and global, serially linked to a focal attention stage, which is considered
and slower. A pre-attentive filter blocks out irrelevant information and attention is drawn towards the most visually prominent areas; an idea that finds its roots in Broadbent's eponymous 'filter theory' (1958). During this first attentive stage pop-out can occur, which is a bottom-up process based on the simple local properties of the image area. The second stage of the model occurs in an iterative manner, where parts of the image or scene are scrutinised until evidence accumulation exceeds a functional threshold, or search is restarted into a different location. Itti and Koch (2001) argue that visually salient areas are attended to first based on conspicuity, with an "inhibition of return" feedback modulation mechanism in place to stop attention returning to already-examined areas.

Most modern interpretations emphasise comparison of the image presented to stored items in visual knowledge (Treisman, 1998). The second stage of the model is driven by matching internal representations of known target exemplars stored in memory to that found in the image space (Graf et al., 2002; Schwaninger, 2005). This is important as it reinforces the reasons why training should improve performance when workers or subjects become adept with visual search but remain naïve to a number of potential stimuli (Hofer and Schwaninger, 2005). Another model derived from the central theorem suggests that focused areas of interest tend to be the most informative for disambiguating identity (Schill et al., 2001). Studies such as these have been used to understand what makes an object difficult to find and form the framework for which many of the concepts that are discussed later depend on (e.g. Schwaninger's image based factors, Drury's model etc).

1.3.2 From A Radiological Background

The first study to explicitly look at x-ray screening performance is dated as far back as 1947 (Birkelo, Chamberlain and Phelps, 1947; as reported by Manning, Gale and Krupinsky, 2005). More recent studies have shown that performance is based on expertise, with experienced radiologists being able to fixate onto target areas of a mammogram quicker than interns (Nodine et al., 2002). On top of being quicker to fixate, experienced medical image inspectors are more likely to spot abnormalities (Sowden, Davies, and Roling, 2000), again supporting the idea that visual knowledge of target items is important for training.

Until the terrorist attacks of September 11th 2001, the focus on x-ray screening studies lay mainly in the radiological domain (Lui et al., 2006). Figure 3 shows the rapid increase in publications from 1981 up to the attacks of 2001 that have investigated visual search ability. Radiologists need to be alert and highly trained to be able to spot the emergence of cancers as there are 1.5million women screened each year in the United Kingdom alone, returning around 10,000 cancer diagnoses (Gale, Purdy and Wooding, 2005). Between 30-40% of mistakes made in clinical radiology are false negatives (Kundel, 2004), meaning that actual abnormalities are missed and are therefore considered to be the most serious errors. The economical and sociological importance of radiological research cannot be overstated.
1.3.3 X-Ray Luggage Search

Due to the similar nature of radiological and luggage screening tasks there is an overlap in literatures, with the same basic ideas being applicable to both. According to Schwaninger, Hardmeier and Hofer (2004) the general visual cognition model can be split into two aspects: knowledge-based factors and image-based factors. These two factors are necessary to understand how screeners manage to successfully match the stimulus representation with the visual memory representation (Hardmeier, Hofer and Schwaninger, 2006). Knowledge-based factors reflect the previously discussed notion that screeners need to know what to look for, being a function of training and memory retention. Prohibited items such as tear gas canisters, semtex and improvised explosive devices are not likely to be encountered in normal daily life (Schwaninger, Hardmeier and Hofer, 2005) and so have to be made familiar through training. As “terrorist threats are both productive and diverse,” (Lui, Gale and Song, 2007, p.301), potential plane hijackers tend to hide objects to the best of their ability, and may disassemble parts for later reconstruction to fool luggage screeners.

Secondly, there must logically be factors that make certain threats harder to detect than others. In other models, such as Itti and Koch’s saliency model (2001), target items are determined by their base characteristics. In Schwaninger, Hardmeier and Hofer’s (2004) three-component model, target items are determined specifically by the presentation of the threat item and its surroundings. As x-ray scanning machines used at airports only output images in two-dimensions, it means that target occlusion and rotation are randomly determined. The first factor, target rotation, refers to the way the
threat item is presented in $z$-coordinate space. Threat items are invariably harder to detect when they differ from the canonical view (Palmer, Rosch and Chase, 1981), which again is the product of a two-dimensional limitation. The second factor, superposition, refers to the degree that the target is occluded by other items in the bag, and the final factor is simply bag complexity. All three of these factors have been mathematically modelled (Schwaninger, Michel and Bolfing, 2007) with automatic prediction of image-based factor difficulty being as reliable as human estimates. However, McCarley, Kramer and Wickens (2004) question whether observers' improvements with ability to recognise targets will transfer when novel items are introduced.

There is a paucity of sleep studies in the x-ray luggage search literature due to the relative newness of the area, with only one study actually using an x-ray luggage search task. Tests of simple visual search ability have been proven to be immune to sleep deprivation for periods up to 28 hours using a letter-search task (Williamson et al., 2001). This is a typical result for tests that are more engaging than the usual performance tests used throughout sleep research, and this topic is discussed in greater detail in Chapter Four. Basner et al. (2008) found during a 34-hour constant routine paradigm that subjects were less able to detect threats and made more false-positive errors attributable to fatigue than when they were rested (during the first 11 hours). Hit rate, false alarm rate and bias also decreased with time on task, but looking at data in this manner contravenes the signal detection theory (SDT; Green and Swets, 1966) these measures are based on. It is theoretically unsound to split performance during a task as SDT measures assume that the operator is equally likely to detect an item unless the signal properties themselves change. Thus, to split data into segments from a test would require proof that ability to detect items had been systematically altered between the bins and not within to explain differences.

Investigating Sleep In A Controlled Environment

1.4.1 Task Durations

The longer participants are exposed to a task, the higher the likelihood that sleepiness will affect performance (review: Dinges and Kribbs, 1991), which can occur early in a task (Thiffault and Bergeron, 2003). Several studies show a linear trend between task duration and length of sleep deprivation affecting performance in conjunction (reaction time task: Lorenzo, 1995; MATB: Caldwell, Ramspott and Gardner, 1998). Further to this, in a seminal paper written by Gillberg and Åkerstedt (1998), it was concluded that monotonous tasks exhibit no 'safe' duration when operators are sleepy, with more than half of missed signals being attributable to electroencephalographically-defined sleep. Describing tasks vulnerable to sleep loss, Gillberg and Åkerstedt list four main characteristics: duration, monotony, extent to which the task is paced, and signal uncertainty.

Due to time and financial limitations it is often not practical to test subjects for extended periods of time that would accurately mirror typical working hours. Thus, guideline task durations are hotly contested with many studies settling
Sleepiness and Distraction 18

on tasks that only last ten minutes or even less (e.g. Lamond et al., 2003; Lisper and Kjellberg, 1972). Tasks as short as this surely misrepresent the effect of sleepiness, with studies using tracking tasks showing an interindividual range for first behavioural microsleep being in the order of 5.5-51.3min (mean 22.3min; Peiris et al., 2003; comparable to Thiffault and Bergeron (2003) who found a peak of fatigue between 20-25min of driving. Reliability suffers in short-duration tasks (Bonnet, 1989) and comes at the price of reduced sensitivity (Mullaney et al., 1983). The amount of practice that participants need to counteract learning effects is an obvious factor to include, which can be dissociated from circadian influences by testing participants in different orders across times of the day (Kryger, Roth and Dement, 2001).

A more suitable question would be to discover the minimum amount of time necessary to elucidate effects of sleepiness. Gillberg and Åkerstedt (1998) report historical studies by Wilkinson (1961), who by using a 40-min visual vigilance task found no performance loss in the first 5-10mins, and Donnell (1969) who concluded the minimum a task should last is half an hour to be sensitive to a night's sleep loss. Other researchers suggest an hour as the minimum for a vigilance task (Galinsky et al., 1993), which would logically have no detractions as the results can be analysed for the first ten minutes or half hour if required for cross-study concomitance. As a rule of thumb, when task duration increases impairment of performance increases (Purnell, Feyer, and Herbison, 2002), which implies that vocational performance may actually suffer more than laboratory performance.

1.4.2 Lapse Hypothesis

Throughout the psychology literature is the notion that a reaction time can be considered so poor that the subject failed to attend to the signal beyond merely being a poor reaction split (Williams, Lubin and Goodnow, 1959). These brief periods of non-responsiveness, often labelled "microsleeps" (Priest et al., 2001), are arbitrarily defined as any reaction time over 500ms in the PVT and any missed signal in the MATB (typically 15s duration dependant on scripting). Lapses have proven to be a popular method to quantify in simple terms a number that represents an amount of times an individual has failed to respond to a stimulus (Dijkman et al, 1997).

In early studies it was believed that performance decrement only occurred during these lapses with intermittent periods being unaffected (Lubin, 1967), although this theory has largely been disproved since (Dinges and Kribbs, 1991). Overall performance declines with sleepiness regardless of which metric is used (Jewett et al., 1999), with this study utilising dose-response curves to conclude that slow-wave sleep, subjective alertness and psychomotor vigilance declines asymptotically during sleep loss. An accelerated state instability is typical of monotonous performance tasks, where an oscillation occurs between alertness and impaired vigilance with accompanying drowsiness and emergence of microsleeps (Binks, Waters and Hurry, 1999). PVT performance changes during the nadir of the attention cycle brought on by sleepiness, reflected by an increased commingling of
errors of omission (lapses) and commission (false responses; Drummond et al., 2005). On this task, the difference between fastest and slowest reaction times coincide neatly with changes in homeostatic sleep pressure (Graw et al., 2004), suggesting an ultradian zone of vulnerability dependant on task novelty and interacting sleep factors.

Alternatively, a lapse could be compartmentalised into various categories of lapse (e.g. attention lapses or microsleeps; Frey, Badia and Wright, 2004) much like we have done behaviourally (see Section 2.2.4). Another method of performance measurement is to report the 90th percentile (Graw et al, 2004), which logically removes ambiguity regarding operationally defining what is a genuine performance lapse and what is simply poor performance relative to baseline. The central problem with counting lapses is that it can be biased towards a select few individuals with a disproportionate amount of lapses (Dinges et al., 1998), while those with little to no lapses are effectively omitted. Because of this, the current construct of a lapse is still a work in progress as it is impossible to deduce a behavioural response with any degree of accuracy by only examining raw data. Chapter Three deals with a proposed solution to this problem, and the first steps towards redefining the lapse.

Performance Changes With Age

1.5.1 Sleepiness Affects Older Subjects

In a recent review of ageing and performance, Philip et al. (2004) went as far as to describe age as having a "very negative effect on performances" (p. 109), citing worse experimental performances than baseline as confirmation for the validity of that statement in subjects aged 20-25 and 52-63 years old. Killgore, Balkin and Wesensten (2006) compliment this idea by stating plainly that "the older the volunteer, the greater the adverse effect of sleep deprivation on decision making" (p. 10). Cognitively demanding tasks are gradually more affected by sleep loss in older participants (Webb, 1995; 18-22 against 50-59 year olds), with EEG studies showing higher cortical arousal levels during a performance task but not at resting levels (Smulders et al., 1997; 18-24 and 62-73 year olds).

Contrary to this, there is a growing body of research that suggests older people suffer less from the adverse effects of sleep loss than younger people (Smulders et al., 1997), and are better capable of pacing their performance to match their abilities (Jarvis, 1993) as they work further below their ceiling than younger people do. And while older people inevitably do suffer worse performance on reaction time tests, their declines aren't as marked as with younger participants following sleep deprivation (Philip et al., 2004). Older participants maintain performance better irrespective of the amount of prior sleep obtained (Bliese, Wesensten and Balkin, 2006; range 24-64 years old). In a study looking at shift workers, Reinberg et al. (1980) found that although the older group (46-56 year olds) slept as long as the younger group (21-36 year olds), they reported much higher fatigue on waking and lower sleep efficiency.
1.5.2 Elderly Reactions

Reaction time is highly susceptible to age (Mani, Bedwell and Miller, 2005) and begins to show decrements as early as thirty years of age (Wilkinson and Allison, 1989) independent of gender. Sustained attention, selective attention, and inhibition performance all show decrements with age into adulthood (Mani, Bedwell and Miller, 2005), although older participants seem to cope with the damaging effects of sleep loss relative to baseline better than younger subjects (Bonnet, 1989; 55-70 year olds). In a study of ten young and ten old participants, the old group exhibited worse reaction times under normal conditions than young but managed to remain stable under sleep deprivation, while the young subjects' performances worsened (Philip et al., 2004). The largest effects on ageing on a reaction time test tend to be errors of omission (Smulders et al., 1997), which is similar to findings from choice reaction times and not evident under normal sleep conditions.

Podlesny and Dustman (1982) found increased electroencephalogram (EEG) P3 activity in elderly subjects (m = 71yrs) compared to young (m= 24yrs), which is a positive event-related wave that appears between 250-500ms after a stimuli is presented (Pritchard, 1981). Similar EEG studies using simple unprepared reaction times as dependent variable find that overall spectral power is suppressed as age increases (Polich, 1997), particularly in the alpha band (e.g. Ciganek, 1961). An fMRI study using subjects aged under 35yrs and older than 50yrs concluded that the older group recruited additional cortical and subcortical areas of their brain when performing a simple motor task, possibly reflecting compensation for structural and neurochemical changes (Mattay et al., 2002). Several studies show that older subjects either demonstrate greater activity levels in the same brain areas as younger subjects when attempting to perform the same task (e.g. D'Esposito et al., 1999), or use different brain regions altogether (e.g. Grady, 2000). These studies tend towards the conclusion that age effects are attributable to neurocognitive changes with age, although it should be stressed that these studies are not longitudinal and that “general dissatisfaction” (Botwinick, 1970, p. 241) arises when exploring ageing effects cross-sectionally.

Not all data are in accordance with these findings, with some studies finding no effect of ageing at all on reaction times (Härmä et al., 2006). Furthermore, trait analyses by Van Dongen and colleagues (2004) single out sleepiness, impatience and perceptions of stressful conditions as factors contributable to individual variation in performance and not age. It may simply be a cause of worsened sleep in the elderly, who are known to have earlier bed and wake times (Reyner and Home, 1995) with meta-analyses concluding that ageing detrimentally affects sleep maintenance and length (Floyd et al., 2000). Even during the day the elderly are more vulnerable to instability of circadian rhythms (Van Goold and Mirmiran, 1986), and performance ability varies greatly between older persons compared to between younger persons (West et al., 2002).
1.5.3 Elderly Distraction

Chao and Knight (1997), utilising two groups aged between 20-22 and 57-71 years, deduced that "increased distractibility and impaired sustained attention with aging may be due to altered prefrontal cortex function" (p. 63). These authors attribute the attentional system as enabling cognitive processing to focus on task demands, of which a portion includes ignoring or inhibiting irrelevant aspects, conceptualised as distractions. One of the early studies looking at this hypothesis was conducted by Rabbitt (1965), who found that older participants were less able to cope with task-irrelevant information; a finding mirrored over thirty years later by Li et al. (1998; 63-75 years old). Interestingly, some researchers have gone as far as to claim a diminished working memory leads to specific functional impairment from distractions (Hasher and Zacks, 1988), claiming that a "person with reduced inhibitory functioning can be expected to show more distractibility" (p. 215). Variants of the stroop task (Eriksen and Eriksen, 1974) confirm that older subjects suffer greater interference from incompatible distracters than younger subjects in the same condition (Scialfa and Kline, 1988).

In the appropriately titled paper “problems in testing the elderly,” Aiken (1980) issues guidelines explaining that older subjects are prone to distraction, possibly linked to lowered motivation. When asked to ignore irrelevant stimuli in ‘oddball’ and related bitalon paradigms, electrophysiological changes are found between young and old age groups (Amenedo and Diaz, 1998). The two processes in memory performance can be split into automatic and effortful (Hasher and Zacks, 1979), of which the latter is most susceptible to event-related potential (ERP) changes in older subjects (Amenedo and Diaz, 1998). Horváth et al (2007) found in an auditory-distraction study (19-24yrs and 62-82yrs old groups) that, along with longer reaction times and lower detection performance, P3 and RON (reorienting negativity) were uniformly delayed by over 100ms in the older group. The RON response occurs when attention is directed towards an unexpected task-irrelevant stimulus, which becomes unbalanced with increasing age and increasingly involuntary (Escera et al., 2000).

1.5.4 Visual Search Studies Using Older Subjects

Visual search performance is known to slow down with increasing age (Hartley, 1992; Madden, Gottlob and Allen, 1999) as well as detection accuracy decreases for target stimuli (Mani, Bedwell and Miller; 2004; McCarley et al., 2004). Errors of omission and commission (including false alarms) were found to increase in older subjects (range 19-82yrs; Mani, Bedwell and Miller, 2004) in a continuous performance task. Armstrong (1997) found that older subjects were significantly less able to respond correctly for auditory discrimination targets (age 60-94yrs) than a younger adult group. Older subjects also have more difficulty in discerning targets when they are stored amongst noise in the periphery of vision (Cerella, Plude and Milberg, 1987). Thus, the pattern of elderly performance deficits compared to young can be attributed mainly to problems with selective attention and inhibition of distracters, compounded by slower performance.
There are wide variations in the theories of why this occurs (Groth and Allen, 2000), such as the theory that with increased age comes a decrease in the energy that fuels cognitive processing (Groth and Alien, 2000). Alternative theories include Hasher and Zacks (1988), who suggest that older adults demonstrate reduced inhibitory functioning resulting in an increase in intrusions from distractions. Thirdly, impairments in spatial localisation may be the root cause of poor performance in visual search studies (Plude and Hoyer, 1985). Physical changes in eyesight and useful field of vision (Cerella, 1985) may simply be the cause of these decrements given that stimuli tend to only be shown for a limited time in most visual search methods, as well as a diminished inhibition of return mechanism that occurs 300ms later in elderly persons (Castel et al., 2003). Cerella (1991) proposed a 'general slowing model' to provocatively claim that a parsimonious model can explain all differences by simple cognitive slowing; a deduction that was rapidly and fiercely criticised (Fisk, Fisher and Rogers, 1992).

Neurobiology And The Electroencephalogram

1.6.1 The Pre-Frontal Cortex

Looking at the effects of experimental intervention on localised brain regions such as the PFC can be achieved by using expensive fMRI equipment for a blood-oxygen-level dependent response (BOLD), or it can be deduced from electroencephalographical data. One particular region of the brain that has been the subject of extensive research interest is the prefrontal cortex (PFC). The PFC is functionally interconnected with virtually all sensory neocortical and motor systems (Fuster, 1997); being uniquely elaborate in humans. Suppression of distraction is theorised to originate in the prefrontal cortex (Anderson and Horne, 2006), the evidence for which mainly derives from sustained attention paradigms. Vendrell and colleagues (1995) state in the opening line of their paper looking at lesional impairments of the PFC that "protection against interference from both inside and outside the organism has been described as one of the most important functions of the prefrontal cortex" (p. 341).

Much of neurobiological research into sleepiness and performance ability relates to the PFC (Muzur, Pace-Schott and Hobson, 2002), viewed as the area that enables generation and execution of novel goal-directed behaviour (Blatter et al., 2005) and the 'seat' of higher level processes that modulate lower level ones (Shallice and Burgess, 1996). Blatter and colleagues emphasise the usage of the term "novelty," as involvement of the PFC diminishes and moves towards lower sensory and motor cortical regions with increased practice (Fuster, 1997). Findings such as these have led to the popular PFC vulnerability hypothesis (Horne, 1993), although Carskadon and Roth (1991) in their review suggest that the studies reviewed suffered from a narrow focus on a small number of neurobehavioural outcomes; effectively oversimplifying what is a highly complex area of study. Miller (2000), in a review, stated plainly that "one of the classic signs of PFC damage is increased distractibility: subjects seem unable to focus on a task
when other, irrelevant events compete for their attention" (p. 62). Coupled with the impairment to the PFC described earlier in sleep deprivation studies, it seems logical to conclude that these two may be inter-related. Working memory is implicated in cognitive distraction models, where studies using primates show that normal working memory can retain information in lieu of distracting (non-essential) information (Miller, Erickson, and Desimone, 1996). However, with frontal-lobe lesions this ability can be lost (Petrides, 1991) and is indeed seen in elderly humans (Chao and Knight, 1997), who naturally exhibit diminished frontal lobe ability.

1.6.2 Electroencephalogram

The electroencephalogram (EEG) is the gold-standard physiological measure of sleep and sleepiness (Coenen, 1994), producing a summated time-segmented representation of voltage fluctuations occurring during synaptic excitations of the dendrites in large populations of pyramidal neurons in the cerebral cortex (Teplan, 2002). These weak signals are greatly amplified and stored for analysis, of which a popular method involves viewing the signal in terms of frequency against intensity by squaring the Fourier coefficients to create a power spectrum (Fisch, 1999). In sleep research the EEG is conventionally analysed using Rechtschaffen and Kales' criteria (1968) in order to create a uniform paradigm, thereby reducing cross-study error variance. For waking EEG data it is most useful to look at specific power spectrum band activity, primarily theta (~4-8 Hz) and alpha (~8-13 Hz) band activity (Klimesch, 1999; Hayashi, Watanabe and Hori, 1999).

1.6.3 Alpha And Theta Bandwidths

Alpha increases are associated with increased drive for sleep (Dijk et al., 1990) while suppression is often associated with increased attentional demands (Ray and Cole, 1995). Studies show strong evidence that increased alpha is related to impaired speed on reaction time tests (Tassi et al., 2006). Increased theta is associated with reductions in performance and has even been shown to increase around ten seconds prior to a lapse (Makeig and Jung, 1996). Sustained wakefulness and subjective fatigue are often marked by obvious increases in theta power density (Finelli et al., 2000). Similarly, decreased beta activity is held as a useful discriminant for impaired performance in sustained attention tasks (Belyavin and Wright, 1987).

Theta increases recorded from midline frontal sites seem to indicate attentional effort and increased cognitive workload (Scerbo et al., 2001), while the same researchers report topographical changes in alpha are more related to the type of information being processed with high alpha (10-12Hz) desynchronisation reflecting semantic or memory demands. These effects change with age as alpha shifts frontally (Kolev, 2002), although the larger frontal activation in the elderly probably reflects compensation rather than deficiency (review; Grady, 1998).

1.6.4 Neurobiology In Vigilance Research
If the Psychomotor Vigilance Task (PVT) represents the basal brain activity of vigilance, then the numerous studies into brain activity during this task should offer an insight into the fundamental processes occurring during sustained attention. Drummond et al. (2005) posit that optimal PVT performance should rely on attention-related regions such as the bilateral ventral lateral thalamic nuclei (short-term) or right prefrontal and left inferior parietal regions (divided attention) after total sleep deprivation. They go on to cite a study conducted by Oguz et al. (2003) that looked at fMRI with simple reaction times and noted increased activation volume for participants with fast reaction times within left sensorimotor and left visual cortices. Further to these findings, Drummond et al. (2005) found that cerebral response during fast reaction times did not change with total sleep deprivation although slow reactions increased in several brain regions. Rather, fast reaction times were indicative of increased responses within cortical and subcortical motor systems.

Alertness and attention are considered to be global functions (Blatter et al., 2005), prerequisite to any specific cognitive performance. Early studies into cats suggested an importance of an intact brain-stem and structures above the midpons to maintaining arousal (Moruzzi and Magoun, 1949). More recent studies into sustaining the alertness state implicates right frontal and parietal areas to be of particular importance (e.g. Posner and Petersen, 1990), as well as other areas of the brain stem. The motivational system has an obvious role to play in alertness and attention, which includes portions of the frontal lobes, much of the dopamine system, as well as limbic and subcortical structures (Robbins and Everitt, 1996). Even the amygdala is implicated in vigilance research and is associated with hypervigilance in post-traumatic stress disorder as well as difficulty in concentrating, as reported by Davis and Whalen (2001) in their Millennium paper.

Dimensions Of Personality

1.7.1 Personality Scales

Personality scales have long been used to explain human behaviour by attempting to score underlying personal dispositions (Ajzen, 1988). Personality traits are latent, meaning that they are inherently unobservable and only explainable by inference. Therefore personality scales measure components of behaviour that typically form a battery of standardised questions that have passed various reliability tests (e.g. predictive, content, convergent validity etc). These questions are hypothesised to reflect the personality dimension they were designed to probe, but it can never be proven whether they are reflective of an entirely different dimension or a subcomponent of another dimension. Personality scales have the benefit of being able to explain inter-subject variance as a trait component, or within-subject variance as a state component. In larger studies the variance caused by these factors can be removed as covariates or used to independently explain differences in results.

Unsurprisingly, there are many criticisms and detractions that accompany psychometric scales, even when the subject is fully compliant and
encouraged to provide honest answers. For example the scales used to measure subjective sleepiness are heavily influenced by biases and may not reflect what they purport to measure (Chervin and Aldrich, 1999), often being dependant on the type of task being used (Gillber, Kecklund and Åkerstedt, 1994) and susceptible to fluctuations in motivation and external factors (Curcio, Casagrande and Bertini, 2001). Scales such as the KSS (Karolinska Sleepiness Scale; Åkerskrdt and Gillberg, 1990), which measure a state component of sleepiness, are bipolar (from ‘extremely alert’ to ‘extremely sleepy, fighting sleep’) and it is this bipolarity that makes larger effects possible (Åhsberg et al., 2000). Scales like the KSS may serve as an alerting factor when asking subjects how awake they feel, or they may reinforce the participant expectation effect (e.g. when participants in sleep studies report enhanced sleepiness to appease experimenter expectations).

1.7.2 Introverts And Extraverts

Extraversion, and its opposite introversion, have been empirically measured and recognised in importance since around the end of the second world war (Eysenck, 1947). Since then a body of research has confirmed that extraversion is a powerful predictor of performance in a number of research domains (Gray, 1970). Introverts are thought to be highly malleable to suggestion, relating to a susceptibility of processes of inhibition or to a higher degree of general arousal (Eysenck, 1957). They are able to focus attention for longer periods of time than extraverts (Davies and Parasuraman, 1982), and are thus ideal candidates for exceptional vigilance performance. Most studies into vigilance and personality have used Eysenck’s later works (1967) to determine extraversion/introversion as part of a three-factor personality model. This model predicts that corticoreticular arousability differs between extraverts and introverts, with introverts preferentially performing best in environments of low stimulation. While this refers mainly to task demands, it can also be extended to surroundings (e.g. peripheral distractions). The extraversion/introversion duality remains one of the most powerful personality dimensions for explaining vigilance performance and there are several appropriate scales that can be used. We have elected to use the EPQ-R (Eysenck, Eysenck and Barrett, 1985) because of its ubiquity and strong theoretical background, which measures Extraversion, Psychoticism and Neuroticism as the “big three” personality factors that are independent of age, sex or geographical origins.
1.7.3 Anxiety

Highly anxious persons tend to be slower than low anxiety persons to target stimuli when distracters of any form are part of a task (Mathews et al., 1990). Trait anxiety, which represents a stable tendency of an individual to experience anxiety on a regular basis (Landers, 1999), has been linked to distraction in several personality studies. Eysenck and Byrne (1992) found that high trait-anxious subjects were more affected by distracting stimuli than either medium or low, although the authors bemoan that previous research is marred by failing to elucidate the potential for other personality dimensions being a factor. The distraction anxiety group had massively elevated saccadic and fixation activity towards peripheral locations as opposed to low-anxious groups. Eysenck and Byrne cite previous research linking the attentional system to anxiety, where solid arguments have been made both ways to argue for a positive effect of distraction, where participants attend to the most relevant stimuli first (Easterbrook, 1959), and a negative effect, where participants spend more resources examining peripheral events (review in Janelle and Singer, 1999). Studies into anxiety and vigilance tend to compare persons with clinical anxiety or personality disorders with people without anxiety complaints, or use emotionally affective targets compared to unloaded targets (MacLeod and Matthews, 1988).

1.7.4 Attention, Arousal And Alertness

Attention is a multidimensional cognitive ability (Posner and Petersen, 1990) involving sustained attention, selective attention and shifting attention (Riccio et al., 2002). Success in tasks or vocations that involve constant attention require individuals to remain alert and avoid processing irrelevant information (e.g. distractions; Fernandez-Duque and Posner, 2001), sometimes for great lengths of time. Hence, persons with above average selective attention should be able to appropriate between relevant and irrelevant stimuli better than normal (Armstrong, 1997). Constant failures of attention eventually manifest themselves as global disorders (e.g. Attention Deficit Hyperactivity Disorder, schizophrenia; Riccio et al., 2002), so studies using ‘normal’ subjects do not probe the extremes of attention. Vigilance tasks determine how successfully humans can attend to monotonous stimuli and demonstrates, even in the simplest of tasks, the wide variation in human preparedness and tolerance for sustained attention. Multi-component tasks like the MATB have sustained attention as an undercurrent to the selective attention required to monitor tasks that are running in parallel.

Arousability is a term used to describe the responsiveness of individuals to variations in environmental conditions (Coren and Mah, 1993). It is an individual trait that can be seen as a basal response to extraneous stimuli, and therefore ties into the anxiety construct and by extension distractions. Sleepiness is known to directly affect arousal (Davies, Bennett and Stradling, 1997) as the onset of sleep-initiating processes cause a decline in state arousal. Trait arousal can be approximated by similar subjective scales as those used to estimate other personality dimensions (e.g. Coren, 1988).
Arousal and alertness are two interchangeable words that describe the same behavioural response to the environment (Bonnet and Arand, 1999). Alertness can be measured subjectively or physiologically, although an incontrovertible measure linking the two with backwards-reliability has yet to appear and it is typical for sleep latencies and subjective scores to be uncorrelated (Johnson et al., 1991). Numeric and analogue scales present stable levels of intra-subject subjective alertness during the daytime (Lafrance and Dumont, 2000), although it has been suggested that subjective sleepiness and objective alertness measures may in fact be assessing independent brain functions (Frey, Badia, and Wright, 2004) such is the desynchrony between the two.
CHAPTER TWO

Methodology
Tasks Utilised

2.1.1 The Psychomotor Vigilance Task

The PVT is ubiquitous in the field of sleep research, having been heavily validated for sensitivity to various forms of sleep loss (e.g. Dinges et al., 1997), circadian variation (Cajochen et al., 1999), and subjective ratings of sleepiness (Frey, Badia and Wright, 2004). The PVT is a monotonous reaction time test that is experimenter-paced (Casagrande et al., 1997) and gives instant knowledge of results, as well as offering an extremely shallow learning curve (Belenky et al, 2003; Dinges et al., 1997). It is designed to measure vigilance defined as “a state of readiness to detect and respond to certain specified small changes occurring at random intervals in the environment” (Weinger, 1999).

As sleepiness increases a progressive decline is seen in reaction times, response time variability and false starts (Roehrs et al., 2000), with recent research demonstrating only sixteen hours of wakefulness causes PVT reaction times and lapses to become unstable (Durmer and Dinges, 2005). In fact, some studies have shown that sleepiness effects are evident even two minutes into the test (Loh et al., 2004), although this view is contentious as sensitivity is compromised in shorter-duration tasks (Mullaney et al., 1983). By means of analogy, 17-24 hours of sustained wakefulness can result in psychomotor deficits comparable to blood alcohol concentrations of 0.05% to 0.1% (Dawson and Reid, 1997).

However, due to the artificial nature of the test and low ecological validity there are question marks over its usefulness as a predictor of performance in real-life situations where operators only have to attend to a stimuli on seldom
occasions, and reaction times over 500ms aren't deemed critical lapses in concentration. Howard and colleagues went as far as to say that "psychomotor tests provide easy and unambiguous scoring but do not represent actual work skills" (p 1353). For this reason the PVT is considered to be a neurobehavioral task (Blatter et al., 2006) that reflects neurobehavioral ability to respond to stimuli, effectively boiling down vigilance tasks into its simplest component to indicate the arousal and attentional state of the individual.

2.1.2 The Multi-Attribute Task Battery

There is an increased recognition of sleepiness' impact in the field of aviation, such as statistics suggesting that human factors account for around 68% of all accidents (Nagel, 1988) due to the constant vigilance requirements for which humans are naturally poorly suited for (Molloy and Parasuraman, 1996). Kirsh (1996) reported a figure of around 4-7% of all air incidents are caused by overly-tired pilots in civilian aviation in the US per annum, while electroencephalographic (EEG) measures support this with studies showing decreases in alpha concomitant with increases in theta band activity indicating increased cognitive workload (Brookings, Wilson and Swain, 1996; Sterman and Mann, 1995). Building on this research, we opted to use a basic cockpit simulator as one of our performance tasks as well as using EEG recordings to approximate parts of the studies above.

Figure 5: The Multi-Attribute Task Battery (Comstock and Arnegard, 1992). A screenshot of the MATB near the beginning of the task. Dial F1 is out of place (it is greater than one notch from centre), light F6 requires action and pump 3 is broken.

The MATB is a multi-component computer task developed by Comstock and Arnegard (commercial release, 1992) that simulates a cockpit in an accessible manner for naïve ('non-expert') participants. The task is pitched at a level that avoids underload, which increases the probability of an error in response to an emergency (Arnegard, 1991), but also avoids the opposite extreme through custom scripting that can be changed to alter the difficulty. There are
three main components that subjects need to balance concomitantly: a central resource allocation task, a systems monitoring task and a communications task. It is a divided attention task as subjects need to focus on the resource task by adjusting pumps to keep the oil in two tanks stable while looking out for deviations in the dials, light failures and audio signals.

Allocating mental resources to this task is an important aspect to consider and likely improves as time-sharing efficiency improves with time on task (Wickens, 1984), while the PVT and other tests of reaction arguably encourages participants to 'automate' attention to the task (Schneider, 1985) due to its inherent predictability. By reducing mental workload an operator turns from novice to expert quicker (Eysenck and Calvo, 1992) and requires the operator to become adept at timesharing skills (Braune and Wickens, 1986). The lapses that occur when a signal is missed during the systems monitoring aspect of the task are not directly comparable to PVT lapses as they can often occur when too much attention or focus is placed on the primary resource task; something unpractised subjects find initially difficult. How long it takes participants to reach their ceiling on this task depends on which subcomponents are being used, although we elected to automate the tracking task (thereby eliminating it) to retain the simple monotony ascribed to systems monitoring tasks, thus reducing the overall demands.

Interestingly, in Arnegard's paper (1991) a steady learning curve was demonstrated on this task as performance improved with time on task. However in Caldwell, Ramsott and Garnder's study (1998), using sleep-deprived volunteers, ability on the task declined with time on task. Comparison of these two studies tends towards the conclusion that sleep deprivation on the MATB impairs performance that would otherwise improve. Because this task is mentally stimulating and engaging, the results can be better appropriated to similarly demanding tasks (e.g. the X-Ray ORT) which are less prone to the effects of sleep deprivation.

2.1.3 The X-Ray Object Recognition Task

The X-Ray ORT is a computer based program that presents subjects with two randomly intertwined sets of simulated luggage items containing 128 threat items and 128 items without threats. Threat items were either knives or guns (8 of each) of varying difficulty, which are shown prior to the test to reduce knowledge effects (Hardmeier, Hofer and Schwaninger, 2005). The threat items are shown in various viewpoints, bag complexities and superpositions so that subjects cannot simply remember overall images, requiring flexible thought and attention. The items are only shown on screen for a duration of four seconds as research has shown that luggage screeners only have between 3-6s to inspect each bag during busy periods (Singh and Singh, 2004; Schwaninger, Michel and Bolfing, 2007).
Figure 6: The X-Ray Object Recognition Task (from Schwaninger, Michel and Bolfing, 2007). Subjects only have two choices: ok, or not ok, and then rate their confidence. No initial position is set for confidence. In this example a gun is clearly visible and so the bag is “not ok.”

Before each session subjects are presented with eight practices and given response feedback, although they are not given response feedback during the experimental tests. Its relative simplicity is useful for using naive raters, as piloting showed that four trials were sufficient to bring participants up to their ceiling. Images are only shown in monochrome to reduce the complexity of the task. Subjects only have two options (“OK” or “NOT OK”) and provide confidence ratings after each item, having as much time as they like to make their response once the images disappear or even before that (commonly the case for hits). It is therefore a subject-paced task; one of the factors that reduces the expected sensitivity to sleep (Gillberg and Åkerstedt, 1998). Task durations vary around 20-30 minutes although there are three scheduled ‘breaks’ during the task so that subjects can divert their gaze from the screen. This task length accurately represents the job that actual screeners perform because they tend to work in groups of three, rotating positions every 20 minutes to two other non-screening positions (e.g. putting bags onto the conveyor). The X-Ray ORT boasts internal, convergent, discriminant and criterion-related validity when compared to various other tests and crucially on-the-job measures from Threat-Image Projection data (TIP; Hardmeier, Hofer and Schwaninger, 2005).

Measurements

2.2.1 The PVT

PVT measures are, by definition, simple measures of the distribution and magnitude of reaction time that change according to experimental manipulation. In sleep research, the lapse has been shown to be the measure that is most sensitive to increasing sleepiness (Dinges and Lim, 2008). Lapses for the PVT are those reaction times above 500ms, which is deemed
to be an adequate amount of time for a non-sleepy subject to respond within. Alongside lapses are measurements of the mean and median reaction time, as well as the 10th and 90th percentiles for reaction times. The 10th percentile is preferential to the 'slowest 10%' concept used in similar studies because it removes the outliers, which often skew data to a large proportion between individuals. False starts, which occur when subjects anticipate an upcoming event by pressing the response button early, can also be used for analysis.

2.2.2 The MATB

For the central resource task, the most useful measurement is the root-mean square value of the average deviation from the target levels, combined for tanks A and B. The larger the deviation from the target, the worse the subject has performed. For the dials, the response time to dial fluctuations and lapses are used for performance measurement, where lapses are a failure to correct the problem within the allotted 15 seconds, upon which time the system corrects itself. The same measurements are used for the red and green dials, and also 'false starts' where the user has attempted to correct an already-functioning system. Finally, the communications task can be used to determine whether subjects have responded to the required auditory commands. Given how simple this task was for trained subjects, it was only used as a means to increase the task load at periodic intervals.

2.2.3 The X-Ray ORT

Statistics used to evaluate performance in two-choice tasks are discussed in detail in Chapter Five, but the basics are explained here. The evaluation of two-choice task performance has traditionally been measured in terms of Signal Detection Theory (SDT; Green and Swets, 1966) which categorises responses into four categories: hits, misses, false alarms and correct rejections. The reasons for these are simple: if performance was measured purely as a count of the amount of items correctly identified, it would only be reflective of a tendency to choose the "threat present" option (Schwaninger, 2003). A rater could click "threat present" for every item thinking it an ingenious way to ensure perfect results, but doing so would result in high false alarm rates. In practical terms, it is not economically viable to check every single bag at an airport for threats, so screeners who adopt this policy would be in for a performance evaluation very quickly! Thus measures that incorporate both hits and false alarms (e.g. d', Δm, Az, A' etc) as decision variables are a popular way of joining both together but carry various assumptions for different measures (heterogeneity between both distributions, fixed sampling intervals etc; for an excellent review see Pastore et al., 2003).

The second statistic that commonly accompanies a decision variable is the criterion, otherwise known as bias, depending on which model is used (e.g. Luce's choice model, 1963; Grier's nonparametric model, 1971). This is a likelihood ratio where a given value represents completely unbiased performance (e.g. β = 1 for d' measures), and values above or below this value represent unbiased or biased performance respectively. In real terms, this means that a subject with unbiased performance requires a stronger
degree of certainty to make a positive decision (e.g. threat present), whereas
a subject with biased performance is more likely to report a threat item being
present when in fact no threat item was present. We have elected to use the
statistics A' and B" as they make less assumptions about the distributions of
both threat and non-threat items (Pollack and Norman, 1964), and the
mapping of results in isopleth-space reflects more accurately the formation of
results from actual data (Zhang and Mueller, 2005). While d' ranges from −ve
to +ve with 0 representing chance performance, A' ranges from 0 to 1
with a value of 0.5 representing chance. Even though reaction time is an
important dimension of performance, there is no theory based in SDT that is
capable of resolving time into a decision variable parameter.

2.2.4 Video Ratings Of Behaviour

Due to economic and technological constraints video data is largely missing
from sleep research literature despite, as Ogilvie (2001) rightly points out,
sleep and sleepiness being a behaviour and as such should be recorded
rather than over-reliance on EEG and dependant variable analyses
characteristic of quantitative research. Often studies retrospectively analyse
data to make inferences, such as the lapse hypothesis, without knowing how
a subject has actually responded to stimuli behaviourally. Using video data
helps move study results from a black box of dependent variables to
uncovering the core behavioural principles that experimental psychology is
apt to address. A driver report financed by the European commission lists
valid video-recorded sleepiness behavioural indices as: eye closures, low
muscle tone, gaze deviations, yawning, and change in body postures
(Sagberg et al., 2004). We have elected to use two of these measures, eye
closures and gaze deviations, as behavioural correlates to sleepiness as the
former reflects a flat-spot from activity (an inability to respond to signals) and
the latter a compensatory strategy.

Of the few studies to tackle video rating outside of driver sleepiness research
Peiris et al. (2006) is the stand-out example, utilising a scoring method from
Wierwille and Ellisworth (1994) although crucially adding a distraction
component. Intervals when the subject diverted from the task were scored as
'distracted,' which accounted for only 0.2 ± 0.1% of overall activity but
explained the reason for cessation of movement of the response cursor for a
predetermined limit of time (known otherwise as flat spot rate) more than four
of the six other measures (130.6 ± 125.8/hr, behind behavioural microsleep;
179.3 ± 27.4/hr) including levels of sleep and forced eye closure. Driver
simulator studies sometimes use offline analysis of video recordings to make
inferences on the nature of accidents, an example of which is Rimini-Doering
et al. (2001). These researchers stated three stages of behavioural
sleepiness: fully awake, 'nodding off' (accompanied by the behaviour of loss
of tonus around the neck), and serious loss of control (typified by a full-blown
sleep episode). Because of the subjective nature of these deductions and the
large amount of time it takes to annotate such behaviours to specific times or
actions, researchers may be less inclined to attempt to record behaviour
during tasks.
2.2.5 The Electroencephalogram

Theory and research that underlies the EEG experimental method was described in Chapter One, but methodological details were largely omitted. The EEG is a highly standardised tool used to measure cortical brain activity that is widely used throughout sleep research to empirically measure sleepiness, sleep and event-related neuronal responses to stimuli (Teplan, 2002). Electrodes are arranged in a standardised montage based on the international 10-20 system (Jasper, 1958) so that results can be compared between laboratories. The 10-20 system is shown in Appendix C. Monopolar and bipolar derivations are two popular ways of comparing potential difference, where the former uses an area of the head or body that is closest to being electrically neutral compared to a cortical area of interest, while the latter compares electrodes that are placed so that they both have a favourable chance of measuring activity, meaning that the specific difference between regions is calculated and not the absolute value (Fehmi and Sundor, 1989). Monopolar electrode montages are often termed common-reference or unipolar as they refer to the process of trying to isolate an electrical point of reference.

All EEG recordings use a differential amplifier to magnify the recording using a common ground (Goff, 1974). Since the potentials have to travel through brain matter, the skull and eventually through to the electrode (which is coated in conductive paste to reduce impedance), the generated signals are very weak and the problem of source localisation happens where it is impossible to specify exactly where in the brain a potential began (Fisch, 1999). During normal wakefulness beta waves (>13Hz) are dominant, whereas in relaxation or drowsiness alpha waves (8-13Hz) become the most conspicuous waveform when spectrally analysed using Fast Fourier Transform (Cooley and Tukey, 1965); a technique first applied by Lubin, Johnson and Austin only four years later (1969). The alpha rhythm exhibits a reduction in power as subjects experience progressively higher sleep pressure with increased time spent awake (Ferreira et al., 2006), and may prove to be the most useful determinant to measure objective sleepiness if technological limitations can be removed in the future.

As has been discussed earlier in the 'measuring distraction' section, there is a wealth of information that can be gleaned from simple eye movements. As well as using direction of gaze to make attention inferences, the behaviour of eye closures can be measured as an accurate analogue to sleepiness; something the electrooculogram (EOG) can derive. EOG techniques use subtly different methods to affix the electrodes around the outer canthi of each eye than EEG measures on the scalp. As the eyeball is a moving dipole with a background electrical activity of 50μV (Aserinsky and Kleitman, 1955), deflections in EOG power can be stored that automatically traces eye movements on a midsagittal plane. Slow eye movements (SEMs) are characteristic of the sleep-initiating process that filter through wakefulness (Aserinsky and Kleitman, 1955), becoming larger and more sinusoidal when simultaneous slowing of the EEG occurs (Cajochen et al., 1999). Eye blink data has historically been used to characterise sleepiness (Caffier, Erdmann
and Ullsperger, 2003), with blink duration and opening time being the two parameters most sensitive to increases in drowsiness. Further to this, pendular horizontal eye movements were found to be the first determinant of sleep onset in more than half (50.5%) of a 200-person sample of US air force personnel (Maulsby, 1966). Saccade frequencies and velocities also tend to decline with time on task as fatigue and motivation begin to decline (Van Orden, Jung and Makeig, 2000).

2.2.6 Actigraphy

Actiwatches are small wrist-worn devices that can be used in conjunction with sleep diaries to ascertain sleep onset and offset, as well as sleep efficiency and other kinetic derivatives (Lockley, Skene and Arendt, 1999). Aggregating measures over multiple nights stabilises differences thus yielding measures with the power to predict other variables (Sadeh and Acebo, 2002), as well as correlating highly (94.5%) with EEG data (Mullaney et al., 1980). Actiwatches are small, unintrusive and useful for when experimental compliance is critical and allows participants to sleep at home, eliminating any 'first night effects' seen in the laboratory (Browman and Cartwright, 1980; Coates et al., 1981). For these reasons actiwatches were chosen as the main assurance that participants had slept the requisite amount and avoided napping during the day.

Building The Laboratory

2.3.1 Planning The Distraction Laboratory

A new performance laboratory was constructed specifically to address the research interests created from our hypotheses, enabling high customisation to our specification. The laboratory was split into two parts separated by a curtain: the experimental side where participants performed their tasks, and an operator control side. A schematic diagram of the final design is shown in Figure 7, and photographs taken of the laboratory can be found in Figures Figure 8 and 9. Only after several incarnations did this finalised design arise, and experimentation could begin proper.
Figure 7: The final layout of the laboratory. This schematic shows the laboratory setup that was used in both studies. Separated by a large curtain (dotted blue line), subjects sat at each terminal with the television at 90 degrees in the periphery. Subjects could neither see nor hear the experimenters.

The experimental side was designed to accommodate two participants simultaneously, split by a divide so that they could not see each other. At the participants’ area were Viglen computers running Windows XP with the monitors set at a 90° angle to a 23” flatscreen television used to provide visual and audio distraction. The reason for this setup was to ensure that head turns were volitional, providing a measure of behavioural distraction. This television was connected to a DVD player that was used to play back the television programs chosen as peripheral distractions. To record these head turns four cameras were strategically positioned: two of which were set above the computer monitors to give a full face-on image, and two cameras that were mounted to the roof to show a wider ‘birds-eye’ angle. Attached to the back of the chairs were four-channel electrode boxes used to connect the EEG signal that were routed through a differential (AC to DC capacitor) amplifier to limit voltage.
At the control side were two quad-split monitors that assimilated the two cameras showing the participant at the top half, with one quadrant showing the EEG feed derived from the operator computer and the final quadrant displaying the actual monitor that the participant was stationed at (Figure 10). Two computers recorded two channels of EEG and two channels of EOG through custom written Dasylab software that stored the data and also calculated real-time fast-Fourier transform of the incoming signals. These two computers were used to control the audio used for the communications task of the Multi-Attribute Task Battery (MATB) and also to record the verbal fluency task created for Chapter Five. Two additional keyboards and mice were connected to the participants' computers so that they could be controlled and manipulated from the operator side, observing the participants' monitors through part of the quad-split monitors. This enabled all task-related changes and operations to be performed on the operator side of the room. Finally, the two blinds in the room were replaced with blackout blinds to create a sterile environment free entirely from time cues.
Running Windows XP presented resolute problems for running the two DOS (Disk Operating System) programs in Study 1, which were chosen so that findings could be directly applicable to previous literature from the 1990s. Given that emulation and the command line had failed, the solution was to run DOS as a virtual operating system on top of Windows XP and to reconfigure parts of the BIOS so that the LPT1 port would change from ECP to bidirectional. Benchmarking tests proved that this system resulted in no loss of performance. The communications task of the MATB had to be completely recreated as the voice synthesiser required had stopped being produced in 1992. For the MATB two new hour-long scripts were written so that there was an equal probability of each event occurring in every ten minute block, and parts of the source code were recompiled in Qbasic to remove the time of day.
Figure 10: The quad-screen monitor setting meant that the task, EEG recordings and behavioural measures were combined into one DVD recording. The experimental controls allowed each of these recordings and tasks to be started in synchronisation, and allowed later analysis of behavioural and task determinants of performance.

2.3.2 Piloting The Distraction Laboratory

All parts of the laboratory had to be extensively piloted before commencement of each study. Early on it became apparent that head turns would have to be quite marked in appearance to distinguish them from involuntary eye reflexes. Thus the stations were altered so that the television derived a 90° angle as opposed to the 45° angle previously used. A divide was erected so that participants could not become a source of distraction to each other during no distraction conditions. Kick boards were also put into place so that participants could not accidentally knock the cameras, which were initially placed in the inner well but later moved when the station viewing angle had to be changed to a 90° angle from the television.

Different metrics were trialled during the piloting sessions, including looking at how distractions could be categorised. In the end it was settled that distractions can only be behaviourally implied and that categorising distractions into long and short duration was the simplest and most reliable measure. For Study 2, using the X-Ray Object Recognition Task (X-Ray ORT), piloting tests made it clear that behavioural distraction measures were of limited use as the test is designed with three breaks for participants to avert their gaze from the screen. It was during these breaks that participants would
relax and look at the television, thereby having no performance implication. These breaks proved to be an effective countermeasure of encouraging sustained vigilance compared to the other tasks, which required constant attention for 30 minutes; something that humans find particularly difficult.

The piloting sessions also allowed for estimation of practice effects so that they could be eliminated from performance. For the MATB an average learning curve of 1hr 15mins was found, so this was factored into the final study design. The X-Ray ORT demonstrated a learning curve of four trials (similar to the MATB at around 60-80mins), necessitating an extra day where participants practiced the task three times in succession before the two experimental days. Lessons learned from the first study were carried over into the succeeding study, where the EEG was changed from monopolar to bipolar as the EOG channels provided no useful data, and an additional level of distraction was introduced to experimentally account for cognitive distractions. These changes are detailed further within the succeeding chapters.

2.3.3 Participant Screening

Screening participants for a study is a labour-intensive and unavoidable aspect of successful laboratory studies. Limits and definitions of ‘normality’ are difficult to operationalise, so it is often most practical to exclude potential applicants based on known extremes. For the studies conducted in the laboratory, recruitment is made via advertisement on intra-university websites and posting boards where participants are invited to contact the laboratory. Then applicants undergo initial screening using a brief list of desirable qualities (e.g. sleep/wake times, age) and if they sleep the regular amount without excessive daytime sleepiness they are invited for a thorough evaluation to ensure:

- No history or complaints of sleep disorder.
- No use of drugs that are known to affect sleep (e.g. sedatives, benzoates, antidepressants etc).
- Limited caffeine and alcohol use of two or less cups of coffee per day or less than 10 units of alcohol per week.
- Regular sleeping patterns around 7-8 hours per night.
- No past experience in overnight shift work.
- Physically healthy and sound of mind, with no evidence of the listed disorders given in the questionnaire e.g. depression, migraines etc.

In addition to the thirty-point questionnaire addressing the above listed criteria, participants were given the Epworth Sleepiness Scale to check for evidence of excessive daytime sleepiness. They must fall within the normal range of sleepiness (<10; Johns, 1991). Following a successful interview, participants are given a sleep diary to complete for three days and an actiwatch so sleep routines can be checked for consistency and indices of sleep efficiency can be calculated (anything above 85% is acceptable). For the recruitment of older participants, slightly modified parameters are used for normality as older people naturally sleep less than younger people (Vgontzas et al., 2003) and may have prescribed medicines. A table of screening
requirements is given in Table 2. Exact protocol and limits for all subjects are detailed within each Chapter.

Table 2: Comparison between screening guidelines for young and old subjects.

<table>
<thead>
<tr>
<th>Screening Parameter</th>
<th>Young</th>
<th>Old</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>18-25yrs</td>
<td>55-65yrs</td>
</tr>
<tr>
<td>Typical Sleep Length</td>
<td>7-8hrs</td>
<td>&gt;6hrs</td>
</tr>
<tr>
<td>Education</td>
<td>University</td>
<td>University/College</td>
</tr>
<tr>
<td>Prescriptions</td>
<td>None</td>
<td>Non-drowsy medication</td>
</tr>
<tr>
<td>Daytime Sleep</td>
<td>None</td>
<td>Infrequent napping</td>
</tr>
<tr>
<td>Eyesight</td>
<td>Normal</td>
<td>Corrected to normal</td>
</tr>
</tbody>
</table>

2.3.4 From The Laboratory To The Workplace

In the laboratory, prospective analyses can be undergone with standardised procedures that reduce noise in a mathematical sense when compared to field studies. Performance is usually found to get progressively worse as wakefulness during the night increases (Folkard and Monk, 1992), although a dichotomy exists between divergent tasks that require innovation and flexibility of thought (Harrison and Horne, 1999), and convergent tasks that are repetitive and resemble automation of mental processes (Wilkinson, 1965). It is the latter that are more susceptible to sleep loss and more relatable to the type of operator vigilance loss that results in the sorts of sleepiness-related incidents reported in the studies cited in the next section. Due to the inherent difficulties in standardising and measuring ‘real life’ performance, tasks such as the PVT have been used as performance proxies even when actual in-the-workplace shiftwork is being investigated (e.g. Harma et al., 2006; Lamond et al., 2003).

As computer simulations increasingly become more sophisticated the convergence between the laboratory and the workplace is narrowing. For example the MATB was originally conceived as a cockpit simulator, albeit a crude substitute limited by an operating system that worked on a 32mb partition. However, findings from studies using software similar to the MATB can be related to vocations that have similar demands for sustained vigilance and divided attention like night-watchmen, CCTV operators or systems monitors etc. By comparison the X-Ray ORT used in the second study provides a strikingly different working environment, approximating the actual job of x-ray screeners very closely (Schwaninger, 2006). Ultimately laboratory studies are a compromise between control and validity, with the choice of task being an important design consideration when making causal inferences extended to the real world.
Preamble To Chapters

2.4.1 From The PVT To The ORT

Careful thought, consideration and years of work have gone into the study designs and executions to probe accurately and reliably the processes that bind sleepiness and distraction. The first study, which forms Chapters Three and Four, was an overnight study that aimed to understand the basic human reaction to high levels of sleep restriction and monotonous demands for work. Over one hundred hours of video footage was collected and analysed for the behavioural effects of monotonous activity; the first study ever to do so in the field of sleep research outside of a driving context. The results that followed would form the backbone of the first paper to question, operationalise and suggest a major revision to the prevailing notion of a lapse that has been used ubiquitously throughout sleep research of the past twenty years.

The second study was a major departure from the first, in that a new sample of subjects were split into two groups (“young” and “old”) to examine age effects in a daytime study. The task used a cutting-edge computer simulator of X-Ray luggage screening and so provided interesting data that had never before been collected for this area of research (e.g. daytime sleepiness, or distraction). This research area in particular is very young and most data derives from field studies, so there are large gaps in the literature where standardised laboratory studies are sorely missed. Due to the nature of the task, which mirrors the screener’s job very closely, the results are more ecological than the first study, which leaned towards the theoretical aspects of performance degradation. The findings from this study form the bulk of Chapter Five and provide the earliest findings from a research area that is bound to flourish over the coming decades.
CHAPTER THREE

Overnight Sustained Vigilance And Distraction

Abstract

Psychomotor vigilance is one of the most studied areas of experimental sleep research as it is highly indicative of sleep need, and reflects the monotony of various vocations such as distance driving that are susceptible to sleep-related incidents. Twenty-four young subjects (m = 23yrs) were tested and video-recorded between 10am-12pm, and again at 4am-6am on the PVT with two levels of distraction- no distraction, and distraction in the form of television showing an enjoyable comedy series. The results showed a high effect of sleepiness and marginal effects of distraction. Power-spectrum EEG analysis showed that alpha increases are related to increases in the mean reaction time, although theta levels are not related to performance. More importantly, the behaviour of the subjects changed with magnitude of reaction time as lapse duration (RT ≥ 500ms) increased, where the percentage chance that participants' eyes were even open during the lapse decreased in a decaying exponential relationship. At the midpoint of the percentage chance of eyes being open are reaction times in the region of 1210ms, whereas above 2740ms there is a 95% chance that the eyes will be closed during the lapse. Therefore a recommendation that these extra dimensions of lapsing are accounted for is suggested, as they are more likely to be reflective of processes of sleepiness that are likely to have a deleterious impact in real-life.
3.1.1 Monotonous Work

Many vocations make only small demands on mental attentional resources, being monotonous or fairly simple to perform and may entail prolonged periods without action before the individual is required to make a response. For example, industrial jobs tend to be repetitive and inherently boring, such as working on an assembly line, as well as security jobs like CCTV monitoring, and of course long-haul drivers who may voluntarily work overnight to avoid heavy traffic on the motorway. These are all domains with a large workforce and jobs that often require overnight or shift work. On top of this, there is a wealth of information from military studies that have examined soldier safety when on long and predictable nightwatch or round-the-clock operations (see Chapter Four). The other major vocational research area in vigilance research is medical residency, which can be repetitive (particularly for anaesthetists) and is a domain where young trainees are expected to endure a 'baptism of fire' of excessive working hours, often topping 80-100 hours per week (Jha, Duncan and Bates, 2001), including routine 36-hour continuous shifts.

Studies during the past half-century have consistently shown human performance declines during the night when undertaking monotonous work, from the speed of connecting calls in a telephone exchange (Browne, 1949) to the frequency of train drivers using their safety devices (Hildebrandt, Rohmert and Rutenfranz, 1974). Injury rates from records kept over a seven-year period in Norwegian offshore rigs (Lauridsen and Tonnesen, 1990) were found to be heightened between midnight and 6am, accounting for over a quarter of all injuries. Even though work in an industrial plant in Ontario, USA, was more dangerous during the day and required more activity it was the night shift that was associated with an increased frequency of accidents (Wojtczak-Jaroszowa and Jarosz, 1987). A later analysis of Polish power industry operators showed an age effect of industrial accidents, where younger operators (< 29yrs) were more frequently accident victims at night than in the morning, although older operators (40-49yrs) had the highest gross injury rate (Zuzewicz and Konarska, 2005). Thierry and Mejiman (1994) argued that absenteeism, productivity and turnover are all negatively affected by having to work at the circadian nadir, but cautioned that the combined studies cannot be used to make unambiguous claims.

The job of the anaesthetist, while clinically important, is routine and unchallenging for suitably trained practitioners (Weinger, 1999). A simulator study using Stanford anaesthetist students found that there was strong behavioural evidence of the effects of sleepiness when they were kept awake for longer than 24 hours, using video data reviewed by ‘blind’ raters (Howard et al., 1998). Anaesthetists have been found to have a base level of sleepiness comparable to persons who suffer from narcolepsy or sleep apnea (Howard et al., 2002). Other vocations in the hospital, using actual real-life performance data, have shown deleterious effects of sleepiness, for example midwifery. Swedish longitudinal data between 1973 and 1995 suggest that infant mortality rates are higher when birth occurs during the night shift than
during the day shift (Luo and Karlberg, 2001). These authors discussed studies conducted in Germany and Great Britain that supports the hypothesis that night-time births are more likely to end in neonatal mortality than daytime births, and attribute this either to excessive workloads or inexperienced staff being given overnight work, possibly attracted by increased pay.

3.1.2 Risks Associated With Loss Of Sleep

To Costa (1994), there were four main spheres of health and well-being that are impacted by shiftwork: 'biological', due to disruptions in circadian function, 'working', or work-related efficiency and risk, 'social' and 'medical'. Besides the increased risk of sleep-related accidents, there are lifestyle differences that are endemic of shiftworkers that can be associated with having to work at times that the human body is not naturally designed to work during. In their seminal review Folkard and Tucker (2003) aggregated the results from eight different industrial datasets to find a normative risk of 'incidents' (a melange of accidents and injuries) at different periods of the day. They found increasing relative risks (RR) from morning (RR = 1.0), to afternoon (RR ~ 1.18) and the highest during the night (RR = 1.3). When the night shift alone was considered, there was a peak around 23:00 (RR = 1.2), declining down to 05:00 (RR = 0.7). This trend increased with successive nights of overnight work, where the RR started at 1.0 normative for the first night and increasing approximately linearly until the fourth night, which capped RR between 1.35 and 1.4. Four studies were then used to approximate risk with hours on duty which increased almost exponentially, and also minutes since last break showed a monotonic increase in RR. These results powerfully demonstrate in a meta-analytic approach the trends confirmed throughout the literature: that shiftworkers are at greater risk of accidents due to fatigue and sleepiness.

In the medical literature the raft of evidence against the brutal hours endured by trainee doctors and medics during their residency and early career has resulted in legislative changes to provide a fairer working environment. A massive 405 regulations in New York were reviewed (known as the 'Bell regulations'; Asch and Parker, 1988) after a high-profile case involving an eighteen-year old college student named Libby Zion, who died while under the care of medical interns (Wallack and Chao, 2001). The authors note that under normal circumstances this would have resulted in a civil malpractice suit at worst, except for that fact that Zion's father was a former federal prosecutor who hired private investigators and expert physicians to cause a major shake-up of the system. As of July 1st 2003, the majority of residents will be limited to 80 hours of work per week in America (Steinbrook, 2003) - falling far short of current European regulations of 56 hours of work per week implemented in 1991 (Jagsi and Surenger, 2004). For a comprehensive chronology of the medical internship working hours debate through the press and literature see Steinbrook, 2003.
3.1.3 Psychomotor Vigilance

Psychomotor vigilance is the ability to respond to unprepared stimuli in a timely fashion, being the simplest assay of neurobehavioural function possible by simply requiring a single acknowledgement that a target has appeared with no onscreen distracters. Performance in psychomotor tasks degrade rapidly when a variety of interventions are given and usually becomes one of the first measures to detect the onset of sleepiness. All of the derived metrics on these tests are impaired by sleepiness (Doran et al., 2001), including fastest and slowest 10% reaction times, frequency of slow reactions, and false start errors; the extent of which depends on the study design. Results from studies utilising psychomotor vigilance tasks are replicable with high levels of cross-study reliability due to the uncomplicated nature of the measures. These tasks therefore present conceptual findings that underpin practically every cognitive theory in sleep research (Lim and Dinges, 2008) by being the measure most closely associated with uncontaminated neurobehavioural performance. The effects of sleepiness become apparent even within the first five minutes of some vigilance tasks (auditory vigilance: Lisper and Kjellberg, 1972), as well as showing clear age effects (Blatter et al., 2005).

There are a number of instruments that can be used to relate physiological vigilance to brain activity; namely EEG, PET (Positron Emission Technology), and fMRI (Functional Magnetic Resonance Imaging). Markers of EEG activity to psychomotor performance relate mainly to the initiation of sleep-onset from drowsiness to a full-on unintended sleep episode, of which there have been many studies. Frontal increases in alpha activity predict impoverished psychomotor performance and is usually the EEG measure most typically associated with declines in psychomotor performance (Strijkstra et al., 2003). During 32-hour sleep studies, the power in the high theta/low alpha band (6-9Hz) progressively increases as the desire to sleep becomes stronger (Cajochen et al., 1995; Caldwell et al., 2004). High alpha (10-12Hz) is more sensitive to variations in workload than low alpha (Scerbo et al., 2001) and can also be used to index attention. Some researchers have shown that overall analysis of the power spectrum, as opposed to looking at specific bands of wavelengths, predicts slowing of reaction times also (Jung et al., 1997). However, caution needs to be exercised when interpreting these findings as EEG band amplitudes and performance varies according to which performance measure is used, the task used, how the subject feels and the site of the electrodes (Makeig and Inlow, 1993).

PET studies have complimented EEG studies in the conclusion that metabolic activity fluctuates and changes according to levels of arousal. Frontal and temporal lobe changes in metabolic activity were reported after sleep deprivation, and slowing of reaction times were linked to an overall increase in brain metabolic activity (Wu et al., 1991). Cognitive performance has been associated with PET results in range of studies, with the prefrontal cortex and parietal lobes being the brain regions showing most response (Thomas et al., 1998). fMRI studies have found that activity in the anterior cingulated cortex is attributable to errors and correct responses in reaction time tests (Carter et al., 1998).
Tomasi et al. (2007) have shown that increased cognitive load activates common networks in the parietal and occipital cortices, and that increased fMRI response in the prefrontal cortex increased during the verbal working-memory task but not the visual attention task. Taken together these studies show the importance and depth of psychomotor measures in sleep research.

3.1.4 Distractions During Office Work

Office workers may not be in any immediate danger when sleepiness or distraction becomes a nuisance factor to their routine, unlike in driving or security settings, but the lack of productivity associated with environments filled with auditory or visual distractions is a major concern for employers (Roper and Juneja, 2007). So-called 'knowledge workers', as opposed to manual workers, are said to contribute to around 78% of an organisation's total operating revenue (US General Services Administration, 1999; cf. Roper and Juneja, 2007). When 13,000 office workers were quizzed in a large-scale survey, the attribute they thought most important to their productivity was "the ability to do distraction-free solo work" (Olson, 2002). With the ever-increasing number of open-planned office space, there is growing concern about whether these are productive environments, or simply cost-cutting at the detriment of employee happiness and productivity. More than half of all offices in America and Canada are classified as "open plan" (IFMA, 1997), giving credence to concerns about the maximal corporate setup. The ability to make private phone calls, concentrate on a reclusive task or simply being allowed space to think are "luxuries seldom afforded by a typical open office configuration" (Mardex, 2004, p.2). Reviews on the open office specification conclude that workers report markedly reduced job satisfaction emanating from a loss of privacy and abundance of visual and audio distraction (De Croon et al., 2005).

The conclusion of many studies in this research area is that employees find that conversations held by neighbouring workers are the most disturbing factor (Heerwagen et al., 2004; Jensen, Arens and Zagreus, 2005; Pejtersen et al., 2006). These interruptions to thought are not just minor nuisances; often these interruptions prevent workers during the critical flow of thought to which they may not regain that day (Demarco and Lister, 1999). Olson (2002) found that workers would spend around a quarter of their time talking in or near adjoining workspaces, being a mixture of business-relevant and personal discourse. Simple tasks such as proofreading or basic mental arithmetic are worsened when irrelevant speech takes place nearby or other peripheral distractions take place (e.g. phones ringing; Evans and Johnson, 2000). Of the multitude of potential auditory distractions present in the modern workplace, Boyce (1974) found that 67% of workers from their sample were disturbed by telephones ringing and more than half were disturbed by people talking. Sunderstrom et al. (1994) found again that more than half of respondents in a 2000-person sample were bothered by other people talking and phones ringing. There is no shortage of evidence then that the sort of environment that open offices provide are a hive of distractions from work.
In particular it is irrelevant or incomprehensible speech which causes most irritation to workers (Banbury and Berry, 2005), accentuated by the fact that workers may be interested to overhear nearby conversations. In the laboratory the "irrelevant speech effect" is well known, where visually-presented information is poorly remembered when irrelevant spoken material is presented during learning (Baddeley, 1992). These findings are extendable to the workplace where clearly audible speech is distracting but not to the point of irritation, whereas inaudible speech causes problems with *iconic* memory and eventual loss of their train of thought. The actual level of the noise is a significant factor, as the louder the noise (so long as it remains incomprehensible) then the larger the 'annoyance' factor (Ellermeier and Hellbruck, 1998). Other studies, by contrast, report that noises within the 48-76dB range are all equally disrupting (Tremblay and Jones, 1999). The theoretical interest stems from the replicable interaction between auditory perception and immediate recall, resulting in mean error rates between 30-329% (Beaman, 2005). A review of the literature for the irrelevant speech effect concluded that the effects are indeed involuntary, perhaps relating to an evolutionary mechanism, and that the effects of extraneous background noise on performance is considerable (Beaman, 2005).

3.1.5 The Psychomotor Vigilance Task

The PVT is comfortably the most studied experimental task in the field of sleep research (Lim and Dinges, 2008), helping researchers to understand the basic processes that drive and contribute to sleepiness-related accidents and loss of performance. The PVT was originally developed back in 1985 (Dinges and Powell, 1985) as a computerised version of a sustained vigilance paradigm that had already been well established (Wilkinson, 1968). Its distribution, ease of use and shallow learning curve (Belenky *et al.*, 2003) made it popular among researchers looking for a task that was highly likely to illicit and quantify sleep effects. It is available in two main packages: as the original DOS software, and a commercial unit that is light, portable and quick to administer (Rosekind *et al.*, 2000). The former is relatively inexpensive and useful for computer-based studies, whereas the latter is more appropriate for field studies where computer equipment cannot be guaranteed and time for performance testing is limited. This test is highly sensitive to the effects of sleep deprivation (Dinges and Kribbs, 1991) no matter whether it was accrued as total sleep deprivation, sleep fragmentation or partial sleep restriction (Lamond, Dawson and Roach, 2005).

To complete the PVT, all that the subject is required to do is to respond to the onset of numbers that appear at random on the screen using the response box provided. Before the task begins there is a visual analogue scale that subjects rate their sleepiness on, and then the task continues for a period of thirty minutes. The interstimulus interval varies between two and ten seconds and the numbers (time in milliseconds since onset) increase up to a limit of thirty seconds. Instant knowledge of results (Wilkinson, 1961) make the PVT a vigilance task that allows subjects to know when they are performing poorly, as well as being experimenter-paced (Casagrande *et al.*, 1997). Experimenter-paced tasks are naturally more sensitive to sleep and mirror
those situations in life where control is out of the individual’s hands e.g. in
duty vocations. The test quickly becomes unchallenging and incredibly boring,
reducing levels of motivation and lacking in any inherent reward structure for
the subject.

3.1.6 The Lapse Construct

Lapses in attention have long entered the layman’s vernacular, but
scientifically quantifying what is and what is not a lapse in concentration has
proven to be an enduring issue. The idea of a lapse dates back almost eighty
years to Bills (1931), but was originally called a ‘block’ to adjoin the familiar
phrase “mental block”. An excerpt from his paper, which like many of these
classic early texts is easily found online, shows that the major research
questions are still hotly contested today:

“Blocking has been frequently observed by students of mental
work, but has not heretofore been systematically studied to
determine
(a) what laws govern the occurrence of these blocks,
(b) what factors determine their frequency,
(c) the relation of blocking to fatigue and the general decrement
(d) its relation to the principle of refractory phase, and
(e) the relation of blocking to the occurrence of errors”

Bills, 1931, p.230

Even then the ‘block’ was “arbitrarily defined” (p.231) as “a pause in the
responses equivalent to the time of two or more average responses.” He
found that blocks tended to last between 2-6 seconds with large individual
differences, and that fatigue increased both the frequency and duration of
these blocks, especially when the seven-minute tasks were replaced with one
that lasted a full hour. When errors did occur they tended to coincide with
these periods of mental blocking, suggesting a change in neural functioning.
While research methods and funding have undergone dramatic changes
since this time, it is somewhat sobering to read that the same issues and
conclusions really have not changed much since the early days of sleep
psychology. With that, it is time to move forward several decades and explore
our current understandings of what makes humans lapse in attention.

The term ‘blocks’ was replaced with ‘lapses’ by Bjerner (1949) and was to
stick due to the work of Williams, Lubin and Goodnow (1959), who
popularised the term with their influential paper. They showed that lapse
duration and frequency was interrelated with distributions of long reaction
times and general slowing. The fastest reaction times never changed much
during the 72 hours of sleep deprivation, but it was the longest reaction times
that showed the most marked change. Later studies utilising very similar
protocols would confirm these findings (Babkoff et al., 1985), showing a
neural response to flat spots of activity called microsleeps and correlating with
slow-eyelid closures (see Dinges et al., 1998). Bob Wilkinson spent much of
his career researching human responses to shape this area of research
during 1950-1990 (e.g. Wilkinson, 1967), and created a portable device to
measure reactions (Glenville and Wilkinson, 1979) but curiously did not measure either lapses or blocks.

Lapses up until this point were still being classified as those reaction times occurring greater than a function of two-times the subject's average until the late 1980s. Inexplicably and without any justification as to why, when the PVT (patented by Bob Wilkinson's student, David Dinges) replaced Wilkinson's reaction time test as the preferred method of response-time study it sold with the lapse construct measured by any reaction time greater than 500ms in the results output. In 1988 Dinges et al. defined lapses as the slowest 10% of reaction times; a metric that found its way into the commercial copy of the PVT but not marketed as a lapse. Even as recently as 2008 there are a few prominent studies that have used this 10% slowest RT as lapses (Chee et al., 2008), though there has never been any justification as to why 10% and not for example 5% or 15%. As the PVT was the software most laboratories would come to use the idea stuck, and the 500ms lapse cut-off has been pervasive throughout the literature ever since.

However the mysterious 500ms lapse definition was derived, there is little doubt that it is highly sensitive to sleep deprivation, showing a dose-response relationship to levels of sleepiness (Van Dongen et al., 2003). The number of lapses also gradually increased with each day of sleep restriction in a similar study (Bekenky et al., 2003), returning rapidly to normal pre-treatment levels immediately after recovery sleep. Lapses are thought to be caused by state instability (Doran, Van Dongen and Dinges, 2001), which asserts that performance is not only a global attribute of attention (shown by slower overall reaction times) but that short periods of inability to respond punctuate overall cognitive slowing. These uncontrolled sleep attacks are characteristic of vigilance tasks, of which the PVT is unquestionably the most difficult for subjects to finish without lapsing due to its base monotony. Sleep deprivation does not eliminate the ability to perform neurobehavioural functions, but it does reduce the ability to perform stably without error (Kleitman, 1923).

3.1.7 A Sleep-Restricted Pilot Study

As a precursor to this research study, Anderson and Horne (2006) conducted a daytime sleep-restriction study using the PVT on a sample of 16 young subjects (age = 21-25yrs). Sleep was restricted to less than five hours on the sleep restriction day and normal sleep allowed as comparison. A statistically-significant increase in head turns occurred only for the sleepy and distracted condition, whereas for lapses there was an interaction effect between sleepiness and distraction, and a large main effect of sleepiness. The large number of head turns was found to occur even within ten minutes of the sleepiness and distraction condition. The main effect of distraction only narrowly missed experimental significance in this study. Anderson and Home concluded that sleepy individuals seek distractions as an attempt to overcome sleepiness or boredom.
3.1.8 Summary

Vigilance is known to be impaired by sleep loss, although the standalone effects of exogenous distraction are less clear. The interaction between psychomotor vigilance and distraction has rarely been studied, and considering the wide literatures of both then we aim to bridge the gap that remains. Concurrently, by using video recordings of participants to gauge distraction then at the same time data can be collected that looks at behavioural response during so-called lapses in behaviour. This is also the first time that the lapse hypothesis has ever been scrutinised or challenged by looking at the behavioural response that is supposed to define a lapse, and not simply totalling reactions. Thus we also aim to reinvent the eighty-year old concept of a lapse and remove the ambiguity that currently defines what is considered to be a lapse.

Figure 11: Top: Average head turns towards the peripheral distraction, and below: the average number of lapses in each condition. From Anderson and Horne’s 2006 study.
Research Questions

1. What extent does psychomotor performance change under low and high sleep pressure?

2. Does performance or attention to a task change behaviourally under increased sleep pressure?

3. What is the mathematical relationship between sleepiness and distraction?

Methods

3.2.1 Screening And Subjects

Twenty-four healthy, young participants (m = 23.17yrs, s.d. = 3.6yrs) were screened for suitability to ensure that all were similar in personal characteristics, sleeping patterns and within the limits of normality on numerous sleep and lifestyle characteristics. Participants were advertised through University mailing lists and also by notifications in various University departments. When potential subjects contacted the research group they were then asked to present themselves to fill out a lengthy questionnaire (Appendix E) and a short face-to-face interview. A total of over seventy volunteers made it to this stage, of which only twenty-four were eventually chosen from a database of over a hundred applicants. Reasons for exclusion included prescriptions to medicines known to impact sleep, history of anxiety disorders, currently experiencing high levels of stress, excessive Epworth Sleepiness Scores (> 10), overnight workers, overuse of alcohol, and routine daytime napping.

Participants who had successfully completed the requisite questionnaire and interview were then entered into a second detailed step of screening where they had to complete a three-night sleep diary and wear an actiwatch (Cambridge Neurotechnology Ltd., Cambridge, UK) for those successive nights. Actiwatches were always issued at the start of the week to avoid the disrupted sleep that typifies the average student weekend. Only once confirmation of routine and ‘normal’ sleeping patterns (between 7-9 hours sleep per evening; sleep efficiency greater than 85%) were obtained were participants then given a testing date to adhere to. Participants were matched for date availability and personality and then given actiwatches to monitor their sleep prior to the testing date, as well as being requested to be awake by 8am on the test date and to avoid all psychostimulants, alcohol, heavy meals and exhaustive exercise. The protocol for this study, including screening techniques, was approved by the Loughborough University Ethical Advisory Committee.
3.2.2 Protocol

On arrival to the laboratory at 6:30pm the subjects filled out ethical forms to acknowledge they understood the lasting effects of twenty-four hours sleep loss and to permit the recording and usage of video data. Their actiwatches were analysed to check that subjects had not napped during the day, and they had arisen from bed at 8am as expected. A random urine sample was taken from a third of all participants to check for recreational drug use (Surescreen Diagnostics, Derby), although no participants failed this test and none were excluded for not complying to instruction. After these formalities were done then the subjects were briefed on the protocol and taken to the laboratory for pre-test practice to eliminate learning effects. A 45min MATB practice was given, followed by a short PVT practice, before a second 30min MATB trial ensued to conclude the practice phase of the study (the MATB segment is discussed in Chapter Four).

![Figure 12: A diagrammatic representation of the study protocol.](image)

Participants were taken to an electrode-placement room where all of the nine electrodes were attached and impedances of less than five micro-ohms were achieved for each scalp electrode, and impedances of less than ten micro-ohms for EOG electrodes. Subjects were then given a light microwave dinner with breadrolls, fruit and water in order that their nutrition and meals were standardised as heavy meals are known to induce sleepiness. Finally, at 10pm the subjects were brought back into the laboratory to commence the first testing block, which were counterbalanced for test type and distractions. For the no distraction condition the television was switched off, but for the distraction condition the television showed a popular BBC comedy series “The Office” to present an enjoyable distraction. After the conclusion of the four thirty-minute tests around midnight the subjects were given supervised ‘free time’ where they could enjoy books, films or games to keep occupied. At 3am they were given a light snack choice of soup or biscuits, and then brought for another testing block at 4am where electrodes were checked for signal. The order of tests and distractions were exactly as they were during the first night block of testing. Finally, around 6am the subjects were discharged and escorted home by taxi and given payment of £100 for their time.
Table 3: Counterbalancing of the conditions meant that there was no bias towards sleepiness or distraction. This Latin-squares design was repeated for alert and sleepy conditions for the same participants.

<table>
<thead>
<tr>
<th>Subjects</th>
<th>No Distraction PVT</th>
<th>Distraction MATB</th>
<th>No Distraction MATB</th>
<th>Distraction PVT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-6</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>7-12</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>13-18</td>
<td>3</td>
<td>4</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>19-24</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

3.2.3 Laboratory Settings

The laboratory setup, described in Section 2.3.1, was calibrated to accommodate two subjects simultaneously for PVT study with concomitant levels of distraction from a television. An experimenter was always present in the room on the experimental side to monitor compliance to the task and to make sure that recording devices functioned properly for the duration of the task. The EEG recording software was altered to permit monopolar recordings from a common ground with the montage C4-A1 and C3-A2 to measure central brain derivations, and two EOG channels. The common ground chosen was centred on the forehead, and the areas A1 and A2 provide the reference for C4 and C3 respectively. The references (A1 and A2) were chosen because it is a part of the body that is close to being electrically stable and provides a sound area to determine electrical potential differences from the cortex. The common ground electrode serves to remove artefact and noise from 50Hz sources. Prior to each task a three-minute relaxed EEG baseline recording was obtained, whereby participants looked at their blank monitors and had to remain silent with their eyes opened.

Due to the recording being stored in a computer, the signal was digitized first via an analogue-to-digital converter and passed through an anti-aliasing filter with a sample rate of 128Hz. To pre-process artefact from radio waves, movement artefact or other sources of nuisance frequency, a high-pass filter eliminated frequencies that occurred less than 0.3Hz caused by noise, eye rolling and muscle artefacts. Similarly, a low-pass filter eliminated signals above 30Hz, which can often occur from electromyographical origins. Because of the low-pass filter, a notch filter for 50Hz frequencies typically associated with power lines was not required. Post-processing was conducted by manual visual inspection of each recording and removing parts that constituted obvious distortion that was typically caused either by movement artefact or electrode fault (e.g. electrode becomes detached from the scalp due to accidental removal by participant). Where there was uncertainty, epochs were removed if they exceeded the mean plus two standard-deviations, which is a measure of outlier influence first proposed by Cook (1977). The epoch length was four seconds, which is the rate at which the fast-Fourier transform updates the frequency domain. Time-domain data was transformed by a fast-Fourier analysis so that the frequency-domain could be
assessed, allowing comparison of alpha (8-13Hz) and theta (4-8Hz) bandwidths to evaluate our hypotheses. This technique is called the “power spectrum analysis” and discards time information, but represents the constituent waveforms in one summated graph. Because all EEG measures are relative to the three-minute baseline measure prior to each trial, any sources of constant interference are naturally partialled out.

3.2.4 Scoring Distractions

Distractions were not scored during the experimental night as they required two raters to be present, so instead they were gleaned from the recorded video data at a later date. This was also necessary as all lapses were printed out so that they could be accurately seen and tallied in the correct locations for the scoring metric (Appendix B). Thus a time-segmented quantification of both lapses and distractions were obtained for all participants in all conditions, collected in thirty-second bins. Head turns were categorised as being either short (< 3s) or long (>3s), which provided the simplest and most reliable procedure during pilot testing. Lapses were categorised as three types: being caused by a distraction (‘head turn’), eyes open (where the eyes were open and engaged for the duration of the lapse) or eyes closed (where the eyes were closed for the majority of the lapse, but did not include blinking which were classed as eyes open). This deceptively simple system was a powerful indicator of performance in pilot testing, and unambiguously managed to categorise every type of lapse without difficulty. DVD recordings were rewound at parts where disagreement between the raters occurred until a consensus was arrived at. A PVT analysis program was written to analyse the results as the original program was not accompanied by suitable outputting software.

3.2.5 Subjective Scales

Various scales were given prior to testing to measure stable qualities of the participants. The first scale was the EPQ-R, which gives an indication of extraversion, psychoticism and neuroticism. Subjects were also given the State-Trait Anxiety Index to complete to measure their normal levels of anxiety (‘trait’) and how they felt during the testing night (‘state’). The PVT includes as part of the performance battery a visual-analogue scale of subjective sleepiness, which is a measure from 1-10 of sleepiness with 1 as the least sleepy and 10 as the most.

Results

3.3.1 Descriptives

Twenty-four subjects aged around 23.17yrs (s.d. = 3.6yrs) participated in the study, split into twelve males and twelve females. Their average scores on the EPQ-R were 16.8 (± 4.3) for extraversion, 5.3 (± 2.5) for psychoticism, and 10.8 (± 5.4) for neuroticism. Their state and trait anxiety levels averaged 32.8 (± 7.8) and 35.3 (± 7.2) respectively. Split-half analyses of these personality dimensions were explored in greater detail to determine whether subgroups
performed differently based on their questionnaire scores. All effect sizes (generalized eta-squared) presented throughout this chapter were calculated according to the formulas given by Olejnik and Algina (2003) and Bakeman (2005) for repeated-measures designs. They should be interpreted as being $\eta^2 = 0.01-0.05$ small, $0.06-0.13$ medium and greater than 0.14 as large (Cohen, 1988).

### 3.3.2 Sleepiness And Fatigue

Figure 13 shows a monotonic increase in both mean response time and raw lapses (reaction times greater than 500ms) as the number of PVT sessions increase. Sessions 1 and 2 represent the two night sessions (10pm-12am), while sessions 3 and 4 take place four hours later between 4am-6am. The x-intercept of the raw lapses has been shifted ten units up so that the similarities between the two measures can be more clearly seen. This relationship can also be seen with the fastest and slowest reaction times across the night and into the morning, where Figure 14 shows the gradual increase in fastest reaction times (10th percentile) compared to the slowest (90th percentile) which increases most as sleepiness is at its highest. Percentiles are best for showing the fastest and slowest reactions as they remove outliers and better represent the average of the two tails than simply looking at absolute best and worst performances. Session 1 represents the first PVT undertaken, Session 2 is the second, Session 3 is the third overall session but the first under sleep pressure, and Session 4 is the final PVT. All error bars throughout this thesis represent standard deviations.

![Figure 13: PVT mean reaction time and lapses across sessions. Sessions 1 and 2 took place between 10pm-12am, while sessions 3 and 4 took place between 4-6am.](image-url)
Similar to Figure 13 and Figure 14 which show a linear trend of lapses and slowness of response, the subjective sleepiness ratings show a similar linear increase with hours awake (Figure 15). The ten-point scale was simply marked "Sleepiness" and set at the mid-point, and no markings were shown on the scale. The average rises from 6.13 around 10am, which deviates little from the centre point, up to 9.50 with most subjects rating at the top end of the scale (point 10) around 6am. There is a direct relationship between subjective sleepiness and performance on the PVT, where subjective sleepiness correlates extremely highly ($r^2 = 0.449$) with performance lapses (Figure 16).
Figure 14: PVT mean reaction time for fastest and slowest reactions. Variability in results increases as fatigue and sleepiness accrue, although increases in the slowest reaction times becomes most noticeable.

Figure 15: Subjective sleepiness across trials from 10am-6am. Subjective sleepiness increases across the sessions, showing support that perceived sleepiness is the factor responsible for performance degradation.
Figure 16: Correlation between lapses and subjective sleepiness. A high correlation ($r^2 = 0.449$) is found between transformed lapses that reduce individual variance and worst performances (lapses).

When each of the PVT metrics is aggregated for the night session and the morning session the gross differences can be viewed (Figure 17). Percentage differences show that the largest magnitude of differences are in the slowest reaction times (243% larger when sleepy than alert) and number of lapses (nearly 2.5 times more lapses), while the smallest are with the fastest reaction times (10% extra of the magnitude of alert reactions) and errors of commission (15% more when sleepy). The median performance serves to eliminate the extremes from both ends of the reaction time spectrum, showing that average performance is really not much worsened when sleepy but it is the extremes that are affected e.g. the mean. These results demonstrate that all reaction metrics are affected by sleepiness even after a four-hour break.

The number of short or long head turns that occurred when alert, irrespective of whether distraction was shown or not, was found to nearly double ($5.79 \pm 9.92$ to $10.33 \pm 17.11$). A paired t-test showed that this difference was highly significant ($t(23) = 3.98, p < 0.05$), indicating that sleepy subjects seek distractions even when distraction conditions are collapsed. Large variations were caused by a floor effect, where some subjects were relatively impervious to distraction and had no head turns at all during the whole PVT testing.
A two-way within-subjects ANOVA (sleepiness x distraction) showed a clear main effect of sleepiness on lapses ($F_{(1,22)} = 36.34$, $p < 0.05$, $\eta^2 = 0.42$), average reaction time ($F_{(1,22)} = 49.77$, $p < 0.05$, $\eta^2 = 0.50$), the 10th percentile ($F_{(1,22)} = 8.98$, $p < 0.05$, $\eta^2 = 0.26$) and the 90th percentile ($F_{(1,22)} = 76.28$, $p < 0.05$, $\eta^2 = 0.48$). These effect sizes are very large and go to show that the sleepy sessions were vastly impaired compared to the alert session. It is interesting that the fastest RTs achieve experimental significance even though in percentage terms they appear not to move far from the no change marker in Figure 17.

3.3.3 Distraction

A stacked column graph of overall distraction results for long and short head turns during PVT time on task are shown in Figure 18 for each distraction level. There is little evidence that distractions increase with time-on task, other than the first five minutes which appear to have a marginally smaller density of head turns than forthcoming 5-minute bins. The vast majority of short head turns occur during distraction conditions (623 instances), whereas the no distraction condition is numerically close in short head turns (92 instances) to the long head turns in the distraction condition (97 instances). Clearly, subjects were prone to distraction during the PVT even when the television was not switched on, although dwarfed in number by the times when the television was on. By comparison, long head turns in no distraction conditions barely existed (3 instances) showing that serious distractions- even under high sleep pressure- can be averted in a sterile environment.
When sleepiness is collapsed and only distraction is considered, the effects of distraction can be appropriated as in Figure 19. Comparison with Figure 17, which showed the effect of sleepiness on the PVT metrics, shows a clear difference between the effects of sleepiness and the effects of distraction. The metrics snake around the 'no change' line but the largest change, that of lapses, managed to only differ by 8% compared to the no distraction condition. By comparison, the average number of head turns was over 509% increased in the distraction (37.25) compared to the no distraction condition (7.31) yet this did not manifest itself as a measureable performance difference due to subject strategizing of the task.
Sleepiness and Distraction 63

Figure 19: Percentage differences between the distraction levels. There is only a negligible change that does not exceed chance.

Two-way within-subjects ANOVA (sleepiness x distraction) demonstrates this clear lack of effect, where main effects of distraction for lapses ($F_{(1,22)} = 1.51$, $p > 0.05$, $\eta^2 = 0.02$), average reaction time ($F_{(1,22)} = 0.58$, $p > 0.05$, $\eta^2 = 0.01$), $10^{th}$ percentile reaction time ($F_{(1,22)} = 8.37$, $p > 0.05$, $\eta^2 = 0.04$), and $90^{th}$ percentile ($F_{(1,22)} = 3.70$, $p > 0.05$, $\eta^2 = 0.08$) all showed small effect sizes. Linear correlation of mean reaction time and head turns in no distraction and distraction conditions showed no relationship whatsoever ($r^2$ both at zero). The question of whether a state instability occurs with distraction is shown in Figure 20, where the total of head turn lapses and head turns occurring in distraction and no distraction conditions for all subjects is graphically demonstrated. Individual plots showed no clear support either way, with random and systematic changes being impossible to dissociate. Plausibility for this theory and results are discussed in the Discussion.
Figure 20: State instability for distraction and head-turn lapses. Peaks in distraction should coincide with peaks in head turn lapses occurring rhythmically to support state instability, and not arise from random variation.

3.3.4 Sleepiness x Distraction

Figure 21 shows over 2700 lapses categorized by time and behavioural correlate for all of the subjects. There is little evidence here that performance degrades significantly beyond the first five minutes, as the number of lapses appears to reach a stable peak after this point. The vast majority of lapses occurred with the eyes opened (n = 1839) compared to eyes closed (n = 692) or head turns (n = 193), although in the morning the relative percentage of eyes opened lapses to eyes closed starts to reduce compared to alert (690% down to 197%). The effects of distraction (over 500% of the no distraction head turns) start to manifest themselves as higher quantities of head-turn lapses, although still comparatively low compared to the other two behavioural measures. Effects of sleepiness are apparent- there are 1907 lapses when sleepy, and only 817 lapses when alert. The effects of distraction are less striking, with 1281 lapses in the no distraction condition and 1441 lapses in the distraction condition. Intriguingly, the distraction conditions appear to have an interactive effect from these graphs that is also corroborated in the inferential analyses.
In the two-way within subjects ANOVA utilized in the previous two sections, the interaction term confirms that there is a small interaction for lapses ($F(1,22) = 6.37, p < 0.05, \eta^2 = 0.06$). This is not the case for any of the other metrics, which have small effect sizes ranging from $\eta^2 = 0.00$ for the 90th percentile and $\eta^2 = 0.03$ for the mean reaction time. Along with Figure 21 above, there is a small case that can be made for lapses interactive with levels of sleepiness and distraction, but the median reaction time does not support this (Figure 22), so this interesting result may only be indicative of a trend. Transformed lapses ($\sqrt{\text{lapses} + \sqrt{\text{lapses}+1}}$) are useful when there is a large dispersal of data and eliminates variance problems caused by empty cells (Dinges and Kribbs, 1991), and makes sense when looking at medians.
Sleepiness and Distraction In The PVT

Figure 22: Boxplots of the transformed lapses per condition. The interaction effect of sleepiness and distraction has a low effect size.

3.3.5 Lapse Durations

Until this point a count of lapses was used or mean reaction times as is convention in the literature, but the magnitude of a lapse should also reflect a behavioural response. As was shown in the previous section, the number of lapses does increase with sleepiness along with our expectations. However, when the nature of the lapse is split a relationship like that found in Figure 23 is shown. Despite there being three times as many eyes open lapses than eyes closed, the difference between the magnitude of the two is the clear difference between a serious performance impairment (mean = 2996.6ms when alert, mean = 3636.2ms when sleepy) and an insignificant moment (mean = 689.2ms when alert, 715.4ms when sleepy). So-called 'outliers' are not removed from any analyses, unless they are demonstrably of technical fault and not of behavioural origin, because it is these extreme results that often provide the most insight into what can go wrong when sleepiness has its hardest impact.
Now that a difference between the two measures (EO vs EC) has been verified, the next step is to model the differences using the majority of the 2700 lapses with head turns (of which there were less than 200) omitted. A moving average can be created by ordering each lapse in terms of magnitude, and then expressing each lapse as a percentage of the previous one hundred lapses that occurred with the eyes open. The reason for this is so that the relationship is constantly updated to accommodate the next lapse, and gives a plot accurate to 1% resolution. Least-squares regression techniques can then be applied to the resulting curve to create a function fit. The results of this technique are shown in Figure 24, which shows a decaying exponential relationship between lapse magnitude and percentage chance of eyes being open during that lapse.
Figure 24: Graphical plot of lapse magnitude against percentage chance that eyes are open. The formula can be used to predict a percentage chance that eyes are open at any given reaction time.

Figure 25: The same graph as Figure 24, but with a major outlier removed. The same relationship is found with small differences for the 5% EO and 50% EO/EC (12ms and 132ms respectively) cutoffs, but an appreciable difference for the 95% EC probability (388ms).
As can be seen, the equation fits the curve extremely closely and can be used to estimate the predicted (percentage EO) from the predictor (lapse magnitude) for this study. A summary of the major landmarks on this curve is provided in Table 4. The midpoint where percentage chance EO and EC bisect each other is found around 1300ms, while the two alpha tails are found just under 600ms (95% chance EO) and near 3100ms (95% chance EC). It should be noted that more than a third of all lapses occurred within the 500-600ms limits. However, later analysis revealed that only one subject (number six) accounted for a vastly disproportionate amount of EO lapses above the 50% sample EO cutoff. A re-analysis that omitted this subject revealed the following changes in Table 5, with a particular reference to the difference at the 95% EC confidence level (3128ms down to 2740ms). These effects are not seen for any other subjects, as shown later in Table 6.

Table 4: Lapse duration and quartile breakdown for the sample of 24 subjects.

<table>
<thead>
<tr>
<th>Lapse Duration (ms)</th>
<th>Description</th>
<th>% Chance Eyes Open</th>
<th>% Chance Eyes Closed</th>
</tr>
</thead>
<tbody>
<tr>
<td>596</td>
<td>5th Percentile</td>
<td>95</td>
<td>5</td>
</tr>
<tr>
<td>3128</td>
<td>95th Percentile</td>
<td>5</td>
<td>95</td>
</tr>
<tr>
<td>901</td>
<td>Lower Quartile</td>
<td>75</td>
<td>25</td>
</tr>
<tr>
<td>1342</td>
<td>MIDPOINT</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>1982</td>
<td>Upper Quartile</td>
<td>25</td>
<td>75</td>
</tr>
<tr>
<td>500-901</td>
<td>All other lapses</td>
<td>Av. 94.1</td>
<td>Av. 5.9</td>
</tr>
</tbody>
</table>

Table 5: Lapse duration and quartile breakdown when the major outlier, subject no. 6 is removed. The largest change is seen at the 95th percentile.

<table>
<thead>
<tr>
<th>Lapse Duration (ms)</th>
<th>Description</th>
<th>% Chance Eyes Open</th>
<th>% Chance Eyes Closed</th>
</tr>
</thead>
<tbody>
<tr>
<td>608</td>
<td>5th Percentile</td>
<td>95</td>
<td>5</td>
</tr>
<tr>
<td>2740</td>
<td>95th Percentile</td>
<td>5</td>
<td>95</td>
</tr>
<tr>
<td>840</td>
<td>Lower Quartile</td>
<td>75</td>
<td>25</td>
</tr>
<tr>
<td>1210</td>
<td>MIDPOINT</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>1780</td>
<td>Upper Quartile</td>
<td>25</td>
<td>75</td>
</tr>
<tr>
<td>500-840</td>
<td>All other lapses</td>
<td>Av. 94.1</td>
<td>Av. 5.9</td>
</tr>
</tbody>
</table>

This formulation can be applied to each subject to demonstrate application and predictability of the model. Table 4 showed the sample measures and was accurate to 1% resolution when used as a moving-point average of all lapses in ascending order mapped to their respective reaction-times. For each subject, then, the formula given in Figure 24 is applied to each reaction time (x) and associated percentage probability of either EO or EC is listed (y). If the actual behavioral correlate was the same as predicted by the model, then the model was scored as correct. If however the lapse was the opposite as predicted (e.g. EC for RT < 1330ms at the 50% level), then the model was scored as incorrect. A simple index from 0-1, with 0 meaning complete inaccuracy and 1 meaning perfect accuracy, was then produced based upon the predictions.
All of the lapses were used for the 50% confidence level, defined as whether a lapse below 1342ms was EO or above 1342ms was EC being correct and the opposite incorrect. At the 95% confidence levels, EO lapses were those reaction times between 500-596ms, and EC for those reaction times above 3128ms. Where there were only one or two reaction times for EC lapses at both the 50% and 95% levels the predictions were not computed because of an inability to separate the predictions from chance and because these binary values greatly distort the variance. All lapses between 596-3128ms are therefore discarded at the 95% levels in order to illustrate the method. Of course, Figure 24 shows a continuous distribution and these cutoffs are merely shown as being indicative of the potential to apply such methods where cutoffs may be desirable e.g. in cases where video recording technology is unavailable.

Using just a 50% split on the data provides an 83% predictive accuracy for EO lapses and 78% for EC. At the 95% confidence levels, predictive values increased to 93% predictive validity for EO lapses and 90% for EC lapses. The same moving-average technique was then applied to each subject with a window of ten reaction times as opposed to 100, giving a 50% cutoff accurate to 10%. This 50% midpoint represents the point in the data where the probability of an EC lapse exceeds the probability of an EO lapse, and the values found are given in the column “Individ. (ms).” This moving-average technique is especially useful when the shape of the distributions is not known or is not likely to be linear. One notably extreme outlier, subject six, was removed from the average given for the individual 50% cutoffs. It should be noted that although there are no counts of EC at the 50% or 95% level for the model fit predictions for subjects 22 and 23, this does not mean there were no EC lapses at all, just that they occurred when EO would be predicted. The “No.” column represents the total number of lapses that contributed to the prediction e.g. if EC = 0.97 and No. = 178, then there must have been 173 EC lapses (173/178 = 0.971) and 5 EO lapses (5/178 = 0.029). Table 6 shows the results from the original model that used all subjects, whereas Table 7 shows the modified results when subject no. 6 is removed. Table 6 shows why this subject should be removed from this analysis as his 50% chance of EO was at 5458ms, which is almost a full four seconds larger than the next highest 50% EO level. The influence from this outlier can be seen in the drop from the overall probability of an EC lapse.
Table 6: The decaying exponential model for all subjects is applied to individuals and shown for predictability. The 95% confidence limits were highly predictive of EC and EO lapses, and the 50% confidence limits performed well. The final column shows the midpoint in milliseconds where the probability of an EC lapse exceeds that of EO for each subject. Subject six is removed from the average of the individual midpoints.

<table>
<thead>
<tr>
<th>Subject</th>
<th>EO No.</th>
<th>EC No.</th>
<th>EO No.</th>
<th>EC No.</th>
<th>Indiv. (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>50%</td>
</tr>
<tr>
<td>1</td>
<td>0.79</td>
<td>140</td>
<td>0.97</td>
<td>178</td>
<td>0.94 52 1.00 127 1156</td>
</tr>
<tr>
<td>2</td>
<td>0.82</td>
<td>122</td>
<td>0.93</td>
<td>61</td>
<td>0.95 44 1.00 25 1012</td>
</tr>
<tr>
<td>3</td>
<td>0.91</td>
<td>187</td>
<td>0.98</td>
<td>124</td>
<td>0.98 58 1.00 59 1290</td>
</tr>
<tr>
<td>4</td>
<td>0.97</td>
<td>68</td>
<td>-</td>
<td>1</td>
<td>0.98 41 - 1 1610</td>
</tr>
<tr>
<td>5</td>
<td>0.82</td>
<td>100</td>
<td>0.77</td>
<td>13</td>
<td>0.98 49 1.00 4 932</td>
</tr>
<tr>
<td>6</td>
<td>0.99</td>
<td>231</td>
<td>0.34</td>
<td>56</td>
<td>1.00 108 0.44 18 5458</td>
</tr>
<tr>
<td>7</td>
<td>0.93</td>
<td>29</td>
<td>-</td>
<td>1</td>
<td>0.95 20 - 1 910</td>
</tr>
<tr>
<td>8</td>
<td>0.46</td>
<td>24</td>
<td>-</td>
<td>1</td>
<td>0.38 8 - 0 923</td>
</tr>
<tr>
<td>9</td>
<td>0.98</td>
<td>128</td>
<td>-</td>
<td>1</td>
<td>1.00 76 - 0 1478</td>
</tr>
<tr>
<td>10</td>
<td>0.97</td>
<td>63</td>
<td>0.75</td>
<td>12</td>
<td>1.00 31 - 1 1392</td>
</tr>
<tr>
<td>11</td>
<td>0.62</td>
<td>47</td>
<td>0.94</td>
<td>32</td>
<td>0.86 14 1.00 7 940</td>
</tr>
<tr>
<td>12</td>
<td>0.86</td>
<td>108</td>
<td>0.83</td>
<td>6</td>
<td>0.94 52 - 1 943</td>
</tr>
<tr>
<td>13</td>
<td>0.50</td>
<td>102</td>
<td>-</td>
<td>0</td>
<td>0.50 74 - 0 907</td>
</tr>
<tr>
<td>14</td>
<td>0.79</td>
<td>34</td>
<td>-</td>
<td>0</td>
<td>1.00 11 - 0 867</td>
</tr>
<tr>
<td>15</td>
<td>0.60</td>
<td>68</td>
<td>0.96</td>
<td>51</td>
<td>0.88 26 0.96 25 960</td>
</tr>
<tr>
<td>16</td>
<td>0.96</td>
<td>24</td>
<td>0.60</td>
<td>5</td>
<td>1.00 11 - 1 1469</td>
</tr>
<tr>
<td>17</td>
<td>0.70</td>
<td>47</td>
<td>0.31</td>
<td>48</td>
<td>0.94 18 0.92 26 898</td>
</tr>
<tr>
<td>18</td>
<td>0.98</td>
<td>56</td>
<td>-</td>
<td>0</td>
<td>1.00 44 - 0 1100</td>
</tr>
<tr>
<td>19</td>
<td>0.93</td>
<td>158</td>
<td>0.88</td>
<td>8</td>
<td>1.00 94 - 0 883</td>
</tr>
<tr>
<td>20</td>
<td>0.72</td>
<td>193</td>
<td>0.92</td>
<td>48</td>
<td>0.96 92 0.75 4 839</td>
</tr>
<tr>
<td>21</td>
<td>0.96</td>
<td>53</td>
<td>-</td>
<td>2</td>
<td>1.00 37 - 2 984</td>
</tr>
<tr>
<td>22</td>
<td>0.76</td>
<td>75</td>
<td>0.75</td>
<td>8</td>
<td>1.00 32 - 0 1612</td>
</tr>
<tr>
<td>23</td>
<td>0.96</td>
<td>54</td>
<td>-</td>
<td>0</td>
<td>1.00 39 - 0 -</td>
</tr>
<tr>
<td>24</td>
<td>0.83</td>
<td>18</td>
<td>-</td>
<td>0</td>
<td>1.00 9 - 0 629</td>
</tr>
<tr>
<td>Average</td>
<td>0.83</td>
<td>0.78</td>
<td>0.93</td>
<td>0.90</td>
<td>1078.82</td>
</tr>
<tr>
<td>S.D.</td>
<td>0.16</td>
<td>0.22</td>
<td>0.16</td>
<td>0.19</td>
<td>238.97</td>
</tr>
</tbody>
</table>
3.3.6 Personality

Grouping variables extraversion, psychoticism, neuroticism, and state anxiety were created by splitting the sample of twenty-four subjects into two groups of twelve by a median split for each psychometric dimension. Independent-samples t-tests showed that the two groups for each of these measures were significantly different from each other, allowing for meaningful use when comparing personality types to performance.

Table 7: The decaying exponential model for all subjects is applied to individuals and shown for predictability similar to Table 6, but with the new cutoffs derived from Table 5. There is a marginal increase in predictability of the model for these new cutoffs.

<table>
<thead>
<tr>
<th>Subject No.</th>
<th>EO No.</th>
<th>EC No.</th>
<th>EO No.</th>
<th>EC No.</th>
<th>EO No.</th>
<th>EC No.</th>
<th>50%</th>
<th>95%</th>
<th>50%</th>
<th>Indiv.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.84</td>
<td>129</td>
<td>0.96</td>
<td>189</td>
<td>0.93</td>
<td>56</td>
<td>0.99</td>
<td>140</td>
<td>1156</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.84</td>
<td>116</td>
<td>0.90</td>
<td>67</td>
<td>0.96</td>
<td>48</td>
<td>1.00</td>
<td>29</td>
<td>1012</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.93</td>
<td>177</td>
<td>0.93</td>
<td>134</td>
<td>0.99</td>
<td>67</td>
<td>1.00</td>
<td>70</td>
<td>1290</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.97</td>
<td>68</td>
<td>-</td>
<td>1</td>
<td>0.95</td>
<td>43</td>
<td>-</td>
<td>1</td>
<td>1610</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.84</td>
<td>97</td>
<td>0.75</td>
<td>16</td>
<td>0.98</td>
<td>57</td>
<td>1.00</td>
<td>4</td>
<td>932</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.99</td>
<td>222</td>
<td>0.29</td>
<td>65</td>
<td>1.00</td>
<td>118</td>
<td>0.43</td>
<td>21</td>
<td>5458</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.96</td>
<td>28</td>
<td>-</td>
<td>1</td>
<td>0.95</td>
<td>20</td>
<td>-</td>
<td>1</td>
<td>910</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0.48</td>
<td>23</td>
<td>-</td>
<td>2</td>
<td>0.38</td>
<td>8</td>
<td>-</td>
<td>0</td>
<td>923</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0.98</td>
<td>126</td>
<td>-</td>
<td>3</td>
<td>1.00</td>
<td>79</td>
<td>-</td>
<td>0</td>
<td>1478</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.97</td>
<td>62</td>
<td>0.69</td>
<td>13</td>
<td>1.00</td>
<td>34</td>
<td>-</td>
<td>1</td>
<td>1392</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>0.64</td>
<td>45</td>
<td>0.94</td>
<td>34</td>
<td>0.88</td>
<td>16</td>
<td>1.00</td>
<td>11</td>
<td>940</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0.87</td>
<td>106</td>
<td>0.75</td>
<td>8</td>
<td>0.95</td>
<td>57</td>
<td>-</td>
<td>1</td>
<td>943</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>0.50</td>
<td>102</td>
<td>-</td>
<td>0</td>
<td>0.50</td>
<td>78</td>
<td>-</td>
<td>0</td>
<td>907</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>0.79</td>
<td>34</td>
<td>-</td>
<td>0</td>
<td>1.00</td>
<td>12</td>
<td>-</td>
<td>0</td>
<td>867</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0.63</td>
<td>63</td>
<td>0.95</td>
<td>56</td>
<td>0.85</td>
<td>27</td>
<td>0.96</td>
<td>28</td>
<td>960</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>1.00</td>
<td>22</td>
<td>0.57</td>
<td>7</td>
<td>1.00</td>
<td>11</td>
<td>-</td>
<td>1</td>
<td>1469</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>0.77</td>
<td>43</td>
<td>0.96</td>
<td>52</td>
<td>0.95</td>
<td>20</td>
<td>0.93</td>
<td>28</td>
<td>898</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>1.00</td>
<td>55</td>
<td>-</td>
<td>1</td>
<td>1.00</td>
<td>44</td>
<td>-</td>
<td>0</td>
<td>1100</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>0.95</td>
<td>153</td>
<td>0.85</td>
<td>13</td>
<td>1.00</td>
<td>102</td>
<td>-</td>
<td>0</td>
<td>883</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>0.75</td>
<td>183</td>
<td>0.91</td>
<td>58</td>
<td>0.95</td>
<td>95</td>
<td>0.90</td>
<td>10</td>
<td>839</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>0.96</td>
<td>52</td>
<td>-</td>
<td>3</td>
<td>1.00</td>
<td>40</td>
<td>-</td>
<td>2</td>
<td>984</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>0.81</td>
<td>70</td>
<td>0.85</td>
<td>13</td>
<td>1.00</td>
<td>34</td>
<td>-</td>
<td>0</td>
<td>1612</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>0.96</td>
<td>54</td>
<td>-</td>
<td>0</td>
<td>1.00</td>
<td>41</td>
<td>-</td>
<td>0</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>0.83</td>
<td>18</td>
<td>-</td>
<td>0</td>
<td>1.00</td>
<td>13</td>
<td>-</td>
<td>0</td>
<td>629</td>
<td></td>
</tr>
</tbody>
</table>

Average: 0.84 0.85 0.92 0.97 1078.82
S.D.: 0.15 0.12 0.16 0.04 273.60
The personality dimensions were compared to performance measures such as lapses in independent-samples t-tests, and also to measures of distraction in the form of proneness to distraction judged by head turns in distraction conditions. It was found that all personality dimensions were insignificantly different from head turns as distraction, meaning that the scales did not have predictive power for distraction. Lapses were collapsed by condition to see whether a proneness to lapsing in general could be uncovered or predicted by personality dimensions. Poor performance on the PVT is not related to any of the psychometrics listed in Table 9 as none of them achieved experimental significance. The worst performances, in the form of the 90th percentile, all returned insignificant results for personality and therefore cannot be predicted using these metrics.

Table 9: Comparing personality dimensions to distraction and performance.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Personality</th>
<th>Means</th>
<th>St.Dev.</th>
<th>t</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lapses</td>
<td>Anxiety</td>
<td>1341.4</td>
<td>1875.2</td>
<td>1.41</td>
<td>p &gt; 0.05</td>
</tr>
<tr>
<td></td>
<td>Extraversion</td>
<td>1198.4</td>
<td>1839.4</td>
<td>0.86</td>
<td>p &gt; 0.05</td>
</tr>
<tr>
<td></td>
<td>Psychoticism</td>
<td>1086.6</td>
<td>1875.3</td>
<td>0.46</td>
<td>p &gt; 0.05</td>
</tr>
<tr>
<td></td>
<td>Neuroticism</td>
<td>828.0</td>
<td>571.3</td>
<td>1.38</td>
<td>p &gt; 0.05</td>
</tr>
<tr>
<td>Distraction</td>
<td>Anxiety</td>
<td>1317.7</td>
<td>166.5</td>
<td>1.22</td>
<td>p &gt; 0.05</td>
</tr>
<tr>
<td></td>
<td>Extraversion</td>
<td>137.7</td>
<td>91.83</td>
<td>0.39</td>
<td>p &gt; 0.05</td>
</tr>
<tr>
<td></td>
<td>Psychoticism</td>
<td>28.83</td>
<td>44.40</td>
<td>0.03</td>
<td>p &gt; 0.05</td>
</tr>
<tr>
<td></td>
<td>Neuroticism</td>
<td>215.0</td>
<td>48.70</td>
<td>0.54</td>
<td>p &gt; 0.05</td>
</tr>
<tr>
<td></td>
<td>Anxiety</td>
<td>31.08</td>
<td>42.12</td>
<td>0.23</td>
<td>p &gt; 0.05</td>
</tr>
<tr>
<td></td>
<td>Extraversion</td>
<td>31.08</td>
<td>42.12</td>
<td>0.23</td>
<td>p &gt; 0.05</td>
</tr>
<tr>
<td></td>
<td>Psychoticism</td>
<td>31.08</td>
<td>42.12</td>
<td>0.23</td>
<td>p &gt; 0.05</td>
</tr>
<tr>
<td></td>
<td>Neuroticism</td>
<td>31.08</td>
<td>42.12</td>
<td>0.23</td>
<td>p &gt; 0.05</td>
</tr>
</tbody>
</table>
3.3.7 EEG Analysis

For distraction measures, fast-Fourier transform of the 30-minute PVT session was compared relative to the 3-minute baseline recording to determine which frequencies are dominant when participants completed the task. These baseline-corrected graphs for each participant in each condition are given in the appendix for C4-A1 and C3-A2 recordings, whereas summaries are provided here. Of a possible 192 comparisons (24 participants x 4 conditions x 2 recording sites), 14 were omitted from analysis because of either data corruption or excessive artifact. Sleepiness measures used a comparison of baseline recordings for pre-test sleepiness judgment as measured by the EEG, and baseline-corrected measures for task sleepiness. All individual EEG measures relative to baseline are given in the appendix (n = 178).

Because EEG measures are subject to large inter-subject variation in both magnitude and direction, it is useful to firstly describe results in a nonparametric manner. Figure 26 graphically demonstrates whether there was a simple increase or decrease in activity for alpha and theta bandwidths compared to baseline for C4-A1 and C3-A2 combined. As can be seen, there is no conclusive deduction that can be made for sleepiness, but there is a trend for an increase in alpha and theta in distraction conditions.

![PVT EEG Bandwidths](image)

**Figure 26:** Comparison of global alpha and theta differences between distraction or sleepiness conditions. If a bandwidth increased in potential difference it was scored as “plus”, otherwise it was scored as “minus.”

Individual frequency bands were binned per 1Hz interval and shown as an increase for a subject, or a decrease. These 1Hz bins are a combination of signals taken at a 0.25Hz frequency. Figure 27 shows the individual bins and simple increases or decreases for all subjects in distraction or no-distraction conditions, while Figure 28 shows the same type of graph for alert and sleepy fast-Fourier transforms. If activity recorded at a site, as measured by potential difference (root-mean square of voltage), was larger for distractions then it was scored as “plus”, otherwise it was scored as “minus” (e.g. activity was higher in the no-distraction condition). Potential difference is changed to the root-mean square value as the direction of potential has no real meaning in EEG power-spectrum analysis, as it is only the magnitude that is important.
Figure 27: Each individual frequency bin is shown as either increasing ("Plus") or decreasing ("Minus") for all subjects for the two EEG recording sites. An increase represents higher voltage in distraction conditions, whereas a decrease represents lower voltage brain activity recorded by the electrodes in distraction conditions compared to no distraction.

Figure 28: Each individual frequency bin is shown as either increasing ("Plus") or decreasing ("Minus") for all subjects for the two EEG recording sites. An increase represents higher potential in sleepy (morning) conditions, whereas a decrease represents lower potential brain activity recorded by the electrodes in sleepy conditions compared to non-sleepy conditions.
As was shown earlier in the results, the subjective sleepiness values were significantly different between the alert and sleepy sessions. To gauge whether the EEG baseline recordings approximated the subjective values given on a global level the baseline alpha (8-13Hz) and theta levels (4-8Hz) are shown in Figure 29. An increase in alpha would be expected in low cognitive workloads, especially given that the subjects were required to stare into a blank monitor. On a global level, Figure 29 does not support this hypothesis. Theta levels are similarly indifferent between alert and sleepy recordings on a global level, although theta increases are best used as indicator of cognitive workload during the task.

![PVT Baseline Measures (C4-A1)](image)

![PVT Baseline Measures (C3-A2)](image)

Figure 29: Global alpha and theta levels for the 3-minute baseline recordings for the C4-A1 and C3-A2 recording sites.

Table 10 displays the average subjective sleepiness and alpha levels for sixteen subjects. Eight subjects could not be used as the subjective sleepiness ratings were missing for one of the sessions. Like Figure 26, there is little congruence in alpha direction. Twelve of the subjects reported an increase in subjective sleepiness between alert and sleepy as well as showing an increase in alpha corresponding to this, although increases are minimal for some subjects. Five subjects reported higher subjective sleepiness between
alert and sleepy while exhibiting lowered alpha levels, while one subject reported higher subjective sleepiness when alert but had highest alpha when sleepy. This subject, number six, was also the major outlier in other analyses when applying the decaying exponential formula, and so proved to be a generally unpredictable person.

Table 10: Baseline alpha levels for C4-A1 recordings for a subsample of sixteen subjects.

<table>
<thead>
<tr>
<th>Sub</th>
<th>Alert</th>
<th>Sleepiness</th>
<th>Sleepy</th>
<th>Sleepiness</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.847</td>
<td>7</td>
<td>1.302</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>1.107</td>
<td>8</td>
<td>1.562</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>1.640</td>
<td>6.5</td>
<td>1.793</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>1.551</td>
<td>7.5</td>
<td>1.523</td>
<td>9.5</td>
</tr>
<tr>
<td>5</td>
<td>0.888</td>
<td>9.5</td>
<td>1.009</td>
<td>10</td>
</tr>
<tr>
<td>6</td>
<td>1.521</td>
<td>9</td>
<td>1.549</td>
<td>8.5</td>
</tr>
<tr>
<td>7</td>
<td>1.549</td>
<td>6</td>
<td>1.206</td>
<td>7.5</td>
</tr>
<tr>
<td>8</td>
<td>1.950</td>
<td>6</td>
<td>1.649</td>
<td>7</td>
</tr>
<tr>
<td>9</td>
<td>0.890</td>
<td>9.5</td>
<td>0.745</td>
<td>10</td>
</tr>
<tr>
<td>10</td>
<td>1.145</td>
<td>6.5</td>
<td>1.581</td>
<td>9</td>
</tr>
<tr>
<td>11</td>
<td>0.856</td>
<td>6</td>
<td>0.981</td>
<td>9</td>
</tr>
<tr>
<td>12</td>
<td>1.187</td>
<td>6.5</td>
<td>1.501</td>
<td>10</td>
</tr>
<tr>
<td>13</td>
<td>1.018</td>
<td>5</td>
<td>0.969</td>
<td>9</td>
</tr>
<tr>
<td>14</td>
<td>1.198</td>
<td>8</td>
<td>1.337</td>
<td>9</td>
</tr>
<tr>
<td>15</td>
<td>1.320</td>
<td>6</td>
<td>1.348</td>
<td>10</td>
</tr>
<tr>
<td>16</td>
<td>1.411</td>
<td>5</td>
<td>1.370</td>
<td>10</td>
</tr>
</tbody>
</table>

To determine whether theta or alpha bandwidths can predict performance an uncomplicated within-subjects measure can be made by ranking individual performance data with corresponding ranks of baseline-corrected theta and alpha levels in ascending order. If the performance rank (mean reaction time in this instance for each condition) matches the alpha or theta ordered data, then a “1” is scored, otherwise a zero (mismatch denoted by “-”) is given. This method 'flattens' the magnitude and instead focuses on direction, which is a stable quality of exploratory data analysis (Tukey, 1976).

Table 13 shows the results from this nonparametric analysis. A score of 6 represents chance results: which is exactly what is found for the theta bandwidths for both hemispheres. Since the nonparametric results are exactly at chance levels, there would be no value in pursuing parametric measures or inferential statistics.

However, the alpha bandwidths give attractive results for the distraction conditions (ND and MD), and even to a lesser-extent the NN and MN conditions. Because of the already-mentioned large intersubject variations, there is no useful parametric measure that can be used to test within-subjects values. When the effect size \( r \) is calculated for this data as either the nonparametric Spearman's rho or the parametric Pearson product-moment
correlation, artificially-inflated values are found due to the small number of 
data points per case (3-4) for both statistics as well as the obvious 
unsuitability of parametric measures on non-binomially distributed data for the 
latter, and so these are not presented in the results but individual results can 
be found in the graphs shown in Appendix D.

Chi-Square cross-tabulations ($\chi^2$) were performed to see whether the 
expected frequencies matched the observed frequencies. Normally this could 
simply be a case of observed frequencies being above or below 6 as 
comparison, but the missing data means that expected frequencies are not 
always 6 for each category. The $\chi^2$ values are then compared to the $\chi^2$ 
frequency table with alpha set to 0.05 or 0.01. Since there are always four 
categories, the degrees of freedom are three and the $\chi^2$ significance cutoff is 
7.82 for $p < 0.05$ and 11.34 for $p < 0.01$. Using the four categories stabilizes 
any spurious results that can occur when examining individual categories. 
Table 11 gives the summary of these results:

**Table 11**: Chi-square cross tabulations for the two recording sites and associated 
significances. Alpha rankings show highly significant results compared to theta, 
which are a poor indicator of performance.

<table>
<thead>
<tr>
<th></th>
<th>C4-A1</th>
<th>C3-A2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Theta</td>
<td>Alpha</td>
</tr>
<tr>
<td>$\chi^2$</td>
<td>1.49</td>
<td>28.49</td>
</tr>
<tr>
<td>Significance</td>
<td>$p &gt; 0.05$</td>
<td>$p &lt; 0.01$</td>
</tr>
</tbody>
</table>

The interhemispheric correlations are tabulated below for the baseline-
corrected conditions and averaged for all subjects by z-score transformation. 
Pearson's product-moment correlations between the C4 (right) and C3 (left) 
hemispheres for all 1Hz bins showed high accordance with each other for 
each subject, so only the grand mean index is given in Table 12.

**Table 12**: Grand interhemispheric correlations per condition for all subjects. A matrix 
of bins per subject for the C4 cortex region was compared to the matrix derived from 
the C3 scalp location. $R = \text{correlation coefficient}, \ N = \text{number of data points}.$

<table>
<thead>
<tr>
<th>Condition</th>
<th>r</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alert, No Distraction</td>
<td>0.644</td>
<td>354</td>
</tr>
<tr>
<td>Alert, Distraction</td>
<td>0.553</td>
<td>419</td>
</tr>
<tr>
<td>Sleepy, No Distraction</td>
<td>0.709</td>
<td>418</td>
</tr>
<tr>
<td>Sleepy, Distraction</td>
<td>0.795</td>
<td>424</td>
</tr>
</tbody>
</table>
Table 13: Ranked performance comparisons between PVT mean reaction time in each condition and the ordered ranks of alpha and theta levels during the task for both hemispheres. "1" means perfect agreement, "-" means disagreement, and "**" represents those trials that were eliminated from the analysis because of excessive artifact or data corruption. Theta results are similar to levels of chance, whereas alpha results show higher accordance with the performance data. NN = Night, No Distraction, ND = Night Distraction, MN = Morning, No Distraction, MD = Morning, Distraction.

<table>
<thead>
<tr>
<th>Sub</th>
<th>NN</th>
<th>ND</th>
<th>MN</th>
<th>MD</th>
<th>NN</th>
<th>ND</th>
<th>MN</th>
<th>MD</th>
</tr>
</thead>
<tbody>
<tr>
<td>C4-A1</td>
<td>RT Theta Rank</td>
<td>RT Alpha Rank</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C3-A2</td>
<td>NN</td>
<td>ND</td>
<td>MN</td>
<td>MD</td>
<td>NN</td>
<td>ND</td>
<td>MN</td>
<td>MD</td>
</tr>
</tbody>
</table>

Total 7 3 4 6 11 13 7 15
Results from this chapter lend strong support to the body of evidence that has linked PVT degradations to sleepiness alone, although the distraction findings are less concrete and the lack of effects may be an artefact of a coping strategy to the task that is discussed later. The video analysis used throughout the results of this chapter have provided a unique insight into the nature of sleepiness as it actually affects performance without being reliant on making backwards-inferences based purely on trial results. Real-time performance as it occurred was systematically logged and stored alongside EEG activity, generating nearly 50 hours of video footage and over 50 hours of EEG recordings. To date this is the most ambitious and focused examination of behaviour during this popular task, and from the results an entirely new way to view the PVT output has been borne that makes meaningful use of what unprepared reaction time changes genuinely reflect. "What laws govern the occurrence of these blocks," (p. 230) as Bill's (1931) sought to understand, at one stage has to account for why subjects close their eyes when they should otherwise be open, and so now we tentatively move towards a more scientific approach to answering Bill's early research questions.

The case for increased PVT lapses being reflective of the processes that govern sleep, which impact on other factors such as motivation as a by-product, is strongly and unequivocally shown in this study. The worst performances in the form of lapse count and 90th percentile (magnitude) both show a nearly 150% change between the alert and sleepy conditions, whereas the 10th percentile (fastest RTs) showed only a marginal change in percentage terms. Lapses also correlated strongly with subjective sleepiness and explained almost half of the variance, indicating that subjects knew when they were likely to be impaired by their increasing feelings of sleepiness. While sleepy, subjects sought distraction or respite from the task even where there were no distractions present; a difference between alert and sleepy conditions that was statistically significant. It appears that even with a four-hour break between testing sessions to help eliminate boredom or motivational effects, bearing in mind that most employment setups would not offer such a large interval between work, the fact that sleepiness was still a factor shows that the effects of sleepiness may be even worse in harsher circumstances. That so many lapses occurred with the eyes closed, even when subjects knew they were being observed via the cameras, shows that vigilance to a task may be involuntarily affected by sleep, whether the subject is aware of the potential danger or not.

By contrast, the effects of distraction are less clear and when interindividual differences are accounted for there is no strong case that any PVT metric is significantly affected by distraction. There is no doubt that increased frequency of head turns, both short and long, occur when the television was switched on, but there is no evidence that these head turns caused a measurable impairment on performance. This may be due to a loophole in the PVT design that subjects quickly learned to adapt to. The interstimulus interval varies between 2-12 seconds on a typical PVT trial, with 2 seconds
being the lowest possible setting. So while subjective sleepiness increased, most subjects learned to sneak a quick glance at the television screen in the period immediately after a number had appeared on screen and then had looked back at the task before the next had arrived. On the occasions where they had misjudged the interval (perhaps the number presented at the lower end of the scale nearer 2 seconds), a head-turn lapse was recorded. As there were only 193 of these head-turn lapses for all subjects out of a total of 2724 lapses, it seems that this strategy was certainly effective. The random nature of the interstimulus interval adds a layer of uncertainty to any unprepared reaction time study, as it is mostly a matter of chance in real-life whether a failure of attention matches up to a moment where attention is required e.g. when driving and a child runs out onto the street. Subjects sought distraction when they were tired, and became increasingly unable to avoid the lure of watching the television- a process of 'perseveration' seen in similar studies (Anderson and Home, 2006). The volume of head turns greatly increased, and although the interaction effect was small, the increase in head turns when sleepy and distracted may cause a precise moment of inattention that causes an accident.

A feature of the video analysis breakdown was that while the number of lapses certainly increased with sleepiness (Figure 21), the nature of these lapses also changed. There were very few eyes-closed lapses during the night (eyes open lapses were ~7 times greater at night to eyes closed, compared to only ~2 times during the morning), but in the morning the number and percentage preponderance increased remarkably. It should be stressed here that during an eyes-closed lapse the subject was completely helpless to respond to whatever was shown on screen, and many times when the subject had their eyes closed this didn’t result in a lapse due to the varying and random nature of the PVT. The average eyes-open lapse is almost inconsequential in real terms at 715.4ms when sleepy, whereas the average eyes-closed lapse averaged almost three seconds (3128ms or 2740ms with outlier removed). To put this into perspective, the average eyes-open lapse would result in a lack of response of 12.8metres in a car travelling at 60mph, compared to 67.8m for eyes-closed. These means also substantially underestimate the real effect of an eyes-closed lapse, which often far exceeds three seconds and is at the mercy of the subject regaining self-control in a timely manner. Even were the lapses to be equal in duration, having the eyes closed is a much more serious proposition than the eyes opened and something that is rarely related to PVT performance with the common assumption for long periods of unresponsiveness being caused by eyes-open microsleep.

To this end the decaying exponential relationship between lapse magnitude and percentage chance of eyes being open clearly reflects a reliable biological process. A third of all data occurred between 500-600ms, yet these so-called lapses result in greater than 95% chance of eyes being open. By contrast, lapses above 2740ms are more than 95% likely to have occurred with the eyes closed. It is obvious that most lapses that last, for example, ten seconds would not be likely to occur when the eyes are open, the same way a 600ms reaction time is not likely to be caused while the eyes are closed.
Linking these two limits gives a clean decaying exponential relationship that can be modelled using the equation given in the results. Further research will uncover how generalizable this relationship is to different populations under different methods of sleep loss.

By using two equal time periods between 10am-12pm and 4am-6am an even curve of nonsleepy and sleepy performance is given, but a qualification exists that the slope of this curve will vary with experimental procedure used. Biasing sleepiness in either direction (more or less sleepiness) will load the results towards either direction, but our design has hopefully provided equal timeframes for alert and sleepy performance to give a fair overall picture balancing the extremes of sleepiness. Our results are only extendable to young populations, and while a decaying exponential is very likely to be replicable in other populations, the exact values will naturally differ. The midpoint of around 1200ms should serve as a meaningful cut-off that reflects a different aspect of sleepiness than simply using 500ms, which delimits possibly more interesting and world-relevant findings.

Overall, the EEG results were difficult to interpret due to the large individual differences. On a sample-wide basis, there is no clear evidence that any particular frequency bands are implicated in the condition manipulations of sleepiness and distraction (Figure 27 and Figure 28), given that the increases in activity are almost proportional to the decreases in activity. However, when analysed individually the alpha bandwidth finds accordance with the performance results, being highly significant ($p < 0.01$) for both hemispheres. This indicates that the alpha frequency can be used as a predictor of performance in the PVT when analysed relative to a subject's own performance, which is a finding in agreement with related research. Teplan called the alpha rhythm “the best-known and most extensively studied rhythm” (p. 3) in sleep research because of its ability to provide a reliable performance indicator. The theta bandwidth did not predict performance in the PVT, although it could be argued that the PVT is not cognitively challenging enough to show as a change in theta activity. Within-subjects analysis is far more useful for EEG than generalising across all subjects.

That the personality dimensions had no predictive value for performance is of little consequence and is likely a product of the comprehensive screening techniques designed to pool together subjects who are as similar as possible to each other. While split-half analysis did produce two statistically-significantly different groups, they were not different enough in real-terms to have any explanatory power. Studies utilising bipolar extremes of these scales will be more likely to elucidate effects than one designed to minimalise individual differences. At the same time, the state instability theory of vigilance was explored but the fluctuations in results are simply not dissociable from random variation. Certainly peaks in lapses, either in no distraction or distraction conditions, do not map satisfactorily to any behavioural measure.

The sheer volume of head turns during the PVT gives some indication that the task is unchallenging, boring and in requirement of compensatory measure. It is typical for PVT subjects to be locked alone in a soundproofed booth, where
it would be all too tempting to relax and seek stimulus from outside the task itself. For this reason it is impossible to know what a genuine lapse might look like when analysing results unless video data is taken at the same time. An excerpt from Drummond et al. (2005) highlights the inadequacies of trying to retroactively make inferences looking at raw data alone:

"Dinges and colleagues reported that, following TSD, lapses (e.g., slow RTs of at least 500 milliseconds) are often commingled with false starts (pressing the response button in the absence of a stimulus or making a response faster than 100 milliseconds). They argue that such errors of commission are anticipatory and reflect subjects' compensatory efforts in response to the feedback that they made a very slow response." (p. 1065)

We would suggest that video playback of these lapses would show participants, with their eyes closed, pressing the response button randomly to ensure compliance with the test (a simple and often-used strategy in this task), especially following total sleep deprivation when the drive for sleep is at its highest. The errors of commission are, of course, inevitable as these random button presses often precede any stimulus being on screen - rather than being any compensation to knowledge of poor results. Even when subjects knew they were being video-recorded in this study they would often 'steal sleep' by closing their eyes for a short period of time, while pressing the response button every few seconds to feign participation on the task. It would not be surprising if many so-called microsleeps were rather more indicative of subjects retreating into their mind to provide stimulus in the same manner as daydreaming. This idea of a cognitive distraction is explored in greater detail in Chapter Five.

When researchers cite the use of the 500ms cut-off they tend to reference early papers that make no mention of why 500ms was chosen (e.g. Graw et al., 2002), if a citation is even attempted. That 500ms is sensitive to sleep intervention may be a historical coincidence that remains loosely fit for purpose but is in need of a scientific rethink should psychomotor vigilance continue to be the cornerstone of experimental sleep studies. Our proposition for a 1200ms cut-off as a further division of lapses, along with 3000ms for severe losses in performance, provides a greater insight into true sleepiness as it borders onto the sleep-initiating process and is clearly more representative of sleepiness than 500ms, which often elucidates results even when subjects are demonstrably non-sleepy. The advent of video and computer technology makes sleep and sleepiness directly observable and shareable, so the time for behaviourally justifying many measures of sleep and sleepiness is long overdue. The eventual take-up of our recommendations will remain to be seen, but should this area of research begin to move towards empirically useful measurements then the tools and methodologies employed need to reflect the phenomena they pertain to represent.
Research Questions Summary

1. What extent does psychomotor performance change under low and high sleep pressure?

Psychomotor performance is greatly affected by increasing sleepiness. The slowest reaction times show an increase in frequency and magnitude with higher subjective sleepiness. Lapses that exceed 3s appear to be part of the sleep-initiating process as the drive for eye-closure increases with increased sleepiness.

2. Does performance or attention to a task change behaviourally under increased sleep pressure?

There is a strong tendency for people to seek distractions under conditions of increased sleepiness. However, these distractions do not always result in a task-long decline in performance, but may result in spontaneous and unprepared loss of performance. Some individuals seem highly prone to distractions, while others appear practically immune to increased distraction under high sleep-pressure conditions. This makes it difficult to generalise across a population when discussing distraction effects.

3. What is the mathematical relationship between sleepiness and distraction?

There is a clear behavioural shift for lapses in attention, as lapses become progressively caused by the eyelids being shut with increasing duration. The relationship is nonlinear, and specifically is a decaying exponential curve. The parameters of this curve are liable to change with study design and sample characteristics, but for this study it was found that reactions above three seconds were 95% likely to have occurred with the eyes closed. Faster reaction times (< 600ms) are 95% likely to have occurred with the eyes open.

The results of this chapter have formed these publications:

Anderson, C., Wales, A.W.J., and Home, J.A. Redefining the PVT lapse in terms of duration, and whether eyes are open or closed. Sleep, in submission.

Anderson, C., Wales, A.W.J, and Home, J.A. Sleepiness increases lapses due to distraction. Fatigue Management in Transportation Operations Conference, accepted for publication.

CHAPTER FOUR
Multi-Tasking And Distraction Under Sleep Pressure

Abstract

The Multi-Attribute Task Battery (MATB) was used to investigate the effects of overnight sleepiness and distraction in 24 young (m = 23.17yrs) subjects on an engaging task. There were no significant effects of sleepiness on any of the MATB components with effect sizes ranging from $\eta^2 = 0.00$ for dials reaction time to $\eta^2 = 0.06$ for lights lapses. Subjects were prone to short and long head turns in distraction conditions (382 and 153 instances respectively), but this did not show as a significant main effect of distraction for any MATB component. Vigilance for the MATB was very high, given that there were only 11 instances of short head turns and no long head turns (> 3s) at all during no distraction conditions. Time-on-task analysis showed no performance decrement between the ten-minute bins. Subjects were significantly quicker to attend to changes in the status of the lights than the dials, which are more difficult to spot. It can be concluded that interesting or otherwise engaging tasks such as plant control or systems monitoring are impervious to 24-hours of sleep deprivation, and are not likely to be seriously affected by peripheral distractions.
4.1.1 Resourceful Work

In between the monotonous and predictable vocations detailed in Chapter Three and highly engaging vocations that are relatively untouched by sleep research (e.g. those of lawyers or academics) there are vocations that require constant attention, but at the same time require flexibility of thought and ability to manage subtasks concurrently. Vocations that fit this description include supervisory positions, plant control, systems monitoring, piloting, and multi-screened security monitoring, amongst many others. The highest-profile disasters relating to sleepiness tend to have involved a mixture of the before-listed vocations, with failures to a devastating effect. The Three Mile Island, Chernobyl, Bhopal, Challenger, Exxon Valdez and the Estonia ferry disasters (Folkard, Lombardi and Tucker, 2003) all took place during the night shift with reviews and official investigations concluding that fatigue and human error combined, at least in part, to the causation of those disasters. While vocations of this nature are unequivocally less impacted by sleep deprivation than lower-order vocations (Hockey, Wastell and Sauer, 1998; Webb and Levy, 1984), the statistics produced from various fields such as the aviation industry and military prove that sleepiness can be and is a factor in performance ability.

4.1.2 Military Operations

Fighting in the military is a twenty-four hour a day, seven day a week job that is life-dependant on being fit to operate and react to sudden actions at any time of the day (Caldwell and Caldwell, 2005). Technical advances in modern warfare like night-vision devices and night-fighting capabilities of ground and air troops make total warfare possible and often desirable at times when the enemy may be nocturnally subdued or unable to cope with the element of surprise (Caldwell et al., 2003). Indeed, one of the strengths of the US military is their ability to co-ordinate surging night attacks to literally “shock and awe” their opposition- as evidenced in recent military coups. One of the qualities of working in the military is that personnel are expected to work until the job is done, allowing small windows where sleep can be recuperated but often the mentality of working in the military means that sleep is seen very much as a luxury than a necessity (Naitoh, 1992).

Patton et al. (1989) reviewed the literature and presented a multitude of sustained operations studies that had used prolonged combat-simulated scenarios. They advised that these studies had worked on the assumption that future conflicts would be “characterized by high intensity operations lasting for periods exceeding an individual’s capability to maintain efficient performance” (p. 1). While most sleep studies concentrate on cognitive impairments there is little evidence to suggest that aerobic power or ability to perform at high levels of physical ability should be impaired by sleepiness. Only two studies had shown a decline in muscular strength or anaerobic capacity (Legg and Patton, 1987; Murphy, Knapik and Vogel, 1984). They too concluded that soldiers given five hours of sleep per day could still perform at high physical levels for up to eight days of continuous performance, but that
the body of sleep literature suggests that there would still be a detrimental cognitive effect.

Strategies to dealing with fatigue in the military tend to take the view supported by the wider field of sleep research but with a parsimonious spin. Sleep management technical papers have shown that power napping of three to four hours per day has been trialled in Marines (Naitoh, Englund, and Ryman, 1983), prophylactic napping in normal subjects at the US Army Medical Research Centre (Dinges et al., 1980), and sleep logs have been examined from Naval officers (Naitoh et al., 1973). Yet despite the research there still remains a general ignorance in the top ranks of military operations. Lauber and Kayten (1988) pulled no punches in their biting riposte to the lax attitudes held by ill-informed generals and air marshals. An excerpt from their excellent keynote address reads:

"The many physicians in this audience ought to be familiar with the phrase, "When I was a resident...", which is always followed by some heroic tale of marathon sessions in the ER (or someplace). I recently heard a former Federal Air Surgeon make such a statement during a discussion of flight and duty time issues.

Or consider the following except from an article in the Washington Post (1987),

The [Nuclear Regulatory Commission] said it found a pattern of sleeping or inattentive operators [at the Peach Bottom Nuclear Power Plant]... especially on the 11 pm to 7 am shift, when the control room is staffed by a skeleton crew... Running a nuclear power plant at full power is largely an automatic operation, but workers are expected to continuously monitor gauges... and they must be alert to abnormal "trends" on gauges... and they must be prepared to handle sudden emergencies"

Lauber and Kayten, 1988, p. 511, original emphasis

4.1.3 Pilot Studies

Pilot fatigue in civilian and military flight operations is a real and significant problem in the aviation industry given the rise of 24/7 flight routines, long duty periods, circadian disruptions and insufficient sleep being commonplace (Caldwell, 2005). Rapid timezone transitions and jet lag (Mohler, 1966) place sleep demands on the modern pilot that are worse than any shift rotation, given the irregularity and changeable nature of transcontinental flights. Rosekind et al. (1994) present anecdotal and observational evidence that sleep does in fact occur on the flight deck despite federal regulations to the contrary. These researchers found that flight schedules caused circadian desynchrony for crew members, being unable to adjust to rapid time-zone shifts. Fatigue in aviators has been linked to lowered response accuracy, speed and forgetting or ignoring of important flight tasks (Perry, 1974). Furthermore, a group of international sleep and aviation experts concluded that fatigue caused by sleepiness and tiredness is the "largest identifiable and preventable cause of accidents in transport operations (between 15 and 20%
of all accidents), surpassing that of alcohol or drug related incidents in all modes of transportation” (Åkerstedt, 2000, p. 395).

Estimation of the actual cause of sleepiness’ impact in aviation accidents ranges from around 4-8% (Kirsch, 1996; Luna, 2003), with human error accounting for upwards of 80% of all commercial and military aircraft accidents (Wiegmann and Shappell, 2001). A quarter of all accidents between 1974-1992 logged by the Air Force’s night tactical fighter Class A were found to be caused by fatigue (Ramsey and McGlohn, 1997). In a breakdown of transport industry costs, Lauber and Kayten (1988) revealed that an astonishing $900 million in insurance payouts were made the previous year alone for worldwide aviation accidents. They quoted estimates that one single major airline accident count incur $500 million in total losses, and that an internal study by Boeing found flight crew error was responsible for 65.5% of all accidents since their jets began operating. Human error is a common theme throughout the statistics produced by flight reports (Latorella and Prabhu, 2000). Given the importance of security and costs attributed to the aviation industry, managing sleep and sleepiness is growing in importance more and more.

4.1.4 Multi-Tasking And Divided Attention

Multi-tasking is a more general term for divided or shared attention processes, which involve being able to allocate mental resources and the development of time-sharing skills (Wickens and Kramer, 1985). The central dogma in performance management has been to claim that the brain acts as a single-communication channel of limited capacity (Allport, Antonis and Reynolds, 1972), where parallel performance is seemingly achieved by flicking between tasks moment to moment. On top of being able to handle sub-components of a task, successful operations require selective attention too, which means prioritising sensory channels and ignoring irrelevant information or distractions that are not conducive to performance (see Wickens, 1991). When distractions are similar in characteristics to the task in hand then their effects will be attenuated (Lavie and Cox, 1997), particularly if these distractions are irrelevant stimuli of the task (or mathematically, noise). Also if distracters are of a different sensory modality than the task then they are less liable to affect performance- much like listening to radio communications while surveying radar, as opposed to listening to two radio channels concurrently.

Total sleep deprivation can substantially reduce performance on certain divided attention tasks by impairing the ability to switch tasks or flexibly alter thought to account for incoming task-critical information (Fisher, 1980). In turn subjects adapt by focusing on one subtask to the detriment of the other. Researchers have made the argument that driving itself is a multi-component task, requiring a combination of visual and auditory attention and control of various simultaneous components (e.g. junctions require gear shifts, indication and wheel manipulation in quick succession) during challenging conditions (Juniper et al., 2000). In these conditions driving can be impaired by moderate sleep loss (Home and Reyner, 1995), and although incidence of sleep-related vehicle accidents do occur during the day they are not as
numerous as during the evening when road conditions are less challenging. Motorway driving at night is more likely to resemble the automated processes detailed in Chapter Three and is more affected by sleep deprivation due to the soporific conditions of low light and low stimulus.

However the more difficult a task is, or the more subcomponents it has, the more difficult it becomes to achieve repeatable, reliable, and ecologically valid results (Jones and Harrison, 2001). Home (2000) made the point that tests do not have to be monotonous and boring to be sensitive to sleep loss, although these tests have become the norm throughout the sleep literature.

4.1.5 Multiple Component Tasks

The Multi-Attribute Task Battery (MATB) was originally conceived to be a simplified cockpit simulator accessible to novices to test and quantify various interventions in a controlled environment (Arnegard, 1991). The MATB has a steep learning curve but once subjects are experienced on the task then it becomes a case of quietly managing the subcomponents as performance approaches automaticity. In this respect, the MATB quickly becomes like many of the systems or operational vocations outlined at the start of this chapter, where subjects monitor and alter the system to take care of unpredictable changes (Comstock and Arnegard, 1992). Where the MATB differs from tasks designed to engage the subject is that it is not a truly divergent task once training has been given. Subjects quickly develop strategies to deal with shifts in each subcomponent (Arnegard, 1991) and once a successful strategy has been formed there is little further expense required of the higher-level cognitive systems. Studies that have specifically used the MATB have concluded that task-specific effects of human automation occur with practice, such that cessation of new strategies are formed when existing strategies seem sufficient (Carmody, 1994; Molloy and Raja, 1996).

Sleep deprivation has seldom been manipulated using this task, but the few studies that have used it have shown that time on task effects occur. Caldwell and Ramsport (1998) have shown that interesting and engaging tasks like the MATB have to be used for at least half an hour to see a sleep effect. This is because the more difficult or enjoyable a task is the less likely that sleep effects will deteriorate performance (Wilkinson, 1964). In a study using twelve actual pilots, Caldwell et al. (2001) showed that MATB indicators of cognitive skill showed effects with sleep loss, which was validated by EEG theta and delta increases. Workload effects by manipulating the MATB difficulty settings are seen in upper alpha EEG and eye blink rate (Fournier, Wilson and Swain, 1999), in that lower difficulty levels are indicative of lowered physiological arousal. Of course these are laboratory findings; it could be that performance in the field is likely to be altered with successive days of sleep deprivation or that the real-life consequence of failing to take action to warning signals will affect performance.
4.1.6 Psychophysiology Of Sleepiness And Task Load

Brain responses to demanding tasks under sleep deprivation are not always the same as when simple tasks are used, and often background brain EEG can be used to predict the perceptual load experienced by a subject. Increased working memory load is typified by increases of power in the beta bandwidth and a *phasic* suppression of alpha (Gevins and Smith, 2003). This makes intuitive sense if alpha increase is to represent levels of arousal. A study using the MATB to predict psychophysical responses found that sustained response to task demands was shown by a reduction of parasympathetic inhibition, reduced eye blink durations, and increase of theta activity at parietal sites (Fairclough, Venables and Tattersall, 2005). Increased theta activity at the frontal lobe is also indicative of higher cognitive workload (defined as 4-7.5Hz, Klimesh, 1999), and indeed working memory is typically linked to a functional network that connects the prefrontal cortex with posterior association cortices (Owen, Evans and Petrides, 1996). Patients with frontal lobe damage are unable to plan appropriately and have lost many executive functions such as prioritising tasks and selectively attending (Russo *et al.*, 2005). So with cognitively-challenging tasks that require executive functions then examination of the frontal lobe by EEG should provide an insight into how well that area is being recruited to deal with executive issues, with topographical changes in alpha indicating type of information being processed (Scerbo *et al.*, 2001).

4.1.7 Summary

The MATB was created to address the pervasive issues and challenges that are posed by the most important transport sector in terms of security and profile- the civilian airlines and military air force. Fatigue is known to be a factor in performance incidents, but the subcomponent sleepiness is nearly impossible to quantify using retrospective accident reports. Seeing that so many of these accidents are purely caused by human error (or as one academic noted, pilot error), and that sleepiness impacts human cognitive ability to function, it's clear that fatigue and pilot performance must be related. While civilian airline pilots may suffer sleep fragmentation, it is unlikely that the majority will be expected to endure sleep deprivation much greater than 24-hours. Similarly, plant and control operators may experience similar levels of sleepiness on a routine basis when working overnight shifts. For this reason we have used the 24-hours of wakefulness routine and not an extreme constant-routine approach to reflect sleepiness as it is most often encountered in real life.

*Research Questions*

1. To what extent is performance impacted in overnight work on an engaging task?

2. Is the sleepy mind prone to distraction when the task is inherently rewarding and interesting?
3. What are the differences between monotonous tasks and engaging tasks with respect to the impact that sleepiness has on them?

Methods

4.2.1 Study Protocol

The study design and protocol was identical to the study as described in Chapter Three (section 3.2.2), with only the task (MATB vs PVT) and method of scoring distractions being different. The same subjects were used to ensure that comparisons between the two tasks were fair. The EEG settings remain unchanged from Chapter Three, and all task EEG measures are relative to the three-minute baseline (change from baseline). As the MATB is a multi-component task it is impossible to estimate which aspects of the task are affected by the eyes being open or even head-turn distractions, as there may be several components that require attention but are not greatly affected by a short head-turn away from the monitor. Piloting suggested that the short/long distraction distinction was still useful to measure the impact of distraction, but that analysing behavioural correlates to performance was impossible.

4.2.2 MATB Settings

The MATB package contains two one-hour scripts as standard, but these scripts are unbalanced for ten-minute blocks and piloting showed differences in difficulty during different parts of the scripts. Therefore new scripts were created to balance the probabilities and occurrences of all events per ten minute block (see Section 2.1.2). The tracking task was removed by automation as it was found to elevate the difficulty level of the MATB to a point where learning effects would have been substantial, and also the tracking task turned the MATB from a relatively quiescent exercise in systems monitoring to a highly demanding task that drastically reduced the attention paid to the other components. Therefore the components left were a systems monitoring component, a resource task and a communications task. As the communications task was unable to run on modern computers, it was rebuilt using the original program specification manual to authentically reproduce this aspect of the task.

Participants were instructed to primarily allocate their attention towards the resource task, where various tanks and pumps have to be manipulated to create an ordered balance of 2500 units. Periodically some of the pumps will 'break', to which subjects need to develop strategies to compensate for the discord in the system. Running at the same time as the resource task, the systems monitoring task requires subjects to attend to random changes occurring within the four dials and the two lights at the top-left of the task. The dials will deviate within one-unit up and down but corrective action needs to be taken when they begin to deviate beyond one unit, characterised by a short 'jump' that can sometimes be seen out the corner of the participant's eye. Changes in the lights are more conspicuous as they alter colour suddenly and like the dials failure to respond within fifteen seconds results in
the lights results in the system recalibrating to normal automatically and a 'lapse' is logged in the results output. The communications task is actually very simple as participants only need to alter the frequency and band of a radio switch when their unique call-sign is given. Rather than being a useful dependent variable the communications task was used to force mental resource allocation to listen and attend to signals while the two main tasks were still running, and is thus not analysed in the results. The dependent variable for the central resource task was deviation in units from the target of 2500, while the dials and lights (red and green) have three dependent variables: reaction time to attend to changes, lapses (or missed signals) and false alarms (altering a light or dial that was not dysfunctional). A summary of the dependent variables is given in Figure 30:

**Figure 30:** The dependent variables used in the MATB. Tank deviations from 2500 units (root-mean square) are combined for A and B. Dials are used for lapses (response not within 15s of deviation, e.g. F1 is out of synch in the figure) and reaction time, while the red and green lights are used for lapses and reaction time also.
Results

4.3.1 Data Collection

Three MATB data trials out of a total of ninety-six were removed from analysis due to a technical fault that had corrupted the data.

4.3.2 Sleepiness

Sleepiness effects for each MATB component were deduced by collapsing the distraction blocks so that just alert and sleepy sessions could be compared. A two-way within-subjects ANOVA (sleepiness x distraction) elucidated insignificant main effects of sleepiness for the resource task ($F(1,22) = 1.23, p > 0.05, \eta^2 = 0.03$), dials lapses ($F(1,22) = 0.21, p > 0.05, \eta^2 = 0.01$), dials reaction time ($F(1,22) = 0.10, p > 0.05, \eta^2 = 0.00$), lights lapses ($F(1,22) = 3.77, p > 0.05, \eta^2 = 0.06$) and lights reaction time ($F(1,22) = 1.71, p > 0.05, \eta^2 = 0.02$). This lack of difference can be clearly shown in Figure 31, where the differences between the alert and sleepy sessions on each component only manage to exceed a 60-40% differential for the red light false alarms and lapses. This type of graph is a 100% stacked column graph where 50% represents no change at all between the two values, and as the influence of one value (e.g. alert RT for dials) increases the other (e.g. sleepy RT for dials) decreases proportionally. The results from the two lights, red and green, had been combined but Figure 31 suggests that the red light may be more prone to the effects of sleepiness. A follow-up paired t-test on the lapses for the red light for alert and sleepy showed a highly insignificant result ($t(23) = -0.816, p > 0.05$). This was because the red light was so conspicuous that very few subjects managed to miss one, such that only six subjects lapsed on the red light and three of those subjects lapsed less than three times over two hours of testing. Therefore the difference seen in Figure 31 is an artifact of a lack of observations giving rise to misleading results.
Figure 31: Comparing dependent variables of the MATB for sleepiness. Lapses for the red light show the largest effect, but it is still a statistically insignificant effect of sleepiness. 50% represents no change between the alert (night) and sleepy (morning) conditions. FA = False Alarm.

Reaction time changes across the time periods are shown in Figure 32. Sessions 1 and 2 were the alert sessions, whereas 3 and 4 were the sleepy sessions, such that the first MATB of the night was Session 1, and the second MATB was Session 2 etc. This is useful to determine whether there is an effect of task exposure rather than a main effect within the sleepiness conditions. There is no measurable difference as the sessions went on through alert (1 and 2) to sleepy (3 and 4). Differences in the resource task show a steady decrease in ability to maintain the tanks at the correct level as sleepiness increases, but the large individual variations and small absolute values conclude that no change in this component of the task occurs either during the task or with increasing fatigue and sleepiness. When taken together these two figures show no graphical trend that could have suggested that sleepiness is a factor in this task, nor does time on task have any impact on the central component.
Figure 32: Reaction times for dials & lights, and resource deviation by session e.g. session 1 would be the first MATB, session 2 is the second etc. The lack of change with increasing fatigue and sleepiness is shown for the dials and lights reaction times.

Figure 33: Time on task for the resource deviations for all four MATB trials, between alert and sleepy sessions. Resource deviations do not fluctuate significantly either during the task or with successive sessions.
4.3.3 Distraction

Figure 34 shows a stacked column graph of all head turns that took place during MATB trials. The vast majority of these head turns are short head turns that took place during the distraction condition (382 instances), followed by long head turns in distraction conditions (153 instances). There are remarkably few head turns during no distraction conditions, with only 11 short head turn bins during no distraction for all twenty-four subjects undertaking a total of 24 hours combined MATB testing without the television. As a marker of vigilance, there was not a single instance of subjects looking away from the MATB screen for longer than three seconds when there were no distractions. There is little support here that tendency for distraction increases much after the first five minutes of the task.

Figure 34: MATB head turn bins with stacked head turn frequencies. Short distractions (<3sec) dominate during the task in distraction conditions, followed by long head turns (>3 sec) in distraction conditions.

When distraction is collated across sleepiness conditions the impact of an enjoyable peripheral stimulation is shown in Figure 35. Red and green light lapses are combined to account for the dearth of results alluded to in the previous section, as well as reaction times. Most of the components differ little between distraction and no distraction conditions being close to 100% similar between the two properties. Although lapsing for lights may appear large there were not enough lapses for lights to be able to make much of an inference.
Although the lapse component appears largest, it is an artefact of a small numbers of lapses. Only reaction time for lights genuinely approaches real significance.

Of these components, the reaction time for dials was the only one to achieve experimental significance probably due to the very low variance making a small effect statistically significant ($F(1,22) = 4.98, p < 0.05, \eta^2 = 0.07$). The rest of the components all failed to show a distraction effect, and in fact this sole distraction effect for dials reaction times was the only component to achieve experimental significance out of all the sleepiness, distraction and interactions measured in real terms (not percentages).

4.3.4 Sleepiness By Distraction

Sleepiness interacting with distraction was judged by means of two-way within-subjects ANOVA and looking at the interaction term. None of the interaction components achieved experimental significance, with the interaction between lapses for dials being the highest ($F(1,22) = 2.18, p > 0.05, \eta^2 = 0.02$) and the interactions for the resource task, dials reaction time and lapses for combined lights all showing zero effect size. Figure 36 suggests that there may be a small trend towards sleepiness as the values for resource, dials and lights are all higher when sleepy but the reaction times were transformed (by inversion) and lapses transformed ($\sqrt{lapses + 1}$) to reduce inter-subject variance for the two-way ANOVA, so this deduction has little empirical foundation.
4.3.5 Personality

Personality measures were analysed to identify individual differences on the three main MATB components. A median split of the sample had confirmed significant differences between the two (section 3.3.6). Independent-samples t-tests showed that none of the personality dimensions obtained could be used as a predictor for MATB performance. T-values, as an indication of experimental effect to error variance, are all suitably low, ranging from 0.02-1.96 and thereby suggest that no individual differences are shown as a difference in performance.
Table 14: Personality predictors of MATB performance for the split-half analysis whereby participants were median-split into two groups of high and low values.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Personality</th>
<th>Means</th>
<th>St. Dev.</th>
<th>t</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Anxiety</td>
<td>74.3</td>
<td>42.0</td>
<td>0.02</td>
<td>$p &gt; 0.05$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>72.0</td>
<td>52.9</td>
<td>0.21</td>
<td>$p &gt; 0.05$</td>
</tr>
<tr>
<td></td>
<td>Extraversion</td>
<td>70.9</td>
<td>43.2</td>
<td>0.27</td>
<td>$p &gt; 0.05$</td>
</tr>
<tr>
<td>Resource</td>
<td></td>
<td>76.0</td>
<td>53.7</td>
<td>0.60</td>
<td>$p &gt; 0.05$</td>
</tr>
<tr>
<td>Deviations</td>
<td>Psychoticism</td>
<td>70.5</td>
<td>39.5</td>
<td>0.27</td>
<td>$p &gt; 0.05$</td>
</tr>
<tr>
<td></td>
<td>Neuroticism</td>
<td>79.8</td>
<td>61.3</td>
<td>0.60</td>
<td>$p &gt; 0.05$</td>
</tr>
<tr>
<td></td>
<td>Anxiety</td>
<td>7.15</td>
<td>1.31</td>
<td>0.82</td>
<td>$p &gt; 0.05$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6.57</td>
<td>1.90</td>
<td>0.58</td>
<td>$p &gt; 0.05$</td>
</tr>
<tr>
<td></td>
<td>Extraversion</td>
<td>7.02</td>
<td>1.31</td>
<td>0.58</td>
<td>$p &gt; 0.05$</td>
</tr>
<tr>
<td>Dials</td>
<td></td>
<td>6.61</td>
<td>2.03</td>
<td>0.58</td>
<td>$p &gt; 0.05$</td>
</tr>
<tr>
<td>RT</td>
<td>Psychoticism</td>
<td>6.85</td>
<td>1.36</td>
<td>0.58</td>
<td>$p &gt; 0.05$</td>
</tr>
<tr>
<td></td>
<td>Neuroticism</td>
<td>7.54</td>
<td>1.67</td>
<td>0.58</td>
<td>$p &gt; 0.05$</td>
</tr>
<tr>
<td></td>
<td>Anxiety</td>
<td>2.93</td>
<td>1.36</td>
<td>0.58</td>
<td>$p &gt; 0.05$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.52</td>
<td>0.87</td>
<td>0.58</td>
<td>$p &gt; 0.05$</td>
</tr>
<tr>
<td></td>
<td>Extraversion</td>
<td>2.98</td>
<td>1.44</td>
<td>0.58</td>
<td>$p &gt; 0.05$</td>
</tr>
<tr>
<td>Lights</td>
<td></td>
<td>2.48</td>
<td>0.73</td>
<td>1.06</td>
<td>$p &gt; 0.05$</td>
</tr>
<tr>
<td>RT</td>
<td>Psychoticism</td>
<td>2.94</td>
<td>1.36</td>
<td>1.06</td>
<td>$p &gt; 0.05$</td>
</tr>
<tr>
<td></td>
<td>Neuroticism</td>
<td>2.90</td>
<td>0.73</td>
<td>1.06</td>
<td>$p &gt; 0.05$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.55</td>
<td>1.36</td>
<td>1.06</td>
<td>$p &gt; 0.05$</td>
</tr>
</tbody>
</table>

4.3.6 EEG Analysis

Three-minute baseline EEG recordings were stored prior to the commencement of the task to normalise the EEG data from the task itself, such that changes in brain activity at certain wavelengths would show as either increases or decreases in potential difference. Out of a possible 192 recordings for the main task for both sites C4-A1 and C3-A2 (384 separate recordings if counting the individual baseline files), 8 were rendered unusable due to a combination of excessive artefact or data corruption. None of the six missing MATB results (out of a total of 96) coincided with these eight unusable EEG recordings, therefore 14 comparisons between EEG findings and MATB performance cannot be made. Both alert, no distraction and sleepy, no distraction conditions were most affected, having three comparisons missing each, resulting in 21 comparisons for these two conditions. All data files were visually inspected for obvious artefacts, and a “mean plus two standard deviations” rule of thumb employed when there were doubts as to the origins of the activity, be it artefact or brain activity.

Simple baseline-corrected counts of increases or decreases with sleepiness or distraction at the alpha and theta bandwidths combined for both hemispheres are given in Figure 37, which shows that there is little
congruence between the conditions and these collated bandwidths. Nearly as many subjects experienced increases with alpha and theta as did those who had decreases, so there is no strong evidence against the null hypothesis (H0) which states that brain activity does not alter during sleepiness or distraction conditions. Nonparametric measures such as these reduce the influence from individuals with large changes even when corrected for baseline, as there are always large interindividual differences in EEG analysis.

![MATB EEG Bandwidths](image)

**Figure 37:** Collapsed alpha and theta bandwidths across all subjects, expressed as a simple non-parametric increase or decrease in activity from baseline. There is no case that can be made against the null hypothesis from these data.

Each individual frequency band was summed as an increase or decrease from baseline for both hemispheres across all subjects, and graphed as shown in Figure 38. An increase means that there was larger activity in the distraction condition, whereas a decrease means that there was larger activity in the no distraction condition e.g. a plus value of 14 shows that 14 subjects had the largest activity in the distraction condition, whereas a minus value of 10 shows that 10 subjects had the largest activity at the given frequency bin in the no distraction condition.

On the whole, more subjects experienced increases in potential at most frequency bins, although the relatively large number of subjects who actually experienced decreases at these bandwidths means that no conclusive underlying performance explanation appears to be at work. The upper alpha band (11-13Hz) shows the largest differences as an increase in potential, but there is no theoretical reason why this should be given that there were no differences in performance between distraction and no distraction results. Figure 37 above shows these overall alpha increases. Figure 39 shows the same results as Figure 38, but with sleepiness conditions as comparison. As there was no statistically significant performance differences in the MATB, it is no surprise that the EEG shows no conclusive differences on a global level.
Figure 38: Each frequency bin is expressed as either increasing ("Plus") or decreasing ("Minus") in value from baseline compared between distraction and no-distraction conditions. The top figure shows the data from the C4-A1 derivation, whereas the bottom figure shows the data from C3-A2. More subjects experienced increased in activity in the lower frequencies (0-5Hz), as well as upper alpha (11-13Hz) during distraction conditions.

Figure 39: Count of increases or decreases in fast-Fourier transformed EEG results per individual frequency bin across the alert and sleepy sessions. Increases represent higher brain activity during sleepy conditions, whereas decreases indicate the opposite. C3-A2 shows trends towards increased alpha similar to the distraction findings, although the same cannot be said of C4-A1.
The baseline results for the main conditions of sleepiness and distraction are given in Figure 40. There are no differences for either alpha or theta bandwidths on the baseline recordings. It should be noted that the absolute potential differences have been magnified (known as “gain”) many times over and therefore the units shown are only for comparative purposes.

![MATB Baseline Measures (C4-A1)](image)

Figure 40: Global alpha and theta levels using only the 3-minute baseline recordings for C4-A1 and C3-A2. There are no perceptible differences between the levels when averaged for all participants.

The performance constructs “lapses” and “resource”, given in Table 15 and Table 16 below, relate to the sum of lapses across dials, red and green lights for the former and the deviation in units from target resource level in the latter. When concatenated into a singular lapse construct, this measure represents the total errors that are attributable to the systems monitoring portion of the MATB tasks and thus provides an uncomplicated index to which EEG levels can be referred to. As before, the performance dependent variables were ordinally ranked and then compared to the ranks of theta and alpha changes from baseline per subject in each condition.

Our hypothesis stated that increased alpha and theta levels would provide a marker for increased drive for sleep in alpha, and increases in theta would
indicate increased cognitive workload. Where data couldn't be compared, an asterisk is shown in the table, while a 1 represents a perfect match and a dash represents an incorrect matching. The rankings for theta lapses and theta resource task results hover around chance levels, which would be a sum of 6 assuming no missing data. C4-A1 accordance averages 4.25 for lapses and an even 6 for the resource task, whereas the C3-A2 derivation averaged only 4 for lapses and 6.75 for the resource task. Using 6 as a rough indicator for chance results, the theta changes can be rejected for having no potential to delineate MATB results either as a predictor or covariate.
Table 15: Ranked comparisons between MATB lapses (missed dials and lights), alpha and theta increases. If performances coincided in rank, then a "1" is marked for accordance. Any misalignment results in a "." (counted as zero), whereas damaged data was omitted ("*"). The alpha findings are related to MATB performance, whereas theta does not exceed chance.

<table>
<thead>
<tr>
<th>Sub</th>
<th>NN</th>
<th>ND</th>
<th>MN</th>
<th>MD</th>
<th>NN</th>
<th>ND</th>
<th>MN</th>
<th>MD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>*</td>
<td>*</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>9</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>10</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>11</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>12</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>14</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>15</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>16</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>17</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>18</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>19</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>20</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>21</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>22</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>23</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>24</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>4</td>
<td>3</td>
<td>6</td>
<td>7</td>
<td>15</td>
<td>10</td>
<td>13</td>
<td>14</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>C3-A1</th>
<th>NN</th>
<th>ND</th>
<th>MN</th>
<th>MD</th>
<th>NN</th>
<th>ND</th>
<th>MN</th>
<th>MD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>*</td>
<td>*</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>*</td>
<td>*</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>11</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>12</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>13</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>15</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>16</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>17</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>18</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>19</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>20</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>21</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>22</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>23</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>24</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>2</td>
<td>5</td>
<td>7</td>
<td>2</td>
<td>14</td>
<td>13</td>
<td>12</td>
<td>12</td>
</tr>
</tbody>
</table>
Table 16: Ranked comparisons between MATB resource deviations, alpha and theta increases. The alpha findings are related to MATB performance for both EEG recording sites, whereas theta does not exceed chance.

<table>
<thead>
<tr>
<th>C4-A1</th>
<th>Resource Theta Rank</th>
<th>Resource Alpha Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sub</td>
<td>NN</td>
<td>ND</td>
</tr>
<tr>
<td>1</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>10</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>12</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>13</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>14</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>15</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>16</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>17</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>18</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>19</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>20</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>21</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>22</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>23</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>24</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>3</td>
<td>7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>C3-A2</th>
<th>NN</th>
<th>ND</th>
<th>MN</th>
<th>MD</th>
<th>NN</th>
<th>ND</th>
<th>MN</th>
<th>MD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>11</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>12</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>13</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>15</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>16</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>17</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>18</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>19</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>20</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>21</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>22</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>23</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>24</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>6</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>13</td>
<td>12</td>
<td>12</td>
</tr>
</tbody>
</table>
The alpha results for both lapses and the resource task are encouragingly congruent with each other as indicators of performance. The nonparametric accordance ranges from 10 (lapses, alert distraction) to 16 (resource, alert no distraction), averaging 13 for C4-A1 lapses and 12.75 for C3-A2 lapses. For the resource task, the accordance index averages 14.25 for C4-A1 and 12.75 for C3-A2 across all conditions. Because of missing data, chi-square cross-tabulations were used as a follow-up to determine whether these results can truly be viewed as significant.

Chi-Square cross-tabulations ($X^2$) were used to determine whether the expected frequencies matched the observed frequencies found for both lapses and deviations. These analyses suppress individual differences, and as $X^2$ doesn't measure magnitude it is ideal for EEG fast-Fourier analysis. The ranking data itself made sure that the observed results followed in the same direction (increase for both theta and alpha), so these $X^2$ results do represent significant results in the direction of our hypothesis. The thresholds being compared are 7.82 for $p < 0.05$ and 11.34 for $p < 0.01$ as there were four categories ($df = 3$). It should be noted that although theta appears related to performance for C3-A2 lapses, this in fact indicates that the results were so far below the expected frequency that they approach the 95% tail on the opposite end of the probability distribution!

**Table 17:** Chi-square cross tabulations for the two recording sites and associated significances. Alpha rankings show highly significant results compared to theta, which are a poor indicator of performance.

<table>
<thead>
<tr>
<th></th>
<th>C4-A1</th>
<th></th>
<th>C3-A2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Theta</td>
<td>Alpha</td>
<td>Theta</td>
</tr>
<tr>
<td>Lapses</td>
<td>$X^2$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Significance</td>
<td>p &gt; 0.05</td>
<td>p &lt; 0.01</td>
<td>p &gt; 0.05</td>
</tr>
<tr>
<td>Resource</td>
<td>$X^2$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Significance</td>
<td>p &gt; 0.05</td>
<td>p &lt; 0.01</td>
<td>p &gt; 0.05</td>
</tr>
</tbody>
</table>

Z-score corrected correlations were performed to determine whether there were significant similarities between the two brain regions measured. These Pearson product-moment correlations used the baseline-corrected measured for each condition averaged across subjects for each 1Hz bin. The number, N, given in Table 18 relates the number of data points used per correlation. High correlations were found for all conditions, ranging from 0.61 for sleepy, no distraction and 0.75 for alert, no distraction.

**Table 18:** Grand interhemispheric correlations per condition for all subjects. A matrix of bins per subject for the C4 cortex region was compared to the matrix derived from the C3 scalp location. $r =$ correlation coefficient, $N =$ number of data points.

<table>
<thead>
<tr>
<th>Condition</th>
<th>r</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alert, No Distraction</td>
<td>0.754</td>
<td>353</td>
</tr>
<tr>
<td>Alert, Distraction</td>
<td>0.734</td>
<td>410</td>
</tr>
<tr>
<td>Sleepy, No Distraction</td>
<td>0.611</td>
<td>394</td>
</tr>
<tr>
<td>Sleepy, Distraction</td>
<td>0.684</td>
<td>384</td>
</tr>
</tbody>
</table>
Since significant nonparametric results were found for alpha as a predictor of performance, a correlation between the lapses and alpha magnitude (comparable to Table 15) is given in Figure 41. There are significant outliers that are several orders of magnitude above the bulk of the points for both lapses and alpha activity. There is a small correlation for these points ($r = 0.17$), which is the same for the natural logarithm transformation shown in Figure 42.

![Figure 41: A parametric scatterplot of alpha activity to the number of lights lapses. There is no clear conclusion that can be drawn between alpha magnitude and number of errors on the task.](image)

![Figure 42: The same data from Figure 41 but with a logarithmic transformation of the voltage.](image)
4.3.7 Comparisons With The PVT

Subjective sleepiness was compared to a summation of both dials and lights lapses to test the association between poor performance and the impact of sleep. Compared to the PVT, there is absolutely no association between sleepiness and task degradation (Figure 43). Lapses were transformed the same way for both measures, although the difference in correlations between the PVT ($r^2 = 0.449$) to the MATB ($r^2 = 0.09$) highlight the overall lack of main effects of sleepiness in the latter.

Vigilance differences between the tasks can be shown in terms of head turns during the tasks. Figure 44 shows the total counts of head turn bins between the two distraction conditions for short and long head turns. There were 623 instances of short head turns in the PVT during distraction, compared to only 382 in the MATB. When no distractions were present, the number of head turns during the MATB is staggering low, with no long head turns and only 11 short head turns compared to 92 short head turns in the PVT and 3 long.

Figure 43: Subjective sleepiness by performance lapses for PVT and MATB. While sleepiness is a predictor of performance lapses in the PVT, the same cannot be said for the MATB where there is no association whatsoever.
Figure 44: Head turns between the PVT and MATB conditions. Apart from long head turns in the distraction condition, the PVT shows much higher levels of distraction during the task than the MATB.

Discussion

Overall results from each subcomponent of the MATB show that subjects can overcome the damaging effects of sleep loss when the task is engaging enough to keep the attention focused. There was seldom larger than a 10% difference between any of the metrics in nonsleepy and sleepy conditions, and the red light only managed to show a 15% difference in false alarms and lapses as an artefact of having such low numbers to compare. Indeed, such was the synchrony between the alert and sleepy conditions that no effect sizes exceeding the 'small' cut-off of $\eta^2 > 0.05$ for repeated-measures designs. Reaction times across the sessions stayed incredibly stable, suggesting that subjects had been brought up to their ceiling after 1hr 15mins of practice, and that the sleepiness intervention did nothing to hamper their ability for responding to dials or lights. When the overall results are split into 10-minute time on task effects, the same stable performances are seen throughout the task itself, indicating that neither sleepiness nor fatigue, boredom or motivation plays any role in the MATB outcome. The communications task, which was unanalysed due to the unchallenging nature of the task, may have served as an alerting factor that maintained performance levels. It was hypothesised that multi-tasking the communications task with the other components could cause difficulties in the early morning, but there was no evidence that this was the case as subjects managed to allocate mental resources successfully.
On top of this, subjects were remarkably vigilant during this task. Comparison with the PVT shows large differences in the tendency to look away to the screen as a behavioural correlate of attention, where the twenty-four subjects diverted their attention from the task much less than when they did the PVT (815 head turn bins compared to 546). More importantly, when there was no television subjects only diverted their attention from the television on 11 occasions during 24-hours of combined video footage. The MATB is the sort of task where subjects can afford to look away and not expect to see a major performance implication, as when dials did go out of synch the subjects knew they had up to fifteen seconds to attend to it. Short head turns in the distraction condition dominated the head turn clustered bar chart (Figure 34), but interesting to note is the relatively large number of long head turns during distraction conditions that outnumbered the PVT (Figure 44). Again, this may be a result of the experimental design. During the PVT it is not desirable to divert attention for too long as the subject would almost certainly miss a signal, whereas during the MATB when systems are balanced there are short periods of time where the subject can afford to look away for longer than three seconds without any performance implication. This has implications for real-world activities, where system monitors can become complacent with prolonged inactivity in the system, unlike a commercial driver who must always be vigilant to surroundings. As Lauber and Kayten noted from their personal experiences, this pattern of inattention brought on from unchallenging circumstances often results in a “skeleton crew” manning operations (p. 511).

The means for each component of the resource, dials and lights tasks show a tendency towards being impacted by sleepiness and distraction (Figure 36), but large interindividual variation makes any solid case untenable. Additionally, actual mean differences are relatively small and unlikely to ever cause a serious performance breakdown. The resource task differed by less than twenty units (from a starting point of 2500 units) between the two conditions that differed the most (alert distraction and sleepy no distraction), which is an average of ten units above or below the desired number; a figure that is clearly completely inconsequential. Given that there were no performance differences between sleepiness and distraction conditions, the fact that there were no personality differences doesn’t explain the lack of differences but instead supports the null hypothesis.

The MATB is an ideal tool to evaluate the sleep effects that are representative of the sort of tasks that many surveillance workers, plant workers and shiftworkers are expected to perform. These tasks are always more robust to the effects of sleepiness and this, regrettably, may deter researchers from using such tasks in their studies. Yet there is unease amongst sleep researchers who seek to move from theoretical to applied science, as Jones and Harrison warned, stating that they would “encourage a move towards greater ecological validity in future research to allow generalisations about the impact of sleep loss for the work place and in daily life” (p. 472). With that, the results from this chapter confirm a long-held suspicion that engaging tasks, even if they are not explicitly entertaining, are impervious to sleep effects at least up to twenty-four hours of sleep loss. However that is not to say that
repeated sleep-loss over several days or weeks, or extreme sleep loss as found in military combat, would not elicit an effect on a complex task like the MATB. Even when subjects rated themselves on average as being 9.5 on a sleepiness scale that runs up to 10 (Figure 15), it is not enough to significantly detract their ability to perform on this task. That the PVT results overwhelmingly showed a sleep effect even though it was counterbalanced with the MATB using the same subjects highlights a direct difference between the two.

Rather interestingly, the theta and alpha levels for individual performance showed a similar pattern to that found with the PVT. Theta levels were insignificantly related to performance for both the lapses (dials and lights) and the resource task. Conversely, the alpha increases were closely related to performance for both hemispheres and both dependent variables, resulting in four measures that were all highly significant ($p < 0.01$). This indicates that alpha increases are indeed related to impoverished performance, but the magnitude of these effects is likely to be small. It has already been shown that the MATB changes across conditions were small, so ranked data may be misleading in statistical terms because only the direction of effects is considered.

This research has provided encouraging signs that humans can cope with up to twenty-four hours of sleep loss on engaging or cognitively-demanding tasks with little performance implication. That no singular component of the MATB showed a sleep effect or any solid evidence of a distraction effect, demonstrates that subjects will attend to a task to a high level of ability when it is not so soporific as to be unchallenging and monotonous.
Research Questions Summary

1. To what extent is performance impacted in overnight work on an engaging task?
   
   Speed, accuracy and unforced errors show no change with overnight sleepiness on a multi-component task. Participants were capable of dealing with the effects of overnight sleepiness and managed to maintain high levels of performance even up to 6:30am. These results are encouraging and suggest that one-off overnight work stints will probably not result in increased danger for challenging tasks.

2. Is the sleepy mind prone to distraction when the task is inherently rewarding and interesting?

   Short head turns largely increased in the distraction conditions compared to the other conditions, but there were very few head turns that lasted longer than three seconds. Short head turns away from the task are not likely to result in any measurable performance implication for a systems monitoring task. There is no evidence of any damaging cognitive-distraction effects as a decrease in value of the dependent variables.

3. What are the differences between monotonous tasks and engaging tasks with respect to the impact that sleepiness has on them?

   There is a large difference between how the same participants performed in either task, which were counterbalanced to eliminate treatment order effects. While there was a linear relationship between subjective sleepiness for psychomotor vigilance, but absolutely no relationship between subjective sleepiness and MATB lapses. There were around 600 (50%) more short head turns for the PVT than the MATB across all conditions, suggesting that subjects sought distractions more in the less engaging and more boring task.
CHAPTER FIVE
Daytime Sleep Restriction In Luggage Search

Abstract

The x-ray screening of passenger luggage is a multi-billion dollar security industry (Poole and Butler, 2001) that is growing in importance. The requirements to work during the circadian nadir, and in distracting conditions, are typically experienced at airports, thus we assessed the implications of sleepiness and distraction on the ability to identify threat items amongst benign luggage contents. A simplified luggage search simulator was used with 12 young (m = 20.8yrs) and 12 old (m = 60.0yrs) participants for two levels of sleepiness (<5hrs and normal sleep) and three levels of distraction (none, peripheral and cognitive). Models of the practice day speed-accuracy operating curves showed that the younger group were quicker and more accurate than the older group when learning this task. The experimental days showed no effect of sleep loss, but significant effects of distraction, particularly for the older group ($\eta^2 = 0.07$ for accuracy and $\eta^2 = 0.10$ for RT). Age was found to be a significant factor in performance for detection ability ($\eta^2 = 0.19$) and speed ($\eta^2 = 0.51$). Confidence ratings revealed mostly insignificant correlations to performance, and only extraversion differed as a personality trait between the young and old groups. It is concluded that cognitive distractions do impair performance in x-ray screening, particularly for older workers, and that age plays a highly significant role in performance.
5.1.1 Changing Importance Of Luggage Screening

Since the epochal terrorist attacks of September 11th 2001 there has been renewed political and academic interest into the job of airport luggage screeners. Prior to 9/11 the locus of responsibility lay with the airlines themselves, resulting in cost-cutting procedures and fragmented responsibility (Seidenstat, 2004). Gaps between the different management groups were touted as a major weakness in the system (Poole and Butler, 2001), lacking in a coherent strategy to deal with the threat posed by modern terrorism. Poole and Butler's paper was actually published just ten days after the twin towers fell - testimony to the insatiable desire for accountability for this disaster waiting to happen. The United States Government was similarly quick to react: only five weeks later the Transport Security Administration (TSA) was born and given a $4.8billion budget, of which a massive $4.5billion was spent on aviation security and the rest for waterways, rails, highways, public transport and pipelines (Felcher, 2004). This would prove to be a significant departure from the $4million that the US Government had spent on preventing terrorist attacks prior to the events of 9/11 (Felcher, 2004). A paradigm shift had occurred that fateful day and the landscape in aviation security is now one of the most dynamic and eagerly researched areas throughout applied psychology.

The need for such expense is justified in the wake of major attacks such as the Air India attack of 1985 (329 killed), Lockerbie bombings of 1988 (270 killed), and the TWA Flight 800 disaster (229 killed and suspected terrorist involvement; Felcher, 2004). As an example, more than two thousand weapons were seized during the year 2000 alone in America (Seidenstat, 2004) and doubtless countless more since. Further to this, the TSA announced that in its first year that their screeners had intercepted more than 4.8million prohibited items including 1.4million knives, 125,273 incendiary or flammable objects, and 1,101 firearms (TSA, 2003). The Federal Aviation Industry (FAA) reported that ongoing testing between 1991-1999 resulted in progressively decreasing dangerous objects being unfound (US General Accounting Office, 2001). Yet despite these encouraging figures, airports still have significant progress to make. Undercover tests by the TSA found that a quarter of fake weapons and explosives were able to get past screeners, while television news stations managed to completely conceal bags over 70% of time and undercover federal agents showed that they could bring concealed knives onto planes with consummate ease (General Accounting Office, 2003). On whether these figures accurately represent how efficient airline security is, Felcher had these words to say after an exhaustive investigative review:

"Making one's way through a busy urban airport, it is clear that TSA employees are hard at work. What is less clear is whether what they are doing is anything more than tightly choreographed busywork that provides illusory rather than real security."

Felcher, 2004, p. 80
5.1.2 Screener Selection And Training

Every study to have examined improvements with training, whether it is on the job or in the laboratory, has found positive effects that increase with more training. This is because visual inspection of passenger bags is a challenging task where threats are generally presented with low target frequencies and objects are deliberately hidden, occluded or disassembled in order to be covertly brought onboard (Lui, Gale and Song, 2007). Most threat items are objects that are not typically encountered during every day life, such as gas sprays, improvised explosive devices or tazers (Schwaninger, 2006; Figure 45). Thus screeners need to be taught how to look for items like these because visual search depends on knowledge of target items (Graf et al., 2002), meaning that if screeners don’t know what they are supposed to be looking for then invariably they will not find intricate threat items. Adaptive computer-based training is now the accepted model for screener progress, whereby screeners are certified at many major airports for visual search and knowledge abilities until they ascertain a certain level of competence (Schwaninger, Hofer and Wetter, 2007).

![Figure 45: Effects of viewpoint (Schwaninger, 2006). A gun (a), switchblade (b) and shuriken (c) can look very different in a 2D x-ray image.](image)

Screeners are selected by different criteria depending on airport and location, which is handled by different agencies and private companies. Typically a three-phase process is initiated in the TSA: the first involves computerised test batteries and completion of formalities, and provided the applicants perform adequately here they then go onto a structured interview and undertake physical abilities (e.g. to lift baggage) and medical screening. Finally, a background security check is undertaken and an offer is given (Kolmstetter, 2003). European airports follow an approximately similar procedure except the computerised tests vary depending on the security company responsible for human resources (e.g. CASRA or QinetiQ). Once employed, screeners must undertake routine training and are constantly being tested while working by sophisticated software which superimposes fictional threat items (FTIs) into passenger luggage (Threat Image Projection; Hofer and Schwaninger, 2005). Yearly improvements with on-the-job training have shown that trained screeners outperform untrained screeners by responding
quicker and more accurately, even if untrained screeners have been working for several years in the domain of luggage search (Hardmeier, Hofer and Schwaninger, 2006). These improvements are more attributable to improvements in knowledge-based factors rather than the image-based factors discussed in the next section (Schwaninger, 2005).

5.1.3 Visual Search Abilities

Several prominent theories have linked existing visual search theories with screener performance in luggage search tasks. Knowing how and why humans direct their attention is critical in understanding the limitations that screeners face and how to overcome them with human-factored design. An example of human-factored design is the multitude of image-enhancement factors that screeners have at their disposal- twelve in total- to help pick and selectively alter the image to show only organic items, metal objects, or varieties of luminosity and colour (Michel et al., 2006). Visual props are also available in some packages to alert the screener to when the bag image may be too noisy to be sufficiently analysed by vision alone (McCarley et al., 2003; Wiegmann et al., 2006). These technologies of existing machines are set to be replaced by ever-more sophisticated methods to help overcome the obvious limitations of the human perceptual system.

Searching for and making decisions about objects is a complicated process that can be explained by an overall visual search model. At its core, a multistage model is responsible for orienting alertness firstly to conspicuous areas (Neisser, 1967) in a rapid global-search that only supports obvious pop-out characteristics like colour, size and orientation (Triesman and Gelade, 1980), which is then serially linked to a considered stage that carefully examines localised areas and makes decisions based on subjective decision thresholds (Wolfe, 1998). With specific reference to luggage screening, Schwaninger (2003) isolated three image-based factors that contribute to subjective detection difficulty: threat viewpoint, superposition, and bag complexity. A later evaluation by Schwaninger, Hardmeier and Hofer (2005) found high reliability estimates for these factors that can be estimated using computational formulas (Bolfing, Michel and Schwaninger, 2006).

Attention has been known to affect visual search performance since the late nineteenth century (Palmer, Ames and Lindsey, 1993), implying by extension that visual search performance should be affected by sleep loss, as sleep loss impacts on attentional capabilities. In fact, detection performance in simple visual search tasks are found to be unaffected by forty hours of sleep deprivation, although subjects responded slower (Gennaro et al., 2001), providing evidence towards a speed-accuracy relationship to compensate for sleep loss. Other studies have found small detection decreases (Casagrande et al., 1997) but mainly speed of performance is the first aspect to suffer under sleep pressure. It is common for more complicated tasks to return no sleep effects, as the number of studies that fail to show any sleep effect from visual search tasks have shown (Kraemer et al., 2000; Porcu et al., 1998; Williamson et al., 2001).
5.1.4 Sleepiness And Visual Inspection

Given the relative youth of applied visual inspection research there is only one study in the entire literature of radiological or luggage search that has manipulated sleepiness. Basner and colleagues (2008), utilising a sample of twenty-four volunteers (mean age = 29.9yrs, s.d. = 6.5yrs), conducted a 35-hour continuous wakefulness study where the group was split into two. The first group were awake from 9:30am-17:30pm the following day and then allowed to sleep for eight hours before retaking the tests, while the second group took their baseline tests before midnight of the previous day from testing and allowed to sleep before continuous wakefulness. The Simulated Luggage Search Task used actually mimics the X-Ray ORT used in this study very closely (see Section 2.1.3). A synopsis of this design is provided below:

<table>
<thead>
<tr>
<th>Day</th>
<th>Time of Day</th>
<th>GROUP D1</th>
<th>GROUP D2</th>
</tr>
</thead>
<tbody>
<tr>
<td>09:30</td>
<td>T1</td>
<td>T1</td>
<td></td>
</tr>
<tr>
<td>11:30</td>
<td>T2</td>
<td>T2</td>
<td></td>
</tr>
<tr>
<td>13:30</td>
<td>T3</td>
<td>T3</td>
<td></td>
</tr>
<tr>
<td>15:30</td>
<td>T4</td>
<td>T4</td>
<td></td>
</tr>
<tr>
<td>17:30</td>
<td>T5</td>
<td>T5</td>
<td></td>
</tr>
<tr>
<td>19:30</td>
<td>T6</td>
<td>T6</td>
<td></td>
</tr>
<tr>
<td>21:30</td>
<td>T7</td>
<td>T7</td>
<td></td>
</tr>
<tr>
<td>00:00</td>
<td>Sleep 0:00 - 8:00</td>
<td>Sleep 0:00 - 8:00</td>
<td></td>
</tr>
<tr>
<td>09:30</td>
<td>W1</td>
<td>W1</td>
<td></td>
</tr>
<tr>
<td>11:30</td>
<td>W2</td>
<td>W2</td>
<td></td>
</tr>
<tr>
<td>13:30</td>
<td>W3</td>
<td>W3</td>
<td></td>
</tr>
<tr>
<td>15:30</td>
<td>W4</td>
<td>W4</td>
<td></td>
</tr>
<tr>
<td>17:30</td>
<td>W5</td>
<td>W5</td>
<td></td>
</tr>
<tr>
<td>19:30</td>
<td>W6</td>
<td>W6</td>
<td></td>
</tr>
<tr>
<td>21:30</td>
<td>W7</td>
<td>W7</td>
<td></td>
</tr>
<tr>
<td>23:30</td>
<td>W8</td>
<td>W8</td>
<td></td>
</tr>
<tr>
<td>01:30</td>
<td>W9</td>
<td>W9</td>
<td></td>
</tr>
<tr>
<td>03:30</td>
<td>W10</td>
<td>W10</td>
<td></td>
</tr>
<tr>
<td>05:30</td>
<td>W11</td>
<td>W11</td>
<td></td>
</tr>
<tr>
<td>07:30</td>
<td>W12</td>
<td>W12</td>
<td></td>
</tr>
<tr>
<td>00:00</td>
<td>Sleep 0:00 - 8:00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>09:30</td>
<td>W13</td>
<td>W8</td>
<td></td>
</tr>
<tr>
<td>11:30</td>
<td>W14</td>
<td>W9</td>
<td></td>
</tr>
<tr>
<td>13:30</td>
<td>W15</td>
<td>W10</td>
<td></td>
</tr>
<tr>
<td>15:30</td>
<td>W16</td>
<td>W11</td>
<td></td>
</tr>
<tr>
<td>17:30</td>
<td>W17</td>
<td>W12</td>
<td></td>
</tr>
<tr>
<td>19:30</td>
<td>W18</td>
<td>W13</td>
<td></td>
</tr>
<tr>
<td>21:30</td>
<td>W19</td>
<td>W14</td>
<td></td>
</tr>
<tr>
<td>23:30</td>
<td>W20</td>
<td>W15</td>
<td></td>
</tr>
<tr>
<td>01:30</td>
<td>W21</td>
<td>W16</td>
<td></td>
</tr>
<tr>
<td>03:30</td>
<td>W22</td>
<td>W17</td>
<td></td>
</tr>
<tr>
<td>05:30</td>
<td>W23</td>
<td>W18</td>
<td></td>
</tr>
<tr>
<td>07:30</td>
<td>W24</td>
<td>W19</td>
<td></td>
</tr>
<tr>
<td>00:00</td>
<td>Sleep 20:00 - 21:00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>09:30</td>
<td>W25</td>
<td>W20</td>
<td></td>
</tr>
<tr>
<td>11:30</td>
<td>W26</td>
<td>W21</td>
<td></td>
</tr>
<tr>
<td>13:30</td>
<td>W27</td>
<td>W22</td>
<td></td>
</tr>
<tr>
<td>15:30</td>
<td>W28</td>
<td>W23</td>
<td></td>
</tr>
<tr>
<td>17:30</td>
<td>W29</td>
<td>W24</td>
<td></td>
</tr>
<tr>
<td>19:30</td>
<td>W30</td>
<td>W25</td>
<td></td>
</tr>
<tr>
<td>21:30</td>
<td>W31</td>
<td>W26</td>
<td></td>
</tr>
<tr>
<td>23:00</td>
<td>Sleep 21:00 - 23:00</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 46: Study design for continuous wakefulness paradigm (Basner et al., 2008). T1-T7 are training bouts designed to accustomise subjects to the task.

'Night work' effects were judged as the differences between night work and day work (D1: Night W7-12 to Day W18-23; D2: Night W14-19 to Day W8-13), and 'sleep loss' effects as (D1: No Sleep W13-17 to 8hr Sleep W18-22; D2 No Sleep W20-24 to 8hrs Sleep W8-W12). This complicated design was used to
minimise circadian effects and to counterbalance experimental and control conditions. Both the night work and sleep loss conditions resulted in statistically significant changes in A' detection ability, hit rate and false alarm rates, and response latency decreased during the task. The authors conclude that both night work and sleep loss exhibit significant decrements in performance, although they attribute this to a fatigue effect and not explicitly to loss of sleep as that could only have been deduced if sleep loss alone and not night work had been significantly reduced. They also found that the high-difficulty items were consistently took longer to identify than those classified as easy.

5.1.5 The Ageing Workforce

The economy is ageing at an alarming rate due to various socioeconomic and health improvements (Dixon, 2003). As a consequence retirement is being pushed later in life (Disney, Meghir and Whitehouse, 1994) and older people who would normally have retired are finding that they need to find work to supplement their Government pension. Older workers may have problems with visual acuity compared to young, but this is not a problem that cannot be fixed by appropriate eyewear. However what is not so easily fixed are the incontestable cognitive differences between the two working age limits. The magnitude of age-related deficits in visual-cognitive performance is widely reported to increase with task complexity (Cerella and Hale, 1994). Evidence accumulated from several researchers suggests that differences between older and younger workers are down to differences in efficiency, speed and general-purpose processing resources (Dollinger and Hoyer, 1996). These deficits are reliable across the board regardless of the type of task used, so it is no surprise that the limited luggage search literature concludes that older workers are less capable than younger workers.

A 'Taskforce on the Aging of the American Workforce' (2008) brought together experts from nine different major employment departments to address the issues posed by that ageing of the 'Baby Boomer' generation. They found that older workers tended to re-enter the workforce by switching vocations, often taking a much reduced salary in return for flexible working hours and simpler setup. Security jobs like luggage scanning are therefore ideal target vocations for this demographic as they are relatively straightforward, do not require years of experience and are not physically taxing. In the laboratory Wiegmann et al. (2006) found that threat-item sensitivity was markedly lower in the older (m = 69.8 yrs) group than younger (m = 21.0 yrs) group for all conditions of cueing (given onscreen prompts). The older subjects were just as confident in their judgements as the younger even though their performance did not match the same levels. In the actual workplace, Riegelnig and Schwaninger (2006) found that age and detection performance correlated negatively (r(294) = -0.32) when age was extracted as a covariate. Ghylin, Drury and Schwaninger (2006) also found that detection level was lower in older workers albeit with a small effect size. They concluded that systematic training had overcome the deficit that older workers initially struggle from. Unfortunately there is little published data yet to make a solid case either way, but the studies conducted
so far suggest that older workers not only perform worse but may require a great deal of training to come up to the same standards as younger workers.

5.1.6 Summary

Luggage screening is the most important line of defence when combating terrorists who have historically used aircrafts to mass murder civilians and cause widespread fear and damage. Performance of these screeners has been described as "the weakest link" (Schwaninger, 2006) in the security process, so the vast sums of money spent since September 11th 2001 have to be justified. Screeners operate in the open with large volumes of peripheral distraction in the form of busy travellers passing through the checkpoint; a setup that has never been empirically justified, and an issue that is known to affect performance. This study provided an opportunity to concomitantly assess whether age would interact with daytime sleepiness, or if the benefits of training could bring older subjects up to the same standard as younger subjects.

Research Questions

1. How quickly can a naive subject reach peak performance on a simulated luggage search task? Does the age of the subject matter?

2. Is luggage-search ability impacted by a night of unsatisfactory sleep? How might this apply to other vocations?

3. Is there a performance difference between peripheral and cognitive distractions?

4. What effect does age have when considering the effects of sleepiness and distraction on an engaging task?

Methods

5.2.1 Screening And Subjects

Twelve healthy young participants (m = 20.8yrs, s.d. = 1.4yrs; females = 6) and twelve healthy and active older participants (m = 60.0yrs, s.d. = 3.7yrs; females = 7) at both ends of the working spectrum and similar in personal traits were screened and selected to take part in a three-day study. The younger group were canvassed by departmental mailings and electronic communications, while the older group were largely recruited from the University of the Third Age, which is an organisation designed for active and inquisitive individuals to further their education and socialise when retired. A couple of older participants volunteered from the Loughborough Lions Guild, which is a community outreach programme that is typically staffed by older persons. They were all asked to report to the Sleep Research Centre for a round of screening tests, questionnaires and a brief face-to-face interview (see Section 2.3.3).
Both groups had to fit within limits of normality on the questionnaire, and some potential applicants were considered unsuitable for the study as they either slept too much (>9 hrs), admitted to recreational drug use or lived outside of the taxi catchment area. In particular the older subjects were informed during the recruitment process that we were looking for healthy, normal sleepers, experienced in computer-based tasks and free of physical debilitation (with emphasis on good hearing ability). Once applicants had passed these requirements, they were then given an actiwatch to wear for three nights and a sleep diary to monitor how active they were and to gauge their routine sleeping habits so that successful applicants were similar to each other in sleep need and length. On return of these, a date was set for the first of three experimental days.

5.2.2 Protocol

Participants were required to present themselves at the Sleep Research Centre on three days separated at least by one day so that any fatigue/motivational defects accumulated could be readily recovered before the next test date. The first day was a practice day that had to be factored into the study design when piloting showed that it took subjects between 3-4 full X-Ray ORT tests before they had reached a plateau in performance. Owing to this date being a practice, no sleep or EEG measures were taken and participants were fully debriefed as to the test design and requirements. No distractions were presented so that uninhibited learning could be completed, and participants filled out subjective difficulty scales after the completion of each test. At all times throughout the practice and experimental days an experimenter was present in the room to ensure that participants were completing the test satisfactorily.

On two proceeding occasions the participants returned for the full experimental day and were given actiwatches to check that they had slept the requisite amounts. In addition to this, participants were required to leave an answering machine message when retiring to bed and on awakening to double-check compliance. For the sleep restriction days the participants had to limit themselves strictly to less than five hours sleep by advancing their wakefulness and waking up at their normal time, while on the normal sleep days they were allowed to sleep their normal sleep quota and encouraged to seek the most refreshing sleep they could. Participants arrived at the Sleep Research Centre at midday, and on the first occasion the ethical forms were completed, at which point the electrodes were attached with bipolar derivations on C4-F4, C4-P4, C3-F3 and C3-P3 in accordance with the International 10-20 system (Appendix C). Participants were given a light lunch of sandwiches and fruit, then testing began once their actiwatches were verified.

During testing three levels of distraction were counterbalanced for each of the age groups, as well as counterbalancing the order of sleep restriction and normal sleep. The first level of distraction was simply no distraction, while the second level was television only (showing episodes of the ITV series ‘Airline’), and the third level combined the television with a verbal fluency task. This
verbal fluency task presented subjects with a noun every thirty seconds, with the expectation that three verbs relating to that noun would be orally produced. By doing this a reliable and standardised cognitive distraction was created which ensured that peripheral input had been cognitively processed by the output verbs, effectively mimicking a conversation or the act of being 'lost in thought'. Although the subjects were all video-recorded during the experimental days the outcome from Chapter Four had provided a precursor to what we would find here, in that subjects were able to attend to the task for the duration without a behavioural indication of distraction. Compounding this was the fact that the X-Ray ORT allows three in-built subject-paced breaks during this task, meaning that subjects quickly learned to attend to the television during these breaks causing no behavioural correlate of distraction to performance ability. Thus the cognitive distraction element in the study design helped to add an extra level of distraction that was lacking in the first study, but at the expense of a behavioural distraction measure. Participants were paid an honorarium of £50 for completion of all three test dates and were provided taxis to and from the Sleep Research Centre on the sleep restriction day.

Table 19: Counterbalancing table for the young and old groups. This protocol ensured that distraction and sleepiness were not favoured in the experimental design for treatment order effects.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Young</td>
<td>1-2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3-4</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>6</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>5-6</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>5</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>7-8</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>9-10</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>11-12</td>
<td>5</td>
<td>6</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>13-14</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>15-16</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>6</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>17-18</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>5</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>19-20</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>21-22</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>23-24</td>
<td>5</td>
<td>6</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>1</td>
</tr>
</tbody>
</table>

5.2.3 Laboratory Settings

The laboratory build was left relatively unchanged from previous studies and pilot testing save for changes in the task, daytime setting, and distraction equipment. The X-Ray ORT (Figure 6, Section 2.1.3), detailed below, runs on Windows XP and only requires a mouse to operate, so additional mice were connected to the experimenter side so that the X-Ray ORT could be controlled between-trials by the experimenters. New headsets were provided for the subjects for the verbal fluency task, which again were initiated and terminated on the experimenter side. The series Airline was recorded to DVD and set on a running chapter loop so that subjects never experienced the same episode twice.
The X-Ray ORT differs in many crucial aspects from similar software packages but retains the same principles of two-choice ratings, confidence judgments and self-paced nature. All images are presented in monotone to limit experience effects from the false-coloured images used in airports, and only two threat items (guns and knives) are used with varying difficulties. Before each test eight bag images are shown for practice are given with feedback, and then a 256-item battery is undertaken without response feedback. All threat items used are shown for ten seconds for guns and then ten seconds for knives in order to eliminate knowledge effects. To this end, the X-Ray ORT is practically identical to the Simulated Luggage Search Task used by Basner and colleagues (2008), in that they utilised a 200-item test with prior feedback practice shown in black-and-white. Given that airport screeners typically have between 3-6 seconds to inspect an item (Singh and Singh, 2004; Schwaninger, Michel and Bolfing, 2007) the X-Ray ORT limits image exposure to four seconds, similar to Basner et al.’s (2008) test which was time-fixed to seven seconds.

5.2.4 EEG

The EEG settings were changed from monopolar (in Chapters Three and Four) to bipolar derivations. Four-times as many research articles of EEG studies utilise bipolar derivations (Fehmi and Sundor, 1989) than monopolar due to the ease in set-up and calibration, but the debate continues as to which is preferential. The montage used was F3-C3 and F4-C4 for frontal-lobe insight, along with P3-C3 and P4-C4 for parietal lobe measurement. Three-minute baseline activity was recorded prior to each condition to give baseline-corrected measurements. It was found in data cleaning that frequencies below 0.3Hz had to be removed for artefact, and data was analysed by power-spectrum analysis of 1Hz bins ranging from 0.1Hz to 13Hz. “Power” here does not represent wattage as is typically used in electro physics, but \( \mu V^2/Hz \). Volts are squared because the pen-deflection polarity does not represent a meaningful direction. Data was carefully cleaned by manual inspection of each 100s from each recording, resulting in 144 (24 subjects x 2 waking sleepiness conditions x 3 distraction conditions) x 4 (EEG recording sites) = 576 recordings. Of these 576 recordings, 12 were omitted due to excessive movement artefact which resulted in unintelligible results, representing only 0.02% of recordings. Other telltale signs of artefact include repetitive, irregular or rhythmical waveforms that appear simultaneously in unrelated head regions (Fisch and Spehlman, 1999).

5.2.5 Subjective Scales

In addition to the EPQ-R and State-Trait Anxiety Index, two new subjective scales were appended to the test design to tap into two important aspects of signal detection theory and visual search ability. Firstly, arousability is well known to cause a behavioural impact in visual search studies (Coren and Mah, 1993), so an oft-used scale was found that has been found to reliably estimate subjective arousal correlated to behavioural indices. Coren (1988, 1990) developed an Arousability Prediction Scale (APS) originally to identify
insomnia from cognitive hyperarousal, but later applied this scale to the notion of distractibility in visual search, which is also related to arousal (Aks and Coren, 1990). Thus this scale provides an ideal method to determine which subjects may be differentially impaired by cognitive distraction as determined by heightened arousal.

An important concept in signal detection theory, which is often used to generate Receiver-Operating Curves, is the confidence in which a subject places on their assertion as to whether a bag contains a threat item or not. Unfortunately many subjects race through this part of the task and quickly stop making careful assessment of their confidence per image. So in addition to the per-item confidence scores (which are not analysed due to the proceeding observation that these confidence scores are unreliable), subjects were asked to rate their confidence on a 100-point visual analogue scale asking firstly “How did you find the task? Very Easy – Very Hard” and secondly “How well do you think you performed? Not Very Well – Very Well.” This helped to gauge intrinsic ability to know when subjects had performed at or above their own best levels.

5.2.6 Modelling Results

A scientific theory can wield more predictive power if it can be modelled, tested, and withstand repeated scrutiny from similar and competing designs. When dealing with signal detection theory, itself a modelling framework, significant detractions arise that cannot be resolved using current mathematical methods. Speed and accuracy are two readily-collected components of two-choice tasks, yet they tend to be analysed separately due to difficulties in conjoining both (Ratcliff and Rouder, 1998). Many paradigms throughout the visual and inspection field yield high accuracy values, meaning that experimental manipulations can often be witnessed only in response time changes (Ratcliff, 2001), particularly when tasks are subject-paced. Derivative models such as the Two Component Model (TCM) and Sequential Sampling Models (SSM) are grounded in Signal Detection Theory (SDT; Green and Swets, 1966), which is limited by its inability to account for speed (or its reciprocal time) by assuming a fixed sampling interval (Smith, 2000). Random walk and diffusion theory (Link and Heath, 1975; Ratcliff, 1978) were created to account for this major limitation, although confidence ratings are notable by their absence in both of these theories (Plescak and Busemeyer, 2007) as well as Drury’s TCM. The need to link SDT with a unified theory of operator performance in visual or inspection tasks is highlighted by Drury (1994), who states that “speed measures are so basic to performance evaluation that the need for them is rarely discussed” (p. 748).

The class of sequential sampling models attempt to reconcile the deficiencies of SDT by linking detection performance and underlying reaction time distributions in two-choice tasks (Smith, 2000) while accurately estimating the speed-accuracy trade-off effects based on accumulation of evidence (Ratcliff and Smith, 2004). SSMs account for the different distributions and scales that speed and detection accuracy have (Ratcliff, 2001). From the literature of simple two-choice tasks diffusion models are by far the most prominent
methods of evaluating both detection and reaction split data (Ratcliff and Tuerlinckx, 2002). While SSMs are neurological in nature (Smith and Ratcliff, 2004) and theoretically sound for short duration tasks (~1000-1500ms mean RT; Ratcliff and McKoon, 2008), their application can be found in impaired (e.g. aphasics and dyslexics; Ratcliff et al., 2004) and elderly populations (aged 75-90 years; Ratcliff, Thapar and McKoon, 2007) that tend to exhibit slower reaction times than the recommended maximum, as well as simple visual search tasks (Schwarz, 1993). However, when responses above one second are recorded then conclusions are likely to be based on multiple failed decision attempts (Ratcliff and Rouder, 1998) resulting in a series of unpredictable diffusion processes operating at various times as various evidences are consumed and rejected. For these reasons the X-Ray ORT does not lend itself well to current SSMs and so only the TCM will be applied to the presenting data, although future resolution of these issues will tend towards an extendable model appropriate for X-Ray screening.

Drury's TCM (Drury, 1975) has been widely used throughout ergonomics to estimate model parameters for inspection tasks, combining visual search and decision theory. The model predicts that as the probability of a hit increases with time allowed to inspect an item, the probability of a false alarm increases concomitantly (Spitz and Drury, 1978). Using formulae derived from these early studies, Ghylin, Drury and Schwaninger (2006) have successfully applied the inspection model to the X-Ray Tutor training program (see Koller et al., 2007). A model of estimated speed/accuracy trade-off can be obtained by applying the formulae found in Ghylin et al. (2008), themselves being an extension of formulae derived by Spitz and Drury (1978). Data could not be modelled for subjects with less than five false alarms as these situations did not create enough points for linear correlation.

For each participant, reaction time was sorted in ascending order and assigned an incremental ordinal number per threat item hit starting from 1, and then the probability of a hit at each threat item response (P(Hit(x))) was generated by dividing the ordinal by total threat item count which, essentially, adopts Drury's (1994) p(error) formulation. These probabilities were mapped to the corresponding reaction times and used as the ordinate in the below equation:

1) \[ y = \log\left(\frac{1-P(Hit(x))}{Pdh}\right) \]

Where P(Hit(x)) at any given reaction time (x) is the cumulative probability of a hit for each reaction time, and Pdh is the overall probability of a hit derived by dividing the sum of total hits by total number of threat items. This gives a y-value for each reaction time (x) that can be graphed for each participant's session, such that a best fitting linear equation (judged by \( r^2 \)) can be fit to the various y and x (reaction time) points. The slope (gradient) and x-intercept of the linear equation represent the search time (STh) and non-search time (NSTh) respectively. To graph the overall SAOC using these values, a second formula can be used:

2) \[ P(Hit) = [1-\exp(-(x-NSTh)/STh)]*Pdh \]
Where \( P(\text{Hit}) \) is the y-value that is generated by graphing the function that uses the values of the NSTh, STh and Pdh, and this formula is similar to other speed-accuracy curve equations (e.g. Wickelgren, 1977). This method is then repeated for false alarms by following the exact same method, although to calculate the ordinal only the non-threat items should be considered rather than threat items. The parameters from each participant are then averaged to form the average Speed-Accuracy Operating Curve (SAOC) shown in Figure 48 and data given in Table 20. The stopping-time policy is obtained by linear correlation of average hit time per person (detection) compared with average correct rejection time (stopping policy) to give an account of speed for purely accurate judgments.

Drury's TCM is only applied to practice data here to conceptually model learning effects, with custom signal detection theory programs designed by our research group used for the experimental days. Main study results were also analysed using a new conceptual measure, Balakrishnan's dynamic signal detection theory (Balakrishnan et al., 2001; Wang, 2007) is applied to the dataset, which claims to be able to conjoin detection and reaction time into a single measure. The sensitivity measure \( A' \) was chosen over \( d' \) because it makes no assumptions about signal and noise distribution normality. Overall \( A' \) detection values, \( B'' \) bias and mean reaction times were calculated for young and old participants using the formulas below where \( H \) is hit rate (number of hits divided by number of threat items) and \( F \) is the false alarm rate (number of false alarms divided by number of non-threat items):

If \( H \geq F \):
\[
A' = 0.5 \left[ 1 + \frac{(H - F)(1 + H - F)}{4H(1 - F)} \right]
\]
\[
B'' = \frac{H(1 - H) - F(1 - F)}{(H(1 - H)+F(1 - F)}
\]

If \( H < F \):
\[
A' = 0.5 - \left[ 1 - \frac{(F - H)(1 + F - H)}{4F(1 - H)} \right]
\]
\[
B'' = \frac{-H(1 - H) + F(1 - F)}{(H(1 - H)+F(1 - F)}
\]

Results

5.3.1 Descriptives

Twenty-four subjects split into two age groups of twelve were studied, with the young group (mean = 20.8yrs, s.d. = 1.4yrs) comprising 6 females and 6 males, while the older group (mean = 60.0yrs, s.d. = 3.7yrs) had 7 females and 5 males. The young group slept an average of 509.2mins (s.d. = 30.67) during normal sleep compared to 272.5mins (s.d. = 26.24) in restricted sleep, while the older group slept 485.0mins (s.d. = 38.2) compared to 279.2mins (s.d. = 12.9). This represents a sleep loss of 46.48% for the young group and 42.43% for the older group compared to normal sleeping habits, resulting in a comparable loss of usual sleep.
5.3.2 Learning Model

The practice day, where subjects practiced the test three times consecutively, offered an ideal opportunity to model learning ability in young and old persons. Drury’s TCM parameters for the first and third of these sessions were calculated for each participant and then aggregated to find an overall model fit. Data had to be eliminated in 0.38% of trials as outliers (excessively slow reaction times above ten seconds), and three subjects had less than 5 false alarms where data couldn’t be reliably modelled. $R^2$ values were higher for hits than false alarms (Table 20), although only Non-Search Time improved significantly with practice, and age was found not to be a significant factor for any individual SAOC component. All interactions were insignificant, so only main effects are shown in Table 21.

Table 20: Model fit parameters for Drury’s model. Mean value and standard deviation are given for young and old, untrained (u) and trained (t).

<table>
<thead>
<tr>
<th></th>
<th>Hits</th>
<th>False Alarms</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pd NST ST</td>
<td>r²</td>
<td>Pd NST ST r²</td>
</tr>
<tr>
<td>Young (u)</td>
<td>0.703 1.087 1.002 0.913 0.151 2.552 0.837 0.747</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.089)</td>
<td>(0.259)</td>
<td>(0.340)</td>
</tr>
<tr>
<td>Young (t)</td>
<td>0.785 0.849 1.049 0.965 0.139 1.800 0.615 0.763</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.084)</td>
<td>(0.225)</td>
<td>(0.187)</td>
</tr>
<tr>
<td>Old (u)</td>
<td>0.683 1.872 0.932 0.905 0.223 3.239 0.903 0.795</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.088)</td>
<td>(0.554)</td>
<td>(0.478)</td>
</tr>
<tr>
<td>Old (t)</td>
<td>0.751 1.431 1.003 0.935 0.271 2.803 0.884 0.705</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.113)</td>
<td>(0.276)</td>
<td>(0.333)</td>
</tr>
</tbody>
</table>

Table 21: Inferential tests on Non-Search Time (NST) and Search Time (ST) for hits (h) and false alarms (fa). Only NSTh and NSTfa improved significantly with practice, showing no age effects.

<table>
<thead>
<tr>
<th></th>
<th>Hits</th>
<th>False Alarms</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NST ST</td>
<td></td>
<td>NST ST</td>
</tr>
<tr>
<td>Age</td>
<td>$F_{(1,22)}$ = 1.774, $p &gt; 0.05$, $\eta^2 = 0.07$</td>
<td>$F_{(1,22)}$ = 1.463, $p &gt; 0.05$, $\eta^2 = 0.02$</td>
<td>$F_{(1,19)}$ = 1.625, $p &gt; 0.05$, $\eta^2 = 0.06$</td>
</tr>
<tr>
<td>Practice</td>
<td>$F_{(1,22)}$ = 16.887, $p &lt; 0.05$, $\eta^2 = 0.05$</td>
<td>$F_{(1,22)}$ = 0.861, $p &lt; 0.05$, $\eta^2 = 0.03$</td>
<td>$F_{(1,19)}$ = 5.631, $p &lt; 0.05$, $\eta^2 = 0.07$</td>
</tr>
</tbody>
</table>

Closer inspection of the signal-detection measures corroborate with the findings from the SAOC, with appreciable improvements in both $A'$ and RT (Figure 47). Overall detection ability ($A'$) improved with practice for both young and old participants ($F_{(1,22)} = 27.172, p < 0.05, \eta^2 = 0.15$), as well as showing a significant between-subject factor of age ($F_{(1,22)} = 8.243, p < 0.05, \eta^2 = 0.27$). The shift in detection and reaction time can be seen in Figure 47, as trained subjects complete the task quicker and more accurately than untrained subjects. It took three sessions for the older group to become...
comparably accurate and fast as the younger during their first ($A' 0.858 (0.026)$ to $A' 0.852 (0.051)$; $rt 4.06s (0.469)$ to $4.09s (0.481)$).

![Figure 47: Improvements in A' and RT for young and old groups. There is no loss in performance when subjects respond quicker to the task. In fact, substantial improvements in detection ability come with training.](image)

The model parameters derived in Table 20 can be put back into the original formula to graphically illustrate the SAOC. The output of this graph is shown in Figure 48, where the young were both quicker and more accurate than the older group. Both groups demonstrated an abscissa-shift towards faster reaction times, while increasing their probability of a hit as shown by higher $P(\text{Hit})$ curves. The probability of a correct rejection (as a correct response for non-threat items) remained fairly stable for young subjects, but showed a slight improvement in older subjects in addition to the clear speed improvements.

$$P(\text{Hit}) = [1-\exp(-t/N\text{STh})/S\text{Th})]^{P\text{dh}}$$
Figure 48: Speed-Accuracy Operating Curve for young and old groups. These curves give a representation of non-search time as the time given per search and associated accuracy, and search time where the curves intercept the x-axis.

The stopping time policies of young and old subjects can be investigated by linear correlation of hit detection time and correct rejection detection time. Effects of age and training serve as determinants of slower stopping policies but still showing an orderly relationship that accounts for much of the variance ($F_{(1,22)} = 63.981$, $r^2 = 0.573$). Outliers tend to be from the untrained condition, where subjects were becoming accustomed to the novelty of luggage search, although only three true outliers exist with small influence.
5.3.3 Experimental Day Performance

Performance measured by A' and RT for young subjects is summarised in Figure 50, while older subject performance is given in Figure 51. A three-way mixed ANOVA (age x sleep x distraction) was used to calculate main effects for each of the three dependent variables: A', RT and bias (B'). There was no significant main effect for either sleepiness condition nor distraction for detection accuracy ($F_{1.22} = 3.387, p > 0.05, \eta^2 = 0.02$ and $F_{2.44} = 3.062, p > 0.05, \eta^2 = 0.07$ respectively), although the latter only just missed the 95% cutoff by 0.007% so the medium effect size should be noted. However there was a significant main effect of age for detection performance ($F_{1.22} = 9.631, p < 0.05, \eta^2 = 0.19$).

For RT, sleepiness condition was highly insignificant ($F_{1.22} = 0.669, p > 0.05, \eta^2 = 0.00$) while distraction ($F_{2.44} = 17.292, p < 0.05, \eta^2 = 0.10$) and age ($F_{1.22} = 31.749, p < 0.05, \eta^2 = 0.51$) were highly significant with a large effect size of age. Figure 50 and Figure 51 support a significant linear increase for reaction time distraction ($F_{1.22} = 25.598, p < 0.05, \eta^2 = 0.09$), but not for sleep. Sleepiness condition and age do not interact for either detection ($F_{1.22} = 0.543, p > 0.05, \eta^2 = 0.01$) or RT ($F_{1.22} = 0.059, p > 0.05, \eta^2 = 0.00$), which may have been expected if either group were differentially at risk to sleep.
loss. Bias results were highly insignificant for sleep, distraction and age. A 2x2x3 ANOVA (gender x sleep x distraction) reported insignificant between-subjects factor of gender for A' (F(1,22) = 2.175, p > 0.05, \( \eta^2 = 0.08 \)) and RT (F(1,22) = 0.996, p > 0.05, \( \eta^2 = 0.04 \)). Interactions for sleepiness condition x distraction were significant for RT (F(1,22) = 0.475, p > 0.05, \( \eta^2 = 0.00 \)), but no other interactions approached significance for either A' or RT and given the almost nonexistent \( \eta^2 \) these can be safely rejected.

Using dynamic signal detection measures to create a unity between detection performance (d') and reaction time measures gives Balakrishnan's d results (Figure 52). A two-way ANOVA (sleepiness condition x distraction) for the young group (again not including the practice day) confirmed that there was a large significant effect of distraction (F(1,11) = 7.18, p < 0.05, \( \eta^2 = 0.14 \)), but not sleepiness condition (F(1,11) = 7.18, p > 0.05, \( \eta^2 = 0.02 \)). The same analysis for the old group found that distraction narrowly missed significance (F(1,11) = 2.959, p > 0.05, \( \eta^2 = 0.04 \)) but also that sleepiness condition was resoundingly insignificant (F(1,11) = 0.456, p > 0.05, \( \eta^2 = 0.00 \)). A three-way ANOVA with age as between-subjects factor revealed a large significant effect of age (F(1,22) = 31.278, p < 0.05, \( \eta^2 = 0.51 \)) that is plainly demonstrated in Figure 52 when considering reaction time and detection together.

Figure 50: Detection performance and speed differences in the young group. The darker bars represent A' and the lighter bars RT. ND = No Distraction, D = Distraction (television only), DC = Distraction (Cognitive).
Old A' and RT

Figure 51: Detection performance and speed differences in the old group. The darker bars represent A' and the lighter bars RT. ND = No Distraction, D = Distraction (television only), DC = Distraction (Cognitive).

Figure 52: Balakrishnan's d values for each condition and age group. Dramatic differences between the young and old groups are seen when reaction time and accuracy are joined, as well as for distraction.
Stopping-time policies for the young and old group in distraction and no distraction conditions are shown on separate graphs (Figure 53) for clarity and ease of comparison. Sleepiness conditions are combined (two points per individual per distraction level) as all measures suggest that sleepiness condition was not a significant factor for speed, so that the no distraction condition can be compared to the cognitive distraction condition with higher reliability. A relationship exists between the reaction times for threat and non-threat items as the stopping policy, which is comparable between young and old persons. With increasing distraction the reaction times do worsen but not typically longer than a second, although hit reaction times showed less of a shift than correct rejections. Stopping policy relationships ranged from $r^2 = 0.22$ for old distraction, to $r^2 = 0.52 - 0.55$ for the other three correlations.

**Figure 53:** Stopping policies for no distraction and distraction by age. A marginal shift is evidenced between no distraction reactions and cognitive distraction, but the relationship between threat items and non-threat items remains similar as shown by the likeness in slopes.
5.3.4 Confidence Ratings

Confidence ratings for the two confidence questions, "how did you find the task?" as Confidence 1, and "how well do you think you performed?" as Confidence 2, were correlated to detection ability. A tabulation of these (Table 22) revealed mid-level correlations for some estimates such as confidence for sleep restriction when not distracted ($r^2 = 0.14-0.16$) and normal sleep with cognitive distraction ($r^2 = 0.10-0.12$), although they were not statistically significant ($p > 0.05$). The only statistically significant result from the twelve was for sleep restricted and cognitively distracted for the second confidence rating, which was highly correlated ($r^2 = 0.21, p < 0.05$). A partial correlation controlling for age found similar findings where only the sleep restricted and cognitive distraction confidence ratings correlated significantly ($r^2 = 0.21$). Effect sizes are given as the correlation measure 'r' in Table 22.

Table 22: Correlations between detection performance ($A'$) and confidence. Certain correlations are reasonably high in behavioural science terms (e.g. $r = 0.35-0.39$), but only cognitive distraction under sleep restriction achieves a significant result from the twelve measures.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Statistic</th>
<th>Confidence1</th>
<th>Confidence2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal Sleep, No Distraction</td>
<td>Correlation</td>
<td>0.2016</td>
<td>0.1885</td>
</tr>
<tr>
<td></td>
<td>Significance</td>
<td>$p &gt; 0.05$</td>
<td>$p &gt; 0.05$</td>
</tr>
<tr>
<td>Normal Sleep, Distraction</td>
<td>Correlation</td>
<td>0.1223</td>
<td>0.1836</td>
</tr>
<tr>
<td></td>
<td>Significance</td>
<td>$p &gt; 0.05$</td>
<td>$p &gt; 0.05$</td>
</tr>
<tr>
<td>Normal Sleep, Cog. Distraction</td>
<td>Correlation</td>
<td>0.3510</td>
<td>0.3204</td>
</tr>
<tr>
<td></td>
<td>Significance</td>
<td>$p &gt; 0.05$</td>
<td>$p &gt; 0.05$</td>
</tr>
<tr>
<td>Sleep Restricted, No Distraction</td>
<td>Correlation</td>
<td>0.3704</td>
<td>0.3974</td>
</tr>
<tr>
<td></td>
<td>Significance</td>
<td>$p &gt; 0.05$</td>
<td>$p &gt; 0.05$</td>
</tr>
<tr>
<td>Sleep Restricted, Distraction</td>
<td>Correlation</td>
<td>0.0251</td>
<td>0.1994</td>
</tr>
<tr>
<td></td>
<td>Significance</td>
<td>$p &gt; 0.05$</td>
<td>$p &gt; 0.05$</td>
</tr>
<tr>
<td>Sleep Restricted, Cog. Distraction</td>
<td>Correlation</td>
<td>0.2731</td>
<td>0.4532</td>
</tr>
<tr>
<td></td>
<td>Significance</td>
<td>$p &gt; 0.05$</td>
<td>$p &lt; 0.05$</td>
</tr>
</tbody>
</table>

Confidence values were correlated for each subject using Pearson's product-moment correlations to $A'$ detection performance in each condition and then averaged by group. Averaging was achieved by converting the confidence values into z-scores per participant and then correlating these averages ($A'$ to confidence rating). Young $r^2 = 0.13$ (s.d. 0.18) for the first question, "How did you find the task?" while the old $r^2 = 0.15$ (s.d. 0.13). In answer to the second question, young $r^2 = 0.22$ (s.d. 0.16) and old $r^2 = 0.15$ (s.d. 0.20). Comparing these averages to the individual condition correlations in Table 23 shows that subjects have little insight into their performance capability independent of whether they are young or old. These correlations represent performance after three practice trials (which are not included) and so subjects should be reasonably well attuned to the difficulty of the task and their performance relative to it.
5.3.5 Personality

Personality differences between the age groups showed no significant differences for all measures except extraversion (Table 23), which was highly significant. The effect size for Cohen’s $d$ should be interpreted as $d = 0.2$ (small), $d = 0.5$ (medium), and $d > 0.8$ (large).

Table 23: Personality differences between the young and old group. Only extraversion differed between the groups significantly.

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Young</th>
<th>Old</th>
<th>t(22)</th>
<th>p-value</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trait Anxiety</td>
<td>33.58 (5.45)</td>
<td>37.00 (6.42)</td>
<td>-1.405</td>
<td>&gt;0.05</td>
<td>0.53</td>
</tr>
<tr>
<td>APS</td>
<td>32.58 (5.87)</td>
<td>31.82 (3.43)</td>
<td>0.510</td>
<td>&gt;0.05</td>
<td>0.29</td>
</tr>
<tr>
<td>Extraversion</td>
<td>18.75 (4.45)</td>
<td>12.67 (5.05)</td>
<td>3.129</td>
<td>&lt;0.05</td>
<td>1.20</td>
</tr>
<tr>
<td>Psychoticism</td>
<td>6.17 (2.48)</td>
<td>4.67 (2.64)</td>
<td>1.434</td>
<td>&gt;0.05</td>
<td>0.57</td>
</tr>
<tr>
<td>Neuroticism</td>
<td>7.25 (2.98)</td>
<td>10.25 (5.33)</td>
<td>-1.563</td>
<td>&gt;0.05</td>
<td>0.56</td>
</tr>
</tbody>
</table>

Groups were split by ages and personality type by a median-split and then analysed using independent-samples t-tests. If a group was significantly different from the other half then meaningful analyses could be undertaken, otherwise they would be rejected for not having meaningful utility. Table 24 showed that all split-half groups were significantly different from their counterparts and so could be used to test for personality effects against X-Ray ORT performance.

Table 24: Split-half personality dimensions.

<table>
<thead>
<tr>
<th>Age</th>
<th>Dimension</th>
<th>t</th>
<th>Mean Diff</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Young</td>
<td>Extraversion</td>
<td>2.25</td>
<td>5.25</td>
<td>$p &lt; 0.05$</td>
</tr>
<tr>
<td></td>
<td>Psychoticism</td>
<td>3.84</td>
<td>3.67</td>
<td>$p &lt; 0.05$</td>
</tr>
<tr>
<td></td>
<td>Neuroticism</td>
<td>4.35</td>
<td>6.17</td>
<td>$p &lt; 0.05$</td>
</tr>
<tr>
<td></td>
<td>Anxiety</td>
<td>3.59</td>
<td>7.83</td>
<td>$p &lt; 0.05$</td>
</tr>
<tr>
<td></td>
<td>Arousal</td>
<td>5.00</td>
<td>9.50</td>
<td>$p &lt; 0.05$</td>
</tr>
<tr>
<td></td>
<td>Extraversion</td>
<td>4.11</td>
<td>7.67</td>
<td>$p &lt; 0.05$</td>
</tr>
<tr>
<td></td>
<td>Psychoticism</td>
<td>5.26</td>
<td>4.33</td>
<td>$p &lt; 0.05$</td>
</tr>
<tr>
<td>Old</td>
<td>Neuroticism</td>
<td>3.12</td>
<td>7.17</td>
<td>$p &lt; 0.05$</td>
</tr>
<tr>
<td></td>
<td>Anxiety</td>
<td>5.50</td>
<td>10.67</td>
<td>$p &lt; 0.05$</td>
</tr>
<tr>
<td></td>
<td>Arousal</td>
<td>3.81</td>
<td>5.11</td>
<td>$p &lt; 0.05$</td>
</tr>
</tbody>
</table>

Although personality did not differ between the age groups for most personality traits, the fact that detection performance was markedly worse in the older sample means that predictions of ability based on personality dimensions still need to be split by age when grouping variables. The two most dissimilar conditions, normal sleep with no distraction and normal sleep with cognitive distraction, were chosen to see whether normal ability or
Table 25: Predictors of performance when nonsleepy and nondistracted for detection measures (A'). Several scattered predictors achieve significance such as arousal, anxiety and neuroticism while psychoticism and extraversion do not.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Personality</th>
<th>Means</th>
<th>St.Dev.</th>
<th>t</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Anxiety</td>
<td>0.932</td>
<td>0.00</td>
<td>4.145</td>
<td>p &lt; 0.05</td>
</tr>
<tr>
<td></td>
<td>0.907</td>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Detection</td>
<td>Extraversion</td>
<td>0.921</td>
<td>0.01</td>
<td>0.194</td>
<td>p &gt; 0.05</td>
</tr>
<tr>
<td>Young</td>
<td>0.919</td>
<td>0.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Psychoticism</td>
<td>0.926</td>
<td>0.01</td>
<td>1.317</td>
<td>p &gt; 0.05</td>
</tr>
<tr>
<td></td>
<td>0.914</td>
<td>0.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Neuroticism</td>
<td>0.929</td>
<td>0.01</td>
<td>2.290</td>
<td>p &lt; 0.05</td>
</tr>
<tr>
<td></td>
<td>0.910</td>
<td>0.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Arousal</td>
<td>0.927</td>
<td>0.01</td>
<td>1.600</td>
<td>p &gt; 0.05</td>
</tr>
<tr>
<td></td>
<td>0.913</td>
<td>0.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Anxiety</td>
<td>2.88</td>
<td>1.01</td>
<td>1.63</td>
<td>p &gt; 0.05</td>
</tr>
<tr>
<td></td>
<td>2.08</td>
<td>0.62</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Extraversion</td>
<td>2.60</td>
<td>0.90</td>
<td>0.627</td>
<td>p &gt; 0.05</td>
</tr>
<tr>
<td></td>
<td>2.24</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Psychoticism</td>
<td>2.72</td>
<td>1.03</td>
<td>0.917</td>
<td>p &gt; 0.05</td>
</tr>
<tr>
<td>Speed</td>
<td>2.23</td>
<td>0.77</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Young</td>
<td>Neuroticism</td>
<td>2.52</td>
<td>1.05</td>
<td>0.160</td>
<td>p &gt; 0.05</td>
</tr>
<tr>
<td></td>
<td>2.43</td>
<td>0.83</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Arousal</td>
<td>2.66</td>
<td>1.06</td>
<td>0.689</td>
<td>p &gt; 0.05</td>
</tr>
<tr>
<td></td>
<td>2.30</td>
<td>0.76</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Anxiety</td>
<td>0.878</td>
<td>0.04</td>
<td>0.416</td>
<td>p &lt; 0.05</td>
</tr>
<tr>
<td></td>
<td>0.866</td>
<td>0.06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Extraversion</td>
<td>0.879</td>
<td>0.04</td>
<td>0.435</td>
<td>p &gt; 0.05</td>
</tr>
<tr>
<td></td>
<td>0.866</td>
<td>0.06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Psychoticism</td>
<td>0.879</td>
<td>0.04</td>
<td>0.461</td>
<td>p &gt; 0.05</td>
</tr>
<tr>
<td>Detection</td>
<td>0.866</td>
<td>0.06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Old</td>
<td>Neuroticism</td>
<td>0.878</td>
<td>0.04</td>
<td>0.360</td>
<td>p &gt; 0.05</td>
</tr>
<tr>
<td></td>
<td>0.867</td>
<td>0.06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Arousal</td>
<td>0.915</td>
<td>0.02</td>
<td>3.770</td>
<td>p &lt; 0.05</td>
</tr>
<tr>
<td></td>
<td>0.842</td>
<td>0.04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Anxiety</td>
<td>4.36</td>
<td>0.58</td>
<td>3.085</td>
<td>p &lt; 0.05</td>
</tr>
<tr>
<td></td>
<td>3.33</td>
<td>0.57</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Extraversion</td>
<td>3.94</td>
<td>0.76</td>
<td>0.416</td>
<td>p &gt; 0.05</td>
</tr>
<tr>
<td></td>
<td>3.75</td>
<td>0.83</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Psychoticism</td>
<td>3.86</td>
<td>0.70</td>
<td>0.068</td>
<td>p &gt; 0.05</td>
</tr>
<tr>
<td>Speed</td>
<td>3.83</td>
<td>0.90</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Old</td>
<td>Neuroticism</td>
<td>3.89</td>
<td>0.72</td>
<td>0.196</td>
<td>p &gt; 0.05</td>
</tr>
<tr>
<td></td>
<td>3.80</td>
<td>0.88</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Arousal</td>
<td>4.36</td>
<td>0.58</td>
<td>3.085</td>
<td>p &lt; 0.05</td>
</tr>
<tr>
<td></td>
<td>3.33</td>
<td>0.57</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
distracted ability is affected by personality. Given that sleep restriction did not measurably alter results it was not necessary to include sleep restricted measures. Table 25 shows that arousal is a predictor of performance in older subjects for both detection and reaction time, while anxiety can only predict the speed at which they respond. For younger subjects, anxiety and neuroticism (to a lesser extent) predicted detection ability whilst none of the predictors could explain any differences in speed of reactions.

5.3.6 EEG Analysis

EEG measures for frontal (F3-C3 and F4-C4) and parietal lobes (P3-C3 and P4-C4) were analysed using power-spectrum analysis for overall power (0.1-13Hz), alpha (8-13Hz) and theta (4-8Hz) bands. The differences in conditions are expressed as an increase or decrease in power given across the 1Hz bins for the conditions being compared e.g. NS-D-ND is a subtraction of the results found in the normal sleep distraction condition to the normal sleep no distraction condition. The large individual differences make parametric analysis unsuitable, so general trends are investigated using nonparametric techniques. Changes in power in the bins are summed for all participants and shown in the tables below.

Sleepiness condition effects were gauged by subtracting the root-mean square of each bin for each subject in sleep-restricted conditions to the corresponding bins in the normal-sleep condition. This within-subjects analysis therefore gives an accurate account of the changes in power under sleep-restriction, if any exist. For all bins, there was parity between sleepiness condition and non-sleepy conditions, with no reason to believe that overall power changed after sleep restriction (Table 26). An increase (inc.) represents when a bin was higher in power during sleepiness condition, whereas a decrease (dec.) means the bin was higher in power in the normal sleep category. At the very least, there should be a significant power change within-subjects if EEG measures are to show sleepiness. However, like with the main results there is no indication of either a uniform increase or decrease in power, suggesting that sleepiness is not heightened.

Table 28 shows the count of power increases at all bins for all conditions, where any bin that has an increase between conditions is shown as one in "count of increased" power and the opposite if there was a decrease. There is a general trend towards increases in power for both young and older groups in the more challenging conditions (working on the assumption that DC > D > N). For frontal sites, the younger group had 143 more bins that increased in power than decreased and 168 for the older group, while for parietal sites the figures were larger at 266 bins for young and 218 for older. Put another way, 66% of parietal site bin activity increased with more difficult conditions for the younger group, but this figure was only 62% for the older group.
**Table 26:** Sleepiness condition changes across the conditions for the two brain regions collapsed across the age groups. There is no strong evidence that either brain region is activated with sleep restriction, which follows the main results.

<table>
<thead>
<tr>
<th>No Distraction</th>
<th>Distraction</th>
<th>Cognitive Distraction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Frontal</td>
<td>Parietal</td>
</tr>
<tr>
<td>Inc.</td>
<td>150</td>
<td>131</td>
</tr>
<tr>
<td>Dec.</td>
<td>162</td>
<td>181</td>
</tr>
<tr>
<td>%</td>
<td>48.1</td>
<td>42.0</td>
</tr>
</tbody>
</table>

**Table 27:** Alpha changes with sleepiness condition for the two brain regions, collapsed across the age groups. There are no strong results that might suggest that alpha activity is heightened significantly.

<table>
<thead>
<tr>
<th>No Distraction</th>
<th>Distraction</th>
<th>Cognitive Distraction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Frontal</td>
<td>Parietal</td>
</tr>
<tr>
<td>Inc.</td>
<td>44</td>
<td>39</td>
</tr>
<tr>
<td>Dec.</td>
<td>52</td>
<td>57</td>
</tr>
<tr>
<td>%</td>
<td>45.8</td>
<td>40.6</td>
</tr>
</tbody>
</table>

**Table 28:** Tabulation of overall power bin (0.1-13Hz = 13 bins) changes comparing all conditions. NS = Normal Sleep, SR = Sleep Restricted, ND = No Distraction, D = Peripheral Distraction, DC = Cognitive Distraction.

<table>
<thead>
<tr>
<th></th>
<th>Young</th>
<th>Old</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NS</td>
<td>SR</td>
<td>NS</td>
</tr>
<tr>
<td>Count of Increased Power</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frontal</td>
<td>D-ND</td>
<td>84</td>
<td>101</td>
</tr>
<tr>
<td></td>
<td>DC-ND</td>
<td>68</td>
<td>103</td>
</tr>
<tr>
<td></td>
<td>DC-D</td>
<td>46</td>
<td>110</td>
</tr>
<tr>
<td></td>
<td>D-ND</td>
<td>88</td>
<td>109</td>
</tr>
<tr>
<td>Parietal</td>
<td>DC-ND</td>
<td>74</td>
<td>117</td>
</tr>
<tr>
<td></td>
<td>DC-D</td>
<td>67</td>
<td>107</td>
</tr>
<tr>
<td>Sum</td>
<td>427</td>
<td>647</td>
<td>574</td>
</tr>
<tr>
<td>Count of Decreased Power</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frontal</td>
<td>D-ND</td>
<td>71</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td>DC-ND</td>
<td>87</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>DC-D</td>
<td>110</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>D-ND</td>
<td>68</td>
<td>34</td>
</tr>
<tr>
<td>Parietal</td>
<td>DC-ND</td>
<td>69</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>DC-D</td>
<td>76</td>
<td>36</td>
</tr>
<tr>
<td>Sum</td>
<td>481</td>
<td>184</td>
<td>332</td>
</tr>
</tbody>
</table>

A similar analysis was performed to that in Table 28, with the exception that only alpha band bins were included (8-13Hz) for Table 29. There is a slight trend towards increases in alpha in more challenging conditions for both young and old at frontal and parietal sites with a range of 11 for parietal DC-D (53%) to 50 for parietal DC-ND (63%). Given that 50% represents completely chance findings, these trends are not strong enough to inspire confidence to
conclude that alpha can predict performance. The theta results (Table 30), on the other hand, unequivocally show no trend towards an increase or decrease in any of the performance conditions.

Table 29: Tabulation of alpha power (8-13Hz) changes for all conditions. There is an alpha power increase for sleep restriction in young subjects, and for the more challenging distraction conditions. Older subjects show the opposite results for alpha sleepiness condition.

<table>
<thead>
<tr>
<th></th>
<th>Young</th>
<th></th>
<th></th>
<th>Old</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NS</td>
<td>SR</td>
<td>NS</td>
<td>SR</td>
<td>Sum</td>
<td></td>
</tr>
<tr>
<td>Count of</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Increased</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Power</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frontal</td>
<td>D-ND</td>
<td>25</td>
<td>33</td>
<td>23</td>
<td>24</td>
<td>105</td>
</tr>
<tr>
<td></td>
<td>DC-ND</td>
<td>26</td>
<td>36</td>
<td>26</td>
<td>22</td>
<td>110</td>
</tr>
<tr>
<td></td>
<td>DC-D</td>
<td>15</td>
<td>33</td>
<td>26</td>
<td>30</td>
<td>104</td>
</tr>
<tr>
<td></td>
<td>D-ND</td>
<td>25</td>
<td>32</td>
<td>26</td>
<td>28</td>
<td>111</td>
</tr>
<tr>
<td>Parietal</td>
<td>DC-ND</td>
<td>21</td>
<td>35</td>
<td>29</td>
<td>30</td>
<td>115</td>
</tr>
<tr>
<td></td>
<td>DC-D</td>
<td>23</td>
<td>29</td>
<td>23</td>
<td>22</td>
<td>97</td>
</tr>
<tr>
<td>Sum</td>
<td>135</td>
<td>198</td>
<td>153</td>
<td>156</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Count of</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Decreased</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Power</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frontal</td>
<td>D-ND</td>
<td>23</td>
<td>11</td>
<td>21</td>
<td>24</td>
<td>79</td>
</tr>
<tr>
<td></td>
<td>DC-ND</td>
<td>22</td>
<td>4</td>
<td>22</td>
<td>26</td>
<td>74</td>
</tr>
<tr>
<td></td>
<td>DC-D</td>
<td>33</td>
<td>11</td>
<td>18</td>
<td>18</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>D-ND</td>
<td>23</td>
<td>12</td>
<td>22</td>
<td>19</td>
<td>76</td>
</tr>
<tr>
<td>Parietal</td>
<td>DC-ND</td>
<td>23</td>
<td>5</td>
<td>19</td>
<td>18</td>
<td>65</td>
</tr>
<tr>
<td></td>
<td>DC-D</td>
<td>21</td>
<td>15</td>
<td>25</td>
<td>25</td>
<td>86</td>
</tr>
<tr>
<td>Sum</td>
<td>145</td>
<td>58</td>
<td>127</td>
<td>130</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 30: Tabulation of theta power (4-8Hz) changes for all conditions. There is no clear link between theta power and distraction condition with the exception of frontal D-ND for young participants.

<table>
<thead>
<tr>
<th></th>
<th>Young</th>
<th></th>
<th></th>
<th>Old</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NS</td>
<td>SR</td>
<td>NS</td>
<td>SR</td>
<td>Sum</td>
<td></td>
</tr>
<tr>
<td>Count of</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Increased</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Power</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frontal</td>
<td>D-ND</td>
<td>18</td>
<td>22</td>
<td>4</td>
<td>23</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td>DC-ND</td>
<td>14</td>
<td>24</td>
<td>23</td>
<td>21</td>
<td>82</td>
</tr>
<tr>
<td></td>
<td>DC-D</td>
<td>10</td>
<td>25</td>
<td>29</td>
<td>22</td>
<td>86</td>
</tr>
<tr>
<td></td>
<td>D-ND</td>
<td>22</td>
<td>23</td>
<td>26</td>
<td>25</td>
<td>96</td>
</tr>
<tr>
<td>Parietal</td>
<td>DC-ND</td>
<td>17</td>
<td>27</td>
<td>29</td>
<td>21</td>
<td>94</td>
</tr>
<tr>
<td></td>
<td>DC-D</td>
<td>12</td>
<td>24</td>
<td>25</td>
<td>17</td>
<td>78</td>
</tr>
<tr>
<td>Sum</td>
<td>93</td>
<td>145</td>
<td>136</td>
<td>129</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Count of</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Decreased</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Power</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frontal</td>
<td>D-ND</td>
<td>18</td>
<td>11</td>
<td>20</td>
<td>13</td>
<td>62</td>
</tr>
<tr>
<td></td>
<td>DC-ND</td>
<td>22</td>
<td>6</td>
<td>13</td>
<td>15</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>DC-D</td>
<td>26</td>
<td>11</td>
<td>4</td>
<td>14</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>D-ND</td>
<td>14</td>
<td>10</td>
<td>10</td>
<td>11</td>
<td>45</td>
</tr>
<tr>
<td>Parietal</td>
<td>DC-ND</td>
<td>16</td>
<td>3</td>
<td>7</td>
<td>15</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td>DC-D</td>
<td>21</td>
<td>9</td>
<td>11</td>
<td>19</td>
<td>60</td>
</tr>
<tr>
<td>Sum</td>
<td>117</td>
<td>50</td>
<td>65</td>
<td>87</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Performance in each condition, as measured by A', was ranked into order from best to worst for each participant, split by sleep restriction and normal sleep. If the best performance rank coincided with the smallest alpha result, as per our hypothesis, then it was marked as 1 for agreement. A score of 8 or less represents chance results, with anything more tending towards the alternative hypothesis. Figure 54 shows that alpha results at frontal and parietal sites was a poor indicator of performance, with the largest score being 13 out of a possible 24.

Figure 54: Predictions of performance ranks across the conditions to alpha level rank. The black line represents chance, with any results below it supporting the null hypothesis. None of the conditions are sufficiently above chance to suggest a link between alpha levels and performance.

By contrast to the alpha results, it was expected that higher levels of theta may predict attentional processes that are linked to improved performance. Figure 55 shows that the predictive potential of the theta bandwidth is very poor as few results even exceed the chance levels, let alone significantly improve on them. This holds true for younger and older participants for both frontal and parietal sites.
Figure 55: Predictions of performance ranks across the conditions to theta level rank. The black line represents chance, with any results below it supporting the null hypothesis. None of the conditions are sufficiently above chance to suggest a link between alpha levels and performance.

A parametric analysis was compared between gross brain activity average at different regions for young and older subjects. Figure 56 shows that the older group, as a population, show heightened brain activity at every region during the normal sleep conditions. A similar relationship is found for sleep restriction (Figure 57), where the older subjects show consistently higher brain activity than young subjects. Large individual differences can clearly be seen in the error bars (standard deviations).
Figure 56: Brain power at different regions during normal sleep. Almost across the board the older age group are showing heightened brain region activity than the younger group.

Figure 57: Brain power at different regions during sleep restriction. A similar result is found to that shown in normal sleep, where older subjects have higher brain activity than younger subjects.
Discussion

The learning effects modelled using the practice day of this study suggest that even for the simplest of luggage search tasks it takes more than an hour to approach mastery in visual search. However, the X-Ray ORT is designed specifically to measure visual search ability as opposed to knowledge effects, using a small library of guns and knives that are shown prior to each session and using monochrome display. In much the same vein as the PVT, the X-Ray ORT was built to measure cognitive abilities with as much parsimony as possible to eliminate irrelevant variables and is therefore ideal for laboratory study. Yet despite the relative simplicity of this task compared to the occupation, where any number of complex threats need to be searched for such as improvised explosive devices and other prohibited luggage items, the older group never managed to approach the levels of performance found in the younger group. In fact, they were not only less accurate in all conditions, but they were also appreciably slower.

The rate of information acquisition is a pertinent issue in airport security due to the disturbingly high levels of staff turnover, meaning that screeners need to be brought to a high level of competency fast as airports cannot rely on retention of staff for long terms. Hainmuller and Lemnitzer (2003) argued that staff turnover is one of the major causes of security concern at airports given the unattractive, repetitive and stressful nature of the job. The Department of Transportation (as cited by Hainmuller and Lemnitzer) found that only 14% of 993 screeners were still working at the same airport in the space of a year. Therefore it is important to establish training norms for all ‘risk categories,’ of which older workers are plainly at a disadvantage. It should be stressed that all twelve of our subjects were ex-professionals from the University of the Third Age and were screened thoroughly for normality. With training, both groups became progressively faster and more accurate on the task, although the older group never managed to perform as accurately or as quickly as the younger group’s second practice at the task.

The principle findings from this study centred around the differences between the younger and older working age limits in simple luggage search. There is only a small trend towards distraction being a factor for the young group, but visual inspection of Figure 51 shows the effects of levels of distraction on the older group in both detection and reaction-time domains. The medium effect-size of distraction for detection performance (A') was attributable to the older group and there was a large distraction effect for speed (RT), although the inability to conjoin detection and reaction time makes statistical inferences less utile. This is important because speed-accuracy operating characteristic theory shows that when subjects take longer to evaluate an image or task demand it is usually to offset the image difficulty (Wickelgren, 1977). Because no bias average achieved experimental significance in any condition, there is no evidence that strategy changed during the experimental conditions.

Earlier chapters detailed laboratory studies that have shown older subjects to be less able to prohibit irrelevant information such as peripheral distractions (e.g. Rabbit, 1965). Our findings suggest a similar phenomenon, and post-
study conversations with our older participants anecdotally confirmed the
difficulty they attributed to trying to multi-task the verbal fluency task while
completing the X-Ray ORT simultaneously. This paradigm served to
standardise the level of cognitive distraction for all subjects, requiring a verbal
output to confirm cognitive processing. That there was a near mirroring of the
distraction effect for the sleep-restricted and normal days provides strong
evidence of the reproducibility and reliability of these results. It is likely that
distraction effects in the real world will be accentuated during busy times as
streams of passengers make their way through security, so these results may
underestimate the likely effect of distraction on-the-job.

Under sleep restriction the younger group slept 46% less than their normal
sleep quota, and for the older group this was similar at 42%. It is apparent that
the five-hour limit of sleep on the sleep restriction day, verified by actiwatch
and telephone calls, was not sufficiently damaging to impair mental
functioning in any measurable way on this task. Additionally, there was no
interaction between sleepiness condition and distraction for young or older
participants, eliminating any potential role of sleepiness on performance. Only
extraversion was different between the younger and older participants. Its
converse, introversion, is linked to caution (Tune, 1966), which may in part
explain the differences in speed between the groups. However, it is unlikely
that these personality differences account for very much of the difference in
performance, and other researchers have suggested that introverts, who are
more meticulous and prone to excellent sustained attention, should perform
better in detection terms (Eysenck and Eysenck, 1976; Matthews et al., 1990).
The overall better performance in the younger group might be linked to the
technological skills that younger people find second nature (Laguna and
Babcock, 2000; Wiegmann et al., 2006). Interestingly, trait arousal predicted
performance in the older group for both detection accuracy and speed when a
split-half analysis was undertaken.

The EEG results yielded few notable findings for the X-Ray ORT, showing
neither a sleepiness condition effect nor any strong evidence that alpha and
theta bands can predict performance. Non-parametric analyses were
undertaken to minimise the inter-subject variances, which are considerable,
with the exception of sleepiness condition measures that were analysed
within-subjects. Alpha power appears to be the most likely candidate for
performance prediction in this luggage-search task depending on the task
difficulty. Theta power levels were as likely to be increased in subjects as they
were decreased in any of the conditions, which rules out any possibility of a
link between distraction level and theta activity. Additionally, alpha and theta
levels were poor predictors of performance for both frontal and parietal sites,
rarely managing to exceed chance levels of prediction. There is a good
empirical reason to use non-parametric measures here, as parametric
measures would mistakenly suggest a proportional link between brain activity
and performance e.g. a 0.1μV increase in alpha activity equals a 0.05
increase in A', which is an unreasonable assertion to make. Ranked data
avoids this problem of individual differences and could have shown, simply,
that as alpha levels increase performance decreases.
Notably, the older group showed consistently higher topographical brain region activity than the younger group. This may at first appear to be the converse of what could be expected, as older people are known to have slightly thicker skulls, which inhibits EEG signals (Behrents, 1985). This increased activity is likely to be a compensatory measure to the task difficulty, as the older subjects focused more attentional resources onto the task.

Of course, there are limits to how closely the screeners’ job can be simulated in the laboratory. The target prevalence effect is a known source of controversy in this area of research as most screeners will go their entire careers without ever finding a threat that was deliberately concealed. Wolfe, Horowitz and Kenner (2005) found that only 7% of targets were missed at a prevalence of 50%, whereas 30% of targets were missed when the prevalence rate was 1%. This is not unique to airline screening: in medical imaging by radiologists, a staggering thirty percent of malignancies are missed (Berlin, 1994).

Yet medical imaging and airport screening are not a fair comparison because of technological advances in the latter known as Threat Image Projection (TIP). TIP systems superimpose threat items into actual passenger luggage x-ray images, and the results are stored and used as performance evaluation in most modern airports. The major advantage of this system is to keep screeners alert who otherwise may find little incentive in seeking-out potential threat items that, in all likelihood, aren’t there. Also, a recent paper by Wolfe et al. (2007) deduced that it is not detection ability that is impacted (e.g. A’), but rather it is the subjective bias (B”) that changes. To re-explore an earlier example, 30% of targets were missed at the 1% level, but this was because subjective bias shifted towards clicking "no threat" due to the string of non-threat items. Searches for items terminated increasingly early and so when threat items did occur the subjects did not spend much time looking for them (Wolfe, 2005). This is why a combined detection measure is used for threat and non-threat items such as A’ and d’; because statistics that talk in terms of ‘targets missed’ do not detail the flipside, which is how often a rater tends to click “threat present” when there are no threats shown.

Laboratory studies seek to determine detection capability under different experimental manipulations, so target rates such as 1% are simply unworkable and achieve little when determining whether experimental manipulations were the cause of study effects or whether it was a bias or chance effect. By means of example, a 1% target rate would show just two threat items in an entire X-Ray ORT session: clearly not separable from chance (A’ could only be 0 for no hits, 0.5 for one hit, and 1 for two hits). A 50% target rate is ideal for determining target detection capability and provides enough statistical power to calculate signal-detection measures with high reliability. Persons who struggle to detect targets at 50% prevalence are not more likely to detect items at 1% prevalence, or any other target presentation rate.

Combined measures such as the TCM and Balakrishnan’s d are a tempting measure to use for empirical research, but they are both flawed to some
extent. The TCM helps approximate a search and decision time component of tasks, but trying to ascertain the degree of each of these from retrospective analyses of detection rates and reaction times is bound to be something of an inexact science without the use of eye-tracking devices. Yet, like our probability eye-open lapses construct (Chapter Three), these benchmarks are certainly useful either for comparative purposes or for when expensive technology and data analysis are simply unusable for time or financial purposes. Balakrishnan's d, on the other hand, has one significant flaw; values may increase simply because of increased speed, even if accuracy decreases. This is evidently an undesirable artefact of any proposed 'singular' measure of detection and reaction time. To be truly worthwhile, a singular measure has to control for detection accuracy, which is of far greater importance than the speed of performance. As of now, no such performance metric exists.

The task duration of the X-Ray ORT, typically between 20-30mins, is in line with international airport regulations (Ainsworth, 2003), as screeners work on a three-way rotating roster of tasks every twenty minutes. These enforced breaks from screening are a brilliant human-factors led approach to reducing performance satiation and the effects of distraction. It is also well known that self-paced tasks like luggage search are performed better than forced-pace tasks (Broadbent, 1964; Ratcliff, 2006). This sort of approach to fit the job to human capabilities are what separate airport security from the vocations discussed in Chapters Three and Four, such as long-haul driving and plant control.

**Research Questions Summary**

1. How quickly can a naive subject reach peak performance on a simulated luggage search task? Does the age of the subject matter?

   Data modelling showed that subjects reached their performance plateau somewhere between their third and fourth session. This was sufficient for both young and old subjects to perfect their visual search ability. However, it should be noted that visual search is distinct from visual knowledge, which takes dozens of hours in training to achieve expert status. Although the older group were less accurate comparatively to the young, they also required four sessions to reach their peak.

2. Is luggage-search ability impacted by a night of unsatisfactory sleep? How might this apply to other vocations?

   One night of sleep restriction does not affect how well people can detect threat items in luggage. This may be extended to similar vocations that are perhaps computer based or depend on surveillance. Yet, the impact of multiple-nights of sleep loss or a pre-existing sleep condition cannot be ruled out to affect performance on these tasks.
3. Is there a performance difference between peripheral and cognitive distractions?

There is a slight but insignificant tendency for performance to be less accurate and slower for the young group, and a noticeable difference in the older group. Because of difficulties in conjoining detection and speed measures, graphical representations of both are most useful for making deductions. The largest differences for both groups were between the no distraction and cognitive distraction conditions.

4. What effect does age have when considering the effects of sleepiness and distraction on an engaging task?

The differences between the youngest and oldest working-age groups are significant and appreciable. Older subjects were slower and less accurate in all conditions, and never managed to better the younger group’s third practice at the task. We conclude that older workers should be given compensatory training to overcome this initial deficit.

The results of this chapter have formed two publications:


Related research also in the process of publication:


CHAPTER SIX

Discussion
Research Summary

6.1.1 Research Questions Addressed

1. Does sleepiness interact with distraction? And does distraction alter performance in monotonous or repetitive tasks?

Distraction increases under increasing sleep pressure, with participants seeking out distractions away from their central task. This is true for simple and complex tasks, although to a lesser extent in the latter. However, a task-long decline in performance shows a trend in monotonous tasks, but not for engaging tasks. Rather, it is the risk that comes with spontaneous moments of inattention that are the real danger when distraction increases with sleepiness.

Cognitive distractions are more damaging to performance than peripheral distractions, such that holding a conversation or being lost in thought creates the potential for more errors. Older people are more susceptible to the effects of cognitive distractions than younger people, as younger people can manage concurrent tasks more efficiently.

2. Is sleepiness or distraction differentially affected by task complexity?

Overnight sleepiness affects performance on monotonous tasks but not on more complex tasks. This is also true of sleep restriction, where a night of sleep restriction to less than 5hrs does not impair performance on a complex task. Distraction shows a similar trend, whereby monotonous tasks are more likely to be impacted by sleep loss than complex tasks.

3. Is there an age effect in ecological (job simulated) tasks, and if so does it vary with sleepiness or distraction?

There is a marked age effect in how well older workers can perform a computer-based luggage simulator task, which is significantly slower and less accurate than younger workers. Regarding sleepiness, there does not appear to be a difference in how well younger and older participants coped with the sleep restriction. However, the older participants were more affected by levels of distraction than the younger participants.

6.1.2 Sleepiness Vs Distraction

The findings from Chapter Three provided some unexpected results in addition to the anticipated effects of sleepiness. Distraction measures for lapses did achieve statistical significance as an interaction effect, although the nature of the task may mask the true effects of distraction as subjects learned to sneak glances at the television without a performance deficit. Certainly, the number of head turns greatly increased under sleepiness, but they did not
often result in a specific head-turn lapse. When alert, a pleasurable distraction may actually improve performance in 'boring' tasks, such as listening to music while writing a report. However, under sleepiness these distractions- whether cognitive or peripheral, causing attention away from the central task- are more liable to affect performance than improve it, giving the interaction between sleepiness and distraction. The interaction effect may be small, but it is probable that in longer vigilance durations (such as a truck driver driving eight hours overnight) and being unsupervised will tend to accentuate distraction and sleepiness effects in real-life settings.

After the PVT study, it became clear that peripheral distractions may be the determinant of certain specific performance failures, but that this definition of distraction was unlikely to explain a task-long decline in performance (e.g. mean reaction time). Under the distraction condition, there were vastly more head turns than in the no distraction condition. These head turns may actually be a more appropriate measure of task fitness than the mean reaction time, as they are more likely to relate to a serious performance failure, for example when driving. It is possible that cognitive distractions may be a better conceptualisation of task inattention and therefore task decrement as determined by slowness of response, whereas head turns can represent those short moments of time when momentary and complete inattention to the task may result in a collision or other rapid performance implication.

Peripheral distractions are quantifiable in a behaviourist study paradigm by lapse head turns, but in order to tap into the wider context of distraction a cognitive element would need to be manipulated. Extensive piloting suggested that a verbal-fluency task would approximate a conversation and have the benefits of being both replicable and demonstrative of an input-process-output loop to ensure the distraction had been mentally processed. The results of this cognitive distraction proved to be more damaging to overall performance than a peripheral distraction alone, particularly for the older sample. This inability to cope with multiple demands in older subjects is thought to originate from an altering of prefrontal cortex function (Chao and Knight, 1997), where elderly people are known to have diminished prefrontal cortex function (Albert and Kaplan, 1980). Miller (2000) wrote that “one of the classic signs of PFC damage is increased distractibility; subjects seem unable to focus on a task when other, irrelevant events compete for their attention” (p. 62).

Blagrove, Alexander and Horne (1995) investigated repeated sleep reduction and the ability to ignore distracting, irrelevant stimuli in a series of experiments. They concluded that, even when sleep is reduced to 4.3 hours per night for four consecutive nights, there is no performance deficit on logical reasoning and auditory vigilance. However, there was a significant decline in ability to find figures in a Finding Embedded Figures Task (FEFT), which the authors attributed to increased distractibility. The FEFT intersperses irrelevant stimuli with targets, so is a subtly different conceptualisation of distraction that may share a common underlying mechanism with our distraction manipulations. We have found that the effects of sleep loss manifests itself
differently depending on the task, but that distractibility increases in both simple and complex tasks.

How much sleep an individual requires to perform at satisfactory levels proves difficult to generalise, and is subject to unpredictable human whims and circumstances far beyond experimenter control. When restricting sleep or otherwise requiring participants to work at unusual hours, there is a wide variance in how well people deal with this (Figure 19). For example, subject 11 had 4 lapses in the non-sleepy condition, but 75 under sleepiness. Compare this to subject 24, who had 7 lapses when non-sleepy and 13 lapses when sleepy, and then there was subject 16, who had 82 lapses when non-sleepy and 94 lapses when sleepy. Because of these widely differing magnitudes of effects of sleepiness across subjects, conclusions can become somewhat confusing at a population-level. However, every subject's performance (except subject 6) was worsened under conditions of sleepiness, showing that the effects of sleepiness are likely to impact on the majority of people's performance, albeit by different amounts. Subject six has been shown in the results to be an extreme outlier, suffering a staggering 151 lapses when not sleepy and 139 when sleepy. Video of analysis of his performance showed a subject who was simply disconnected with the task, showing little motivation to get reactions below one second and seemingly unwilling to apply full concentration to the task. It is shown, then, that performance for monotonous tasks is certainly impacted in people who are unaccustomed to working at the circadian nadir. The extent of this performance decline, though, varies largely from individual to individual.

6.1.3 Simple Vs Complex Tasks

The extent that sleepiness will affect performance depends largely on the task parameters, and particularly the nature and difficulty of the task (Dinges and Kribbs, 1991). Wilkinson (1968) noted that for a task to be sensitive to sleep loss it should be boring and last at least half an hour, but preferably an hour (cf. Casagrande, 1997). Yet, acting on this well-intended advice may inadvertently result in study paradigms that are really measuring motivation, even if it is accentuated by sleep loss. Caldwell, Rampsott and Gardner (1998) stated that “at least thirty minutes of continuous performance will be required before generalizations to ‘real-world’ performance can even begin to be made” (p. 14) when using the MATB with sleep loss. These authors found significant sleepiness effects on the MATB using 36-hours of wakefulness compared to our 24-hours, even though they used thirty minutes of task duration the same as we did. Perhaps it is between 24-36 hours of sleep that divergent tasks begin to show sleep-loss effects, as Basner and colleagues (2008) found when using a 36-hour constant routine and luggage simulator. It is unlikely that motivation played any role in the results of Chapters Three and Four as participants were paid a large sum of money to participate and were given four-hours to relax and enjoy snacks between alert and sleepy conditions, which is a luxury that would not typically be afforded in a workplace.
6.1.4 Speed Vs Accuracy

X-ray luggage screening is a time-pressured task that is caught between a difficult balance of speed and accuracy. This is important when viewed from the job perspective, where there is genuine pressure to get customers through a terminal ("throughput") that is placed by the airports themselves. Therefore, luggage screeners have to be able to act quickly and accurately when on the job, or else they risk their performance becoming impaired by the demand to oversee a conveyor-belt of seamless items coming through at busy times. Modelling the changes with accuracy as response-time is particularly necessary for this vocation that depends on efficiency of throughput.

The speed-accuracy operating curve found in Chapter Five was based on the assumption that if screeners take longer to inspect items their accuracy will increase too. Studies that have systematically manipulated detection time, known as a response-signal paradigm (Ratcliff, 2006), have shown this observation to have high validity. Singh and Singh (2004) and Schwaninger, Michel and Bolfing (2007) independently found that screeners take 3-6s to inspect luggage items at airports. Our SAOC findings from the X-Ray ORT for the young and old groups level-off around 5-6 seconds, which is the time taken to achieve maximal performance. It may be then that six seconds is enough time to achieve satisfactory performance in real-life settings also. Any time less than four seconds is bound to cause problems, and it should be remembered that the luggage items used in the study only showed monochrome bags with a known set of guns and knives.

6.1.5 Personality Dimensions

Personality dimensions were carefully collected for all participants to tap into latent dispositions towards aspects of personality that may influence distractibility. These included psychoticism, neuroticism, state and trait anxiety, extraversion, psychoticism, neuroticism, and arousal. It is no surprise that in such a controlled study few of these factors indicated performance differences in split-half analyses (median split). The reason being that statistically significant differences may be small in such a close sample when split into two distinct groups, but the possibility of explaining results by personality differences has to be considered. So although a high and low group were statistically determined, they would not appropriately be described as "highly extraverted" for example.

Rather interestingly, anxiety was a predictor of performance for the young group's detection performance ($A' = 0.932$ for high anxiety, $A' = 0.907$ for low) and the older group's speed ($RT = 4.36s$ to $RT = 3.33s$) in Chapter Five. Easterbrook (1959) stated that highly-anxious individuals tend to focus mainly on task-relevant cues and are less affected by distractions, although Eysenck and Byrne (1992) argue the converse is true. True to this academic discord, our findings are equivocal given that the high anxiety group were more accurate in the younger sample, but slower for the older group. 'Anxious' people, if such a term can be used correctly in a sample of normal subjects,
would be expected to respond quickly and therefore less accurately in a speed-operating model. Arousal levels also predicted performance between the high and low-split older groups for both detection accuracy and response time, although neither were significant for the younger group. With the exception of state anxiety, all personality measures were assumed to measure an underlying trait disposition.

6.1.6 Transport Security

The importance of maintaining a safe vigil has been stated throughout Chapters Three to Five, often expressed in terms of causalities, risk, and financial cost. Frederickson and LaPorte (2002) wrote that “airlines contracted out passenger and baggage screening to firms that paid very low wages, required minimal qualification, and, as consequence, experience high personnel turnover” (p. 35). Despite this, Frederickson and LaPorte note that it is publicly unacceptable to suffer another failure in the air travel system of the magnitude of September 11th 2001, but that high levels of delay are similarly unacceptable in the minds of airline passengers. Even when the risks are known, however, it can be impossible to foresee a potential accident or disaster. Less than a year prior to the terrorist attacks of September 11th, 2001, a representative from the Government Accounting Office testified in front of the House of Representatives:

“A single lapse in aviation security can result in hundreds of deaths, destruction of equipment worth hundreds of millions of dollars, and have immeasurable negative impacts on the economy and the public's confidence in air travel” (p. 1, Dillingham, 2000).

By means of comparison, throughout the year 2001 alone, there were 4,793 fatal collisions involving large trucks in America (Federal Motor Carrier Safety Administration, 2003); and this figure doesn’t include the countless fatalities suffered in car and motorcycle accidents. Yet, it only took one incident in the month of September, 2001 for radical policy changes to sweep across the aviation industry that dwarfed the previous and subsequent budget allocated to preventing motor vehicle accidents. That day, just under 3,000 civilians died in a terrorist attack that is unlikely to be repeated. The budget that is allocated to preventing an airline disaster is massively disproportionate to the budget allocated to preventing motor vehicle accidents (see Section 5.1.1 for overview) when viewed in terms of potential for injuries and fatalities.

Experimental Design: A Discussion

6.2.1 Laboratory And Study Design

Our laboratory provided a sterile environment free from all time cues, temperature deviations, external noises, light levels or any other controllable source of nuisance variable known to affect sleep studies (Bonnet, 1989). In fact, we even went as far as to prepare meals for participants before every testing period, resulting in over 120 meals being made across the studies. Finally, subject characteristics were controlled for by a lengthy screening
process that involved personality assessment, face-to-face interviews and three nights of quantitative sleep assessment through actigraphy and sleep diaries. This was to ensure that the participants and the conditions they undertook the tasks were as similar and free from unwanted variance as could be controlled for.

6.2.2 Electroencephalography

The EEG results did not provide a major insight into the performance of subjects on the tasks used, although the alpha results were a significant predictor of performance in the PVT and the MATB. There was a tendency for alpha power to increase in the distraction conditions compared to no distraction, but there were no significant effects in the theta bandwidth. It may be that the sleep restriction of less than five hours used for the X-Ray ORT was not enough to show as a change in FFT. The hypotheses created before data collection were primarily focused on using alpha and theta bandwidths rather than a focused examination of individual 1Hz bins, so the possibility of these smaller bins being a predictor of performance cannot be ruled out, although this was beyond the scope of this research. The older group showed a clear increase in activation for all brain regions compared to the younger group. It can be suggested that this is due to a compensatory effort for the older group, who may have to allocate additional attentional resources to the task than the younger group.

Future Research

6.3.1 Chronic Sleep Loss

Chronic sleep loss or repeated days of unsatisfying sleep are more damaging to performance than one-night of overnight work or restricted sleep. Yet the more sleep-deprived a subject is, the less relevant their performance is to the general working populace. Studies that use sleep-disordered patients (e.g. insomnia patients) may find that their performance suffers on ecological tasks, but to date this hasn't occurred with either the MATB or a luggage simulator task. Now that it has been established that normal participants are capable of performing these tasks to a similar level as when sleep restricted, more specific populations such as insomnia patients would be a possible next step, or larger bouts of sleep loss could be used in normal subjects.

We posit that overnight work or a bad night’s sleep (< 5hrs) will be the most sleep loss typically experienced by workers in the field of aviation screening, if not the workforce in general. Given that there was absolutely no performance loss, it is safe to say that ‘normal’ people working in comparable vocations (e.g. CCTV operators) can still perform to a high level when losing half of their usual sleep quota for a night. Future research may be concerned with more damaging levels of sleep loss, and it may be necessary to employ a relatively extreme paradigm such as a 48-hour constant routine to see changes in performance. A thinly-veiled critique of existing research by Jones and Harrison (2001) called for “a move towards greater ecological validity in future research” (p. 472) to rally researchers behind more realistic tasks, settings
and sleep manipulations. I believe the studies presented throughout this thesis have addressed these issues, so future research should be concerned with higher levels of sleep loss such as repeated evenings of lost sleep, for example.

6.3.2 Modelling Lapse Behaviour

Often researchers are left to retrospectively determine what behaviours must have occurred at certain reaction times, and this poses significant difficulties. Did a three-second reaction time represent a microsleep, or was the subject not even looking at the screen? While this question can never truly be settled without a time and labour-expensive video-recorded study, it is more than 95% probable that a reaction time given over three seconds was caused with the eyes closed. The potential for microsleeps to explain behaviour diminishes as the reaction times get longer, because a decreasing proportion of reactions occur with the eyes open. Studies that examine microsleeps may be including a large proportion of reactions with the eyes closed, although it is generally assumed that a microsleep involves drooping eyelids (Anderson and Home, 2006). Peiris et al. (2006) found that "the most frequent definite BMs (behavioural microsleeps) are of 2–3 s duration and that the rate of longer events decreases approximately exponentially" (p. 295). Our findings agree with this deduction, but we can go further by saying that as well as the rate of the events decreasing approximately exponentially, so does the very nature of the behavioural responses.

Even though the subjects in this study knew they were being video-recorded in real-time, it was surprising to note just how many reaction times were caused with the eyes closed, especially after a couple of seconds. The decaying exponential curve found between eyes open and eyes closed confirms a long-held suspicion that human behaviour progressively changes as reactions worsen. Very seldomly did the three behaviours of eyes fixed on the task, eyes closed and head facing away from the task fail to categorise the human behaviour during a reaction. Although simple, this behaviourist paradigm summarises the human behaviours that occur during unprepared reaction time tasks almost completely. The decaying exponential curve found in Section 3.3.5 should be extended to a larger sample, different age groups and sleep restrictions to determine a model that is predictive in various circumstances and robust to different conditions. The current model ranges in prediction from 81% accurate for EC lapses at the 50% level to 93% accurate for EO lapses at the 95% confidence level for this sample of young participants. The model should be refined and improved using different predictors to gain a wider insight into what a lapse is. Perhaps then, Bill’s (1931) set of ‘blocking’ questions (Section 3.1.6) may finally be resolved.
6.4.1 Conclusion

In summary, our understanding of the effects of sleepiness and distraction has advanced a lot in recent years with the increasing use of video-technology to understand human behaviour during the task. The task demands are a central factor in the extent to which sleepiness and distraction can affect performance. Monotonous tasks are heavily impacted by sleepiness-induced reduction in attention motivation, which causes people to be more susceptible to the effects of distraction. Cognitive distractions are damaging to performance, as well as the head-turns and eye closures seen in the video footage. Future research will help model the exact probabilities of forced-eye closures under different tasks, settings and sleep manipulations that have begun with the studies conducted for this thesis and Anderson and Horne's (2006) study.
References


Ajzen, I. Attitudes, Personality, and Behavior. 1988, Chicago, IL: Dorsey Press.


Babkoff, H., Torne, D.R., Sing, H.C., Genser, S.G., Taube, S.L., and Hegge, F.W. Dynamic changes in work/rest duty cycles in a study of sleep
Sleepiness and Distraction


Bonnet, M.H., and Arand, D.L. We are chronically sleep deprived. *Sleep*, 1995, 18:10, 908-911.


Bunn, T.L., Slavova, S., Stuttmann, T.W., and Browning, S.R. Sleepiness/fatigue and distraction/inattention as factors for fatal versus


Carmody, M.A. Current issues in the measurement of military aircrew performance: a consideration of the relationship between available metrics and operational concerns. Air Vehicle and Crew Systems Technology Department, Naval Air Warfare Center: Aircraft Division, Warminster, PA.


Chao, L.L., and Knight, R.T. Prefrontal deficits in attention and inhibitory control with aging. Cerebral Cortex, 1997, 7, 63-69.
Chervin, R.D., and Aldrich, M.S. The Epworth Sleepiness Scale may not reflect objective measures of sleepiness or sleep apnea. *Neurology*, 1999, 52, 125.


Chao, L.L., and Knight, R.T. Prefrontal deficits in attention and inhibitory control with aging. *Cerebral Cortex*, 1997, 7, 63-69.


Corsi-Cabrera M, Arce C, Ramos J, Lorenzo I, Guevara MA. Time course of reaction time and EEG while performing a vigilance task during total sleep deprivation. *Sleep*, 1996, 19, 563-569.


Davies, R.J.O., Bennett, L.S., and Stradling, J.R. What is an arousal and how should it be quantified? *Sleep Medicine Reviews*, 1997, 1:2, 87-95.


Gray, Y.A. The psychophysiological basis of introversion-extraversion. Behavior Research And Therapy, 1970, 8, 249-266.


Hartley, L.R. Fatigue and driving—driver impairment, driver fatigue and driving simulation. *Taylor & Francis, London*. 


Lyznicki, J.M., Doege, T.C., Davis, R.M., and Williams, M.A. Sleepiness, driving, and motor vehicle crashes. *JAMA*, 1998, 279, 1908-1913.


Tassi, P., Bonnefond, A., Engasser, O., Hoeft, A., Eschenlauer, R., and Muzet, A. EEG spectral power and cognitive performance during sleep inertia:


## Appendix A: Glossary

<table>
<thead>
<tr>
<th>Word</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bottom-up</td>
<td>Features in the world or surroundings that 'grab attention.'</td>
</tr>
<tr>
<td>Canonical view</td>
<td>The typical view that would be expected to show an object in its most familiar form.</td>
</tr>
<tr>
<td>Nomogram</td>
<td>A two-dimensional diagram used to approximate the computational value of a function.</td>
</tr>
<tr>
<td>Isoleth</td>
<td>A line drawn on a map through all points having the same numerical value.</td>
</tr>
<tr>
<td>Salient</td>
<td>Areas that are most obvious to visual scanning when an image is instantly produced.</td>
</tr>
<tr>
<td>Threat-Image Projection</td>
<td>A system of projecting fictional threat items into actual images of luggage operational in hundreds of airports around the world. Commonly used to assess on the job performance.</td>
</tr>
<tr>
<td>fMRI</td>
<td>Functional Magnetic-Resonance Imaging. A method of brain imaging that is sensitive to changes in oxygen content of the blood.</td>
</tr>
<tr>
<td>Oddball Paradigm</td>
<td>Two auditory tones are intermixed in a series, with participants having to respond to the target tone. This tends to generate a positive ERP response known as the P3.</td>
</tr>
<tr>
<td>ERP</td>
<td>Event-Related Potential. An EEG signal generated after a specific event.</td>
</tr>
<tr>
<td>Latent</td>
<td>Inherent to the organism; observable only by inference.</td>
</tr>
<tr>
<td>Manifest Trait</td>
<td>Directly observable and measurable phenomena.</td>
</tr>
<tr>
<td>State</td>
<td>A flexible aspect of personality that can change rapidly (e.g. alertness).</td>
</tr>
<tr>
<td>Iconic Memory</td>
<td>A very short-term visual memory store.</td>
</tr>
<tr>
<td>Sleep Apnea</td>
<td>A sleep-disorder that occurs from obstructed breathing airways that collapse during rapid-eye movement sleep.</td>
</tr>
<tr>
<td>Prophylactic</td>
<td>Preventative in nature.</td>
</tr>
<tr>
<td>Soporific</td>
<td>Sleep inducing.</td>
</tr>
<tr>
<td>Phasic</td>
<td>Event-related.</td>
</tr>
<tr>
<td>Sleep Inertia</td>
<td>Ill-feeling upon awakening</td>
</tr>
<tr>
<td>Receiver-Operating Curves</td>
<td>A method of presenting signal-detection results that combines detection ability, bias and confidence in one image.</td>
</tr>
</tbody>
</table>
Appendix B: PVT and MATB Distraction Scoring Card

Participant ________________________________

Task PVT / MAT

Condition Night / Morning Distraction / No Distraction

EC = Eyes Closed  EO = Eyes Opened  HT = Head Turn

<table>
<thead>
<tr>
<th>Time</th>
<th>TURNS</th>
<th>LAPSES</th>
<th>Time</th>
<th>TURNS</th>
<th>LAPSES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>short</td>
<td>long</td>
<td></td>
<td>short</td>
<td>long</td>
</tr>
<tr>
<td>00:00</td>
<td></td>
<td></td>
<td>15:00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>00:30</td>
<td></td>
<td></td>
<td>15:30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>01:00</td>
<td></td>
<td></td>
<td>16:00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>01:30</td>
<td></td>
<td></td>
<td>16:30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>02:00</td>
<td></td>
<td></td>
<td>17:00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>02:30</td>
<td></td>
<td></td>
<td>17:30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>03:00</td>
<td></td>
<td></td>
<td>18:00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>03:30</td>
<td></td>
<td></td>
<td>18:30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>04:00</td>
<td></td>
<td></td>
<td>19:00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>04:30</td>
<td></td>
<td></td>
<td>19:30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>05:00</td>
<td></td>
<td></td>
<td>20:00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>05:30</td>
<td></td>
<td></td>
<td>20:30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>06:00</td>
<td></td>
<td></td>
<td>21:00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>06:30</td>
<td></td>
<td></td>
<td>21:30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>07:00</td>
<td></td>
<td></td>
<td>22:00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>07:30</td>
<td></td>
<td></td>
<td>22:30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>08:00</td>
<td></td>
<td></td>
<td>23:00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>08:30</td>
<td></td>
<td></td>
<td>23:30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>09:00</td>
<td></td>
<td></td>
<td>24:00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>09:30</td>
<td></td>
<td></td>
<td>24:30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10:00</td>
<td></td>
<td></td>
<td>25:00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10:30</td>
<td></td>
<td></td>
<td>25:30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11:00</td>
<td></td>
<td></td>
<td>26:00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11:30</td>
<td></td>
<td></td>
<td>26:30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12:00</td>
<td></td>
<td></td>
<td>27:00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12:30</td>
<td></td>
<td></td>
<td>27:30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13:00</td>
<td></td>
<td></td>
<td>28:00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13:30</td>
<td></td>
<td></td>
<td>28:30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14:00</td>
<td></td>
<td></td>
<td>29:00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14:30</td>
<td></td>
<td></td>
<td>29:30</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix C: The International 10-20 Electrode Placement System
Appendix D: Fast-Fourier EEG Analyses
Sleepiness and Distraction

The graphs depict the change from baseline in various frequency bins (Hz) for different subjects. Each graph shows the data for a specific subject, with lines representing different conditions or categories.
Sleepiness and Distraction

Frequency Bin (Hz)

Change From Baseline (%)
Sleepiness and Distraction 201

Frequency Bin (Hz)

Change from Baseline (%)
Sleepiness and Distraction 205

[Graphs showing changes in sleepiness and distraction across frequency bins.]
Appendix E: Study Questionnaires
# Sleep Research Centre

## Screening Questionnaire: CONFIDENTIAL

<table>
<thead>
<tr>
<th>Field</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name:</td>
<td></td>
</tr>
<tr>
<td>Address:</td>
<td></td>
</tr>
<tr>
<td>Phone Number:</td>
<td></td>
</tr>
<tr>
<td>National Ins:</td>
<td>NB: This must be included for payment, which is handled by finance</td>
</tr>
<tr>
<td>Age/D.O.B.</td>
<td></td>
</tr>
<tr>
<td>Sex:</td>
<td></td>
</tr>
<tr>
<td>Occupation:</td>
<td></td>
</tr>
<tr>
<td>R/L Handed:</td>
<td></td>
</tr>
</tbody>
</table>

### GENERAL QUESTIONS

1. Do you smoke?  
   - Yes  
   - Sometimes  
   - No

1a. If yes, How many cigarettes per day?  
   - 1-5  
   - 5 or more  
   - Don't Know

2. How many cups of tea/coffee do you usually drink in a day?  
   - None  
   - 1-2  
   - 3-4  
   - 5-6  
   - Over 6  
   - Don't Know

3. Are you available to come into the Sleep Centre any day of the working week?  
   - No  
   - Yes

3a. If no, what week days are you available?  
   - Monday  
   - Tuesday  
   - Wednesday  
   - Thursday  
   - Friday

### HEALTH QUESTIONS

4. In general would you say your health is:  
   - Excellent
5. Have you ever experienced any of the following medical conditions and if so, when?

<table>
<thead>
<tr>
<th>No = 1</th>
<th>Yes, sometimes = 3</th>
<th>Yes in the past = 2</th>
<th>Yes, at present = 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Asthma</td>
<td>(b) Dyslexia</td>
<td>(c) Eczema</td>
<td>(d) Allergies</td>
</tr>
<tr>
<td>(e) Thyroid Problems</td>
<td>(f) Undue anxiety</td>
<td>(g) Sleepwalking</td>
<td>(h) Loud snoring</td>
</tr>
<tr>
<td>(i) Nightmares</td>
<td>(j) Bruxism</td>
<td>(k) Difficulty reading/writing</td>
<td>(l) Arthritis/Rheumatism</td>
</tr>
<tr>
<td>(m) Depression</td>
<td>(n) Heart problems</td>
<td>(o) Stomach problems</td>
<td>(p) Waking up with a jolt</td>
</tr>
<tr>
<td>(q) Waking up excessively early</td>
<td>(r) Difficulty falling asleep</td>
<td>(s) Stress/anxiety at home/work</td>
<td>(t) Epilepsy</td>
</tr>
<tr>
<td>(u) Migraine</td>
<td>(v) Colour blindness</td>
<td>(w) Hearing Problems</td>
<td>(x) Diabetes</td>
</tr>
<tr>
<td>(y) Chronic Fatigue Syndrome</td>
<td>(z) Restless Leg Syndrome</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

6. Do you regularly take pills or medicines from the chemist or by prescription? Yes | No | Don't Know |

If so can you tell me what they are?

SLEEP QUESTIONS

7. How well do you feel that you sleep generally?

<table>
<thead>
<tr>
<th>Very well</th>
<th>Well</th>
<th>Not very well</th>
<th>Poorly</th>
</tr>
</thead>
</table>

8. What time do you normally go to bed?

9. What time do you normally get up?

10. How long does it normally take you to fall asleep?

<table>
<thead>
<tr>
<th>0-5 minutes</th>
<th>5-10 Minutes</th>
<th>10-20 Minutes</th>
<th>20-30 Minutes</th>
<th>Over 30 Minutes</th>
<th>Don’t know</th>
</tr>
</thead>
</table>

11. How much sleep do you feel you need each night?

| Less than 5 hours |
12. Do you ever miss a night's sleep or have much more sleep than usual?
   - No
   - Yes, sometimes
   - Yes, regularly
   - Don't know

12a) If yes, can you tell me what is the reason for this?

13. How much does your quality of sleep vary from one night to the next?
   - Very much
   - Moderately
   - Slightly
   - Not at all
   - Don't know

14. How many times do you wake, on average, a night?
   - Never
   - Once
   - Twice
   - More than twice
   - Don't know

14a) If you wake up: How long does it take you to get back to sleep again?
   - Less than 10 minutes
   - 10 – 30 Minutes
   - 30 – 60 Minutes
   - Over 60 Minutes
   - Don't know

15. Do you ever feel sleepy during the day?
   - Yes every day
   - Yes, several times a week
   - Yes, several times a month
   - Yes, once a month
   - Never
   - Don't know

15a) If yes, at about what time does this sleepiness usually start?

16. Do you ever nap during the day?
   - Yes
   - No
   - Don't Know

16a) If yes, how often on average?
   - Every Day
   - 2-3 Times per week
16b) If yes, why?
- Boredom
- Inadequate sleep / sleepiness
- Routine
- Hang over / Late night
- No reason

17. Do you ever experience ‘poor sleep’?
- Yes
- Sometimes
- No
- Don’t know

18. If you had a poor night’s sleep, does it affect:
- How you feel
- How you perform
- Both of these
- Neither of these
- Don’t know

19. If you had a poor night’s sleep, when do you feel the consequences?
- The next day
- The day after
- Both of these days
- Neither of these days

20. Please Complete the Following:

How likely are you to fall asleep in the following situations? Please indicate, using the following scale, which is most appropriate given the situation.

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

<table>
<thead>
<tr>
<th>Situation</th>
<th>Chance of Dozing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sitting and Reading</td>
<td></td>
</tr>
<tr>
<td>Watching TV</td>
<td></td>
</tr>
<tr>
<td>Sitting inactive in a public place (e.g. theatre/meeting)</td>
<td></td>
</tr>
<tr>
<td>As a passenger in a car for an hour without a break</td>
<td></td>
</tr>
<tr>
<td>Lying down in the afternoon when circumstances permit</td>
<td></td>
</tr>
</tbody>
</table>
Sleepiness and Distraction

Sitting and talking to someone
Sitting quietly after lunch without alcohol
In a car, while stopped for a few minutes in the traffic

TOTAL
21. Considering your own “feeling best” rhythm, at what time would you get up if you were entirely free to plan your day?

<table>
<thead>
<tr>
<th>Time</th>
<th>Selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>5am – 6.30am</td>
<td></td>
</tr>
<tr>
<td>6.30am – 7.45am</td>
<td>☐</td>
</tr>
<tr>
<td>7.45am – 9.45am</td>
<td>☐</td>
</tr>
<tr>
<td>9.45am – 11am</td>
<td></td>
</tr>
<tr>
<td>11am – 12noon</td>
<td></td>
</tr>
</tbody>
</table>

22. Considering your own “feeling best” rhythm, at what time would you go to bed if you were entirely free to plan your day?

<table>
<thead>
<tr>
<th>Time</th>
<th>Selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>8pm – 9pm</td>
<td></td>
</tr>
<tr>
<td>9pm – 10.15pm</td>
<td></td>
</tr>
<tr>
<td>10.15pm – 12.30am</td>
<td>☐</td>
</tr>
<tr>
<td>12.30am – 1.45am</td>
<td>☐</td>
</tr>
<tr>
<td>1.45am – 3am</td>
<td></td>
</tr>
</tbody>
</table>

23. Do you consider yourself to be a “morning” or “evening” type of person?

<table>
<thead>
<tr>
<th>Type</th>
<th>Selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morning</td>
<td></td>
</tr>
<tr>
<td>More morning than evening</td>
<td>☐</td>
</tr>
<tr>
<td>More evening than morning</td>
<td>☐</td>
</tr>
<tr>
<td>Evening</td>
<td></td>
</tr>
</tbody>
</table>

24. Have you ever worked unsocial hours, e.g. night shift work?

<table>
<thead>
<tr>
<th>Worked</th>
<th>Selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td></td>
</tr>
</tbody>
</table>

a.w.j.wales@lboro.ac.uk