1 Introduction

1.1 Chapter abstract

The definition of sleepiness is elusive. There are several overlapping concepts related to sleepiness: a transition state, wake state instability, sleep propensity, a set of behaviours and perceptual changes, and the need for sleep. There are many methods to measure sleepiness by operationalising various concepts of sleepiness. Questionnaire methods are easy to administer and may relate to time periods longer than the immediate, but are prone to bias. Objective methods, including measures of reaction time slowing, cognition slowing or latency to sleep are still prone to environmental and motivational influences, and may measure sleepiness over a shorter time period. There is no single 'gold standard' measure of sleepiness, and none directly probe underlying neurophysiological processes controlling the transition between wake and sleep.

The electroencephalogram (EEG) has been studied as a marker of sleep pressure. It has been shown to respond to increasing sleepiness from sleep deprivation. A proposed interaction between cortex and thalamic areas explains the process of sleep transition, the modulation of sleep depth, and describes the changes observed on the surface EEG as sleep deepens.

This study will examine a novel method to measure sleepiness, based on a mathematical model that fits power spectra derived from EEG recordings. The measure will be tested in a disease model of sleepiness, obstructive sleep apnea (OSA). A reliable and objective measure of sleepiness would be invaluable in the clinical assessment of those with sleep disorders, to gauge severity of symptoms, fitness to work or drive, and assess response to treatment.

1.2 What is sleepiness, and how is it measured?

What is sleepiness?

> *"Sleep is an active behaviour, which is a reversible, repeating state of unconsciousness that can only be resisted for a limited period of time."* (Rogers, Dorrian, & Dinges, 2003)

As an abstract concept, the definition of *sleepiness* is elusive. It can be described by elaborating similes (see Table 1.1), or by the experiences or behaviours associated with it: a feeling of general lassitude, lagging attention, and loss of interest in the surroundings that one associates with the approach of sleep. Many words are used in connection with sleepiness, including "drowsiness", "hypersomnia", "somnolence", "sleep propensity", "fatigue", "tiredness", "ability to fall asleep", "sleepability", "ability to stay awake", "subjective sleepiness", "objective sleepiness", and "manifest sleepiness" (Johns, 1998).

It is commonly regarded as a transition state between wakefulness and sleep (Pivik, 1991). Another definition incorporates the effort expended to resist sleep (Torbjorn Akerstedt, 1998).

Macquarie Dictionary

Sleepy (adjective) 1. ready or inclined to sleep; drowsy. 2. of or showing drowsiness. 3. languid; languorous. 4. lethargic; sluggish. 5. quiet: a sleepy village. 6. inducing sleep. {SLEEP + -Y1} — *sleepily*, adverb — *sleepiness*, noun (Yallop, 2005).

Oxford English Dictionary The state of being sleepy; drowsiness; inclination to sleep; sluggishness, indolence (Simpson & Weiner, 1989).
Other definitions of sleepiness (A process involving) a succession of intermediate states, part wakefulness and part sleep, in varying proportions (Kleitman, 1963). Sleepiness is a curious construct (Bliwise, 2001). The concept of 'sleepiness' has traditionally referred to a subjective state of need for sleep (Torsvall & Akerstedt, 1988). Sleepiness is an attempt to turn the CNS over to sleep … it reflects an effort at resistance (Torbjorn Akerstedt, 1998).

Table 1.1 Examples of definitions of sleepiness

A working classification of sleepiness

It is clear from the numerous definitions above that sleepiness is not a unitary concept. Rather, there are several overlapping nuances, or constructs, of sleepiness. Some nuances, for example tiredness or fatigue (in terms of a lack of energy or a greater effort needed to undertake activities) are considered by many to be separate from sleepiness (Shen, Barbera, & Shapiro, 2006). A working classification of the definitions of sleepiness is presented below. Tests and questionnaires named here as examples are discussed later.

A transition state between wake and sleep (Kleitman, 1963). In this definition sleepiness is a gradual or graded process.

Lapses and wake state instability. Consistent the proposed existence of a flip-flop sleep switch controlling abrupt transitions between two discrete and stable states (Saper, Chou, &

Scammell, 2001), sleepiness has been described as a process of increasing instability where the state flips quickly between discrete states of wake and brief periods of sleep (microsleeps or lapses), rather than a gradual or continuously graded process (Doran, Van Dongen, & Dinges, 2001). An earlier theory, the lapse hypothesis (Bjerner, 1949; H. L. Williams, Lubin, & Goodnow, 1959), attributing impaired vigilance to momentary lapses between periods of normal performance, has been expanded to allow intermediate states between fully alert and microsleeps where errors of commission (false positives) can occur (Valley & Broughton, 1983). More recently the concept of wake state instability has been described, referring to the sleepy state as the increasing tendency for rapid and uncontrolled sleep initiation due to elevated drive for sleep, accompanied by efforts by the individual to resist, using increasingly greater compensatory effort to perform (Doran et al., 2001).

The propensity to fall asleep, the speed at which someone falls asleep (W. C. Dement & Carskadon, 1982). This type of sleepiness is operationalised by the Epworth Sleepiness Scale (Johns, 1991), and by the Multiple Sleep Latency Test (M. A. Carskadon & Dement, 1979).

Defined by the behaviours or sensory alterations observed with sleepiness: slowing of responses, reduced awareness of external environment, reduced responsiveness to external stimuli (G. O. Morris, Williams, & Lubin, 1960). These perceptions are described in the Stanford Sleepiness Scale (Hoddes, Zarcone, Smythe, Phillips, & Dement, 1973). Reaction time tests show reduced cognitive and psychomotor performance.

Defined by the states that bring about the sleepiness, the need for sleep or the drive for sleep, arising from lack of sleep, the physiological drive resulting from sleep deprivation (Aldrich, 1994). Models incorporating time spent awake and circadian influences have been used to explain sleep drive (Borbely, 1982). Related to this, experimental paradigms inducing sleep

loss or disruption from deprivation or fragmentation of sleep, or observational studies of disease states associated with abnormal sleep have also been used to study sleepiness.

Other concepts of sleepiness

The above classification is incomplete. There are a few other ways to divide the concept of sleepiness that should be noted in parallel to the working classification above. Firstly sleepiness can be described in terms of the time frame, as a **state or trait** component. The state component is a short-term property which refers to consequences of daily vigilance fluctuations or caused by atypical situations such as sleep deprivation. The trait component is steady and a constant aspect of each subject (Curcio, Casagrande, & Bertini, 2001). All 'objective' tests probably measure the sum of state and trait, while questionnaires can be worded to probe either component. Some nuances in the classification can apply to both trait and state sleepiness, for example the propensity to fall asleep (3), or the behaviours observed with sleepiness (4), while others such as the transition between wake and sleep (1) should be regarded as a state phenomenon. Trait sleepiness might be separated from state sleepiness by averaging repeated assessments over a longer time period, by administration of some questionnaires, or perhaps by as yet undiscovered genetic markers.

Secondly, **manifest sleepiness** has been distinguished from **latent sleepiness** (M. A. Carskadon & Dement, 1982). An underlying sleep tendency, the latent sleepiness, can be unmasked by behavioural and environmental factors such as a boring lecture, a long thesis, a heavy meal, or a monotonous car ride (Mary A. Carskadon & Dement, 1987).

Core or optional sleepiness. Analogous to the proposed presence of core sleep, (the part providing restitutive benefit, predominantly slow-wave sleep, occupying the first 3-4 sleep cycles) and optional sleep (non-restitutive drive behaviour maintaining sleep until waking-up time), **optional sleepiness** can be counteracted by sleep or by increasing the motivation to stay awake, whereas **core sleepiness** is that part of sleepiness that cannot be overcome by increasing motivation (Horne, 1991).

Sleepiness has also been categorised as **subjective and objective**. The former refers to symptomatic reports, while the latter to the results of laboratory tests of sleep latency or of neurocognitive performance testing. Sleepiness can be thought of as purely a symptom, or as an underlying biological construct that can be expressed as a symptom or a pattern of behaviours. The latter concept prompts the need to identify this biological construct, so that a 'true' measurement of sleepiness can be made.

There are other features related to sleep that are not covered by the classification scheme above. While some definitions refer to the transition from wake to sleep, an analogous state occurring upon the emergence from sleep, **sleep inertia**, has also been described (D. F. Dinges, 1989; Lubin, Hord, Tracy, & Johnson, 1976). Any sleepiness here should not arise due to sleep need, but due to persistence of some sleep processes after awakening. This phenomenon may be consistent with sleepiness as a transition state, but not as a need for sleep, nor an accumulation of sleep drive after prolonged periods awake. Furthermore, the classification has not clarified whether sleepiness reflects varying levels of a unipolar phenomenon, or whether the phenomenon is bipolar. Models incorporating competing effects of sleep and wake drive are discussed in the next section.

The following text briefly reviews concepts that are relevant to our aim to assess a potential physiologically-based measure of sleepiness. These concepts include the regulation of sleep, the mediation of sleepiness, and currently available methods of measuring sleepiness, by which the new measure will need to be compared. These areas have been thoroughly reviewed elsewhere (For example Cluydts, De Valck, Verstraeten, & Theys, 2002; Curcio et al., 2001; Shen et al., 2006).

Theoretical models of sleep regulation

Several models of sleep regulation have been proposed. The **two-process model** is the most well-known (Figure 1.1) (Borbely, 1982). In this model, sleep drive arises from the combination of a homeostatic process, process S and a circadian process, process C. Process S accumulates with increasing hours spent awake as a saturating exponential function, and is discharged upon sleeping by exponential decay. Process C varies by the time of the day as a sinusoidal function, and leads to maximum sleepiness close to the body temperature minimum, thus helping to concentrate the main sleep period during the night. The original form of the model conceived of two independent processes, with process S rising and falling between thresholds of sleep onset and waking from sleep, as determined by process C. The model was later revised to represent a single process arising from the additive interaction of process S and C (Achermann & Borbely, 1994). A later extension of this model, the **three-process model** added sleep inertia, process W (T. Akerstedt & Folkard, 1995; Borbely & Achermann, 1999). These models explain a tendency to fall asleep. Volitional factors or behaviours that might allow an individual to successfully resist sleep are not included. The models have formed the basis of many other variants, some of

them developed to allow prediction of the level of fatigue from various forms of sleep loss (Borbely & Achermann, 1999).

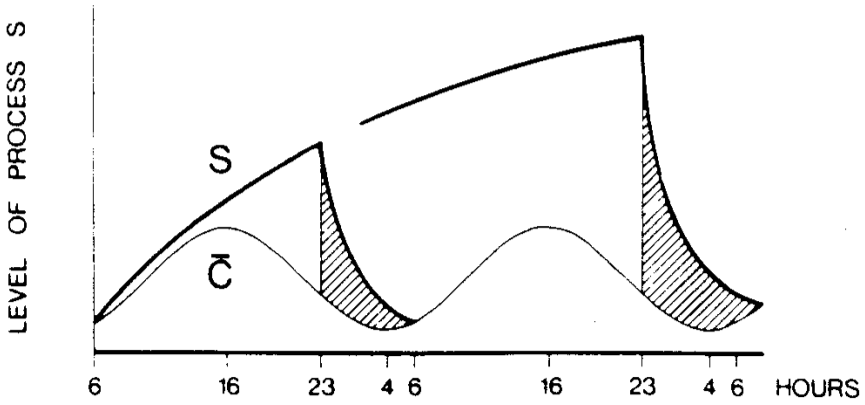


Figure 1.1 Two-process model of sleep regulation

Time course of the sleep processes after regular or extended wakefulness (right side of figure). The homeostatic process (S in the figure) accumulates with increasing time awake, and is discharged with sleep. The circadian process is drawn here as the negative of process C, Ĉ, so that the sleep drive is depicted as the vertical gap between the two curves (Borbely, 1982).

Others have proposed that sleep is not regulated by a monotonic sleep drive, but rather an interaction between two competing drives: a sleep drive and a **wake drive** (Edgar, Dement, & Fuller, 1993; R. B. Sangal, Thomas, & Mitler, 1992). In this model, factors such as the level of interest in the activities being undertaken or motivation towards a goal may influence the ability to fall asleep or stay awake in individuals at the same time of day who have been awake the same length of time. Notably, the original description of the two-process model also found it convenient to conceptualise Process C as its negative function (Figure 1.1), thus describing sleep propensity as the difference between the two opposing processes, although not explicitly naming a wake drive.

A **'four-process' model** (Johns, 1993, 1998)separates the sleep and wake drives into primary and secondary components (Figure 1.2). In this model, primary sleep drive is a process intrinsic to the central nervous system, responsible for the modulation of non-rapid eye movement (NREM) sleep. It may also incorporate circadian and ultradian rhythms. The secondary sleep drive increases progressively during wakefulness, and is analogous to the Process S from the two-process model. The primary wake drive arises from intrinsic reticular activating system activity, and corresponds to Process C from the two-process model. A fourth process, the secondary wake drive, refers to the effects of posture, behaviour, physical and mental activity, and sensory inputs. This latter process is partly under voluntary control.

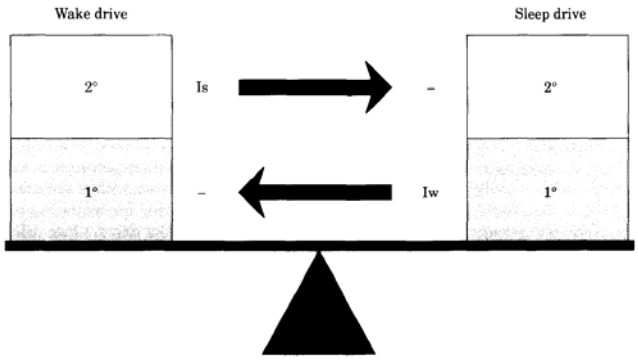


Figure 1.2 Interaction between primary and secondary components of the sleep and wake drives

A four-process model of sleep regulation (Johns, 1998): sleep and wake promoting neural systems (Saper et al., 2001) interact with extrinsic influences, with sleepiness resulting from an imbalance between the two (Shen et al., 2006).

The distinction between **state and trait factors** (Curcio et al., 2001) has been included in another four-process model, represented in Figure 1.3 below (De Valck & Cluydts, 2003). Trait components of wake and sleep drive determine the position of an individual within a two-dimensional space, and forms a starting point from where state wake and sleep drive can exert

their influence. While this model is similar to that of the previous with respect to sleep and wake drives, it differs in subdividing this into state and trait (stability over time), rather than intrinsic and extrinsic (whether influenced by factors external to the CNS).

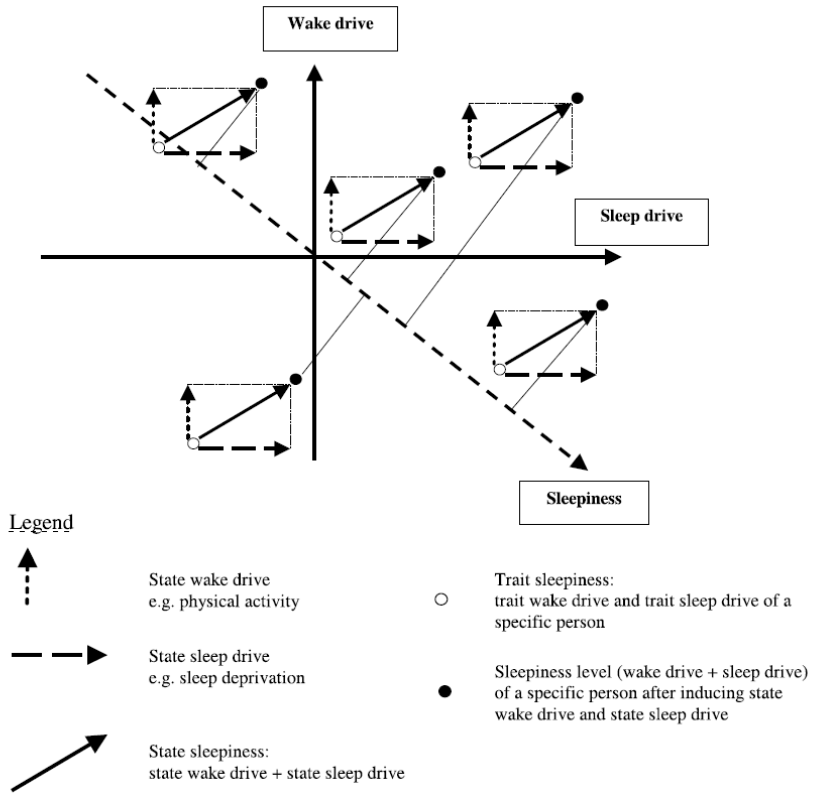


Figure 1.3 A working model for the two-dimensional state-trait conceptualization of sleepiness

Wake drive and sleep drive are depicted as orthogonal to each other. The main axes represent trait wake and sleep drive. Open circles represent an individual's trait level of wake and sleep drive. The addition of state wake drive (e.g. physical activity, behavioural factors, the afternoon dip) and state sleep drive (e.g. preceding time spent awake) gives the current state level (De Valck & Cluydts, 2003).

The above list of models of sleep regulation is hardly exhaustive, yet it brings to light important points. There may be multiple factors influencing sleepiness in any given situation. Also, given the variety of definitions of sleepiness and the likely presence of more than one process or drive governing sleep, it would be unreasonable to assume that a single measure of sleepiness will

encompass all forms of sleepiness. Underpinning the different types of sleepiness, there may be different cellular and chemical substrates governing wake or sleep drive, and the circadian or homeostatic systems. Key to understanding the construct validity of a measure of sleepiness, the following sections briefly review the purpose of sleepiness, and candidate biological substrates of sleepiness.

The purpose of sleep or sleepiness

The function or functions of sleep are unknown, but it has been proposed to be essential to tissue repair and growth, reversal of waking-induced changes in brain function, memory consolidation, energy conservation, and restorative processes in the brain and body (Rogers et al., 2003; Siegel, 2005). Many studies have supported the role of sleep in restorative processes. Drosophila flies carrying a gene mutation that results in one-third the usual sleep time did not demonstrate obvious performance impairment, however they had a reduced lifespan (Cirelli et al., 2005). Epidemiological data has associated reduced hours of sleep to cardiovascular disease (Ayas et al., 2003). It is clear that with prolonged time spent awake, it becomes more difficult and ultimately impossible, for the individual to resist the onset of sleep. During this period between full wakefulness and sleep onset, the range of behavioural and cognitive slowing, and perceptual change occurs that is understood as sleepiness. It is not known if sleepiness is a primary phenomenon, such as from the accumulation of a toxin that needs to be removed during sleep, or the depletion of a resource that needs to be replenished in sleep, or alternatively as a secondary marker, or timer, to remind the organism of the need to obtain sleep. It may be a means to

organise appropriate behaviours conducive to a routine sleep and wake cycle: preventing the organism from suddenly transitioning to sleep in inconvenient, or even dangerous, situations.

The biological substrate of sleepiness

The mechanisms for biological processes within an organism can be understood on various scales or levels. Broadly, they may be divided into the molecular level, the anatomical (brain nuclei) level, and the neural pathway level.

A theoretical **sleep factor**, or biological substrate for sleepiness, would promote sleep if administered, and show variations in levels reflecting sleeping and waking in the organism (Borbely & Tobler, 1989). If found, this sleep factor may form the basis for a physiologically-valid gold standard measure of sleepiness.

Molecular substrates of sleepiness

Adenosine is one of the most extensively studied candidate sleep factors (Porkka-Heiskanen, Alanko, Kalinchuk, & Stenberg, 2002; Porkka-Heiskanen et al., 1997). With prolonged wakefulness, extracellular adenosine concentrations rise, as measured by micro-dialysis in the basal forebrain of cat and rat. With recovery sleep after sleep deprivation, adenosine levels fall. Increasing adenosine levels in the basal forebrain induces sleep, an effect mainly mediated by the adenosine A1 receptor: perfusion of adenosine or an adenosine agonist induces sleep, while a selective A1 antagonist increases waking. It has been proposed that adenosine, triggered by energy depletion in the waking organism, is a means to signal the need to conserve energy. Interestingly, knockout mice with no A1 receptor retain homeostatic responses to sleep

deprivation (Stenberg et al., 2003). This result raises the possibility of other coexistent or subservient systems that regulate the homeostatic sleep drive, or the possibility that other adenosine receptor subtypes may also promote sleep.

Many other molecules have been studied as the sleep factor. Candidate molecules have included prostaglandin D_2, growth hormone, interleukin-1, brain-derived neurotrophic factor (BDNF) and amylase (Borbely & Tobler, 1989; R. Huber, Tononi, & Cirelli, 2007; Seugnet, Boero, Gottschalk, Duntley, & Shaw, 2006). For the circadian process, proteins such as *fos* accumulate intracellularly with a roughly 24-hour periodicity to provide cells with a clock functionality (Vitaterna, Pinto, & Turek, 2005; Zee & Manthena, 2007).

Brain nuclei and anatomical substrates of sleepiness

The role of hypothalamic structures in the control of sleep has recently been reviewed (Szymusiak, Gvilia, & McGinty, 2007). The waking system consists of the raphe nuclei and locus coeruleus of the brainstem (serotonergic and noradrenergic), tuberomamillary nuclei in the posterior hypothalamus (histaminergic), mesopontine laterodorsal and pedunculopontine tegmental (LDT/PPT) nuclei (cholinergic), and the basal forebrain (cholinergic). Orexinergic neurons within the posterior hypothalamus also regulate arousal via widespread projections, including to the basal forebrain. This distributed system is the putative anatomical substrate for the wake drive (Cluydts et al., 2002).

The ventrolateral preoptic area and median preoptic nucleus contain a population of neurons that are active during sleep. These nuclei project to the wake-promoting structures of the

tuberomamillary nuclei, the raphe nuclei, locus coeruleus, and receive reciprocal inhibitory projections from them. This supports the existence of mutually inhibitory sleep and wake drives.

The suprachiasmatic nucleus within the anterior hypothalamus is responsible for the control of a variety of circadian rhythms, including the system regulating wake drive (Miller, Morin, Schwartz, & Moore, 1996; Moore, 1983). It mediates arousal by indirect connections to the locus coeruleus (Aston-Jones, Chen, Zhu, & Oshinsky, 2001).

Neural populations and their interactions.

A **model of thalamocortical interactions** mediating the transitions between sleep and wake has been described (McCormick & Bal, 1997; Steriade, 2003). The model proposes that during NREM sleep, groups of thalamic and cortical neurons interact to generate synchronised oscillations. Hyperpolarisation of the thalamic reticular nucleus leads to hyperpolarisation and inhibition of the thalamocortical relay nuclei, thence blocking transmission of sensory information from the periphery. Hyperpolarised cell groups in the thalamus discharge action potentials in bursts. The 0.5 – 12 Hz oscillations are transmitted to the cortex, resulting in sleep spindles and delta waves on the surface EEG. Spindle rhythms appear at lesser degrees of hyperpolarisation, while delta rhythms arise with more progressive hyperpolarisation as NREM sleep deepens. A below 1 Hz slow oscillation arising in the cortex serves to group the delta and spindle waves, and also allow generation of EEG K-complexes (Amzica & Steriade, 1997; Steriade, Nunez, & Amzica, 1993). Electroencephalographic delta activity during sleep has been proposed as the correlate of the homeostatic sleep drive (Borbely, 1982), and theta activity may

be the waking correlate of this homeostatic process (Finelli, Baumann, Borbely, & Achermann, 2000).

During wake or in REM sleep, the neurons are depolarised, and show tonic spike activity. The depolarised state permits transmission of afferent information through to the cortex. In drowsiness and early sleep, thalamic reticular nucleus activity increases leading to hyperpolarisation (inhibition) of thalamocortical relay cells via inhibitory GABAergic projections (Steriade & Amzica, 1998). This increased gating of sensory information is consistent with experimental data showing that sleep-deprived subjects are less able to perceive and register stimuli (H. L. Williams, Gieseking, & Lubin, 1966).

Pathways from the mesopontine pedunculopontine (glutamatergic) and laterodorsal tegmental (cholinergic) nuclei inhibit the thalamic reticular nucleus, and cause depolarisation of thalamocortical relay neurons, hence promoting activation of the cortex during wake and REM sleep.

The above theory of thalamocortical interactions has been extended by others. A Synaptic Potentiation theory proposes that cortical activation as the organism interacts and learns from the environment induces **synaptic potentiation**, which increases over the course of a waking period (Giulio Tononi & Cirelli, 2003; G. Tononi & Cirelli, 2006). According to the theory, slow-wave activity during sleep is determined by the degree of synaptic potentiation at sleep onset, and the accumulation of the homeostatic drive is not only dependent on the time spent awake, but rather on the type of activities undertaken during wakefulness. This has been supported by experimental data (R. Huber et al., 2007). In this framework, sleep is necessary for energy conservation by downscaling of the synaptic potentiation. The changes in slow-wave activity

could be mediated by synaptic potentiation and synchronisation of cortico-cortical connections, which increase the cortical slow oscillation.

The **Neuronal Transition Probability Model** is another interesting model describing the behaviour of populations of neurons as a three-element radioactive decay process. Individual neurons exist in the mutually-exclusive states of depolarised, spindle activity, or delta activity as described in the thalamocortical model above, with stochastic transitions occurring between these states. The aggregate behaviour of the cell population predicted by the theory provides a good fit to the observed fluctuations in delta and sigma power across NREM sleep (Merica & Fortune, 1997).

A biological substrate may provide an accurate and objective method to measure sleepiness. The existence and identity of such a substance or structure is as yet unknown. Current potential candidates are difficult to measure in the living organism. On a larger scale, theories of interactions of populations of neurons appear to relate to changes observed within the surface electroencephalogram. The next section describes common currently-employed measures of sleepiness, with the exception of EEG-related measures, which will be discussed later (page 30).

How is sleepiness measured?

Measures of sleepiness in current use operationalise aspects of the definitions or nuances of sleepiness alluded to earlier. They are thus phenomenological and empiric – reflecting the consequences of sleepiness – rather than based on knowledge of the underlying neurophysiology of sleepiness. The following section describes a sample of the measures of sleepiness commonly

used in clinical and research situations. They have been mentioned here to describe the different approaches that have been undertaken to measure sleepiness, and also because, in the absence of a gold-standard measure, any new measure of sleepiness will need to be assessed against a variety of existing measures, with the expectation that these existing measures are approximations of the underlying concept of sleepiness. These measures have been reviewed elsewhere (For example Curcio et al., 2001; Dorrian, Rogers, & Dinges, 2005; Shen et al., 2006; Terri E. Weaver, 2001; T. E. Weaver, 2001).

The following list of measures has been grouped into direct and indirect symptomatic, performance-based, physiological, and measures of sleep or wake propensity.

Direct symptomatic measures

These are questionnaire instruments enquiring directly about sleepiness. These instruments can relate to either state or trait sleepiness (or possibly both). Questionnaires can be useful in probing experiences over a longer time period, however as they are based on self-report they rely on an individual's insight into their own state, and can be prone to bias.

Epworth Sleepiness Scale (ESS) (Johns, 1991). This is a measure of sleep propensity and trait sleepiness. Subjects are asked to indicate their probability of dozing given a set of circumstances. This is an 8-item, 4-response scale which yields a total score from zero to 24, with higher values indicating increased sleepiness. Values over 10-11 indicate abnormal sleepiness (Johns, 2000). Factor analyses and item response theory analyses have shown the ESS to measure a single construct (Johns, 1992, 1994; Violani et al., 2003). It is used to measure sleepiness over a longer time period of days or weeks, rather than minutes or hours. The ESS can distinguish between

snorers and OSA patients of varying severity (Gottlieb et al., 1999; Johns, 1993), as well as distinguish healthy workers from those with narcolepsy or idiopathic hypersomnia (Johns & Hocking, 1997). The ESS and sleep latency on the MSLT (see later) show moderate correlations (Spearman's r -0.42) (Johns, 1994). It is the most frequently cited questionnaire of sleepiness (Chervin, 2003; Miletin & Hanly, 2003), but some criticisms of the instrument have included the presence of gaps in the severity of sleepiness that it is able to measure with the items employed at intermediate levels (Violani et al., 2003) or at the highest levels of severity (R. Bart Sangal, Mitler, & Sangal, 1999), the lack of data on the test-retest reliability of the ESS in disease groups such as OSA (Miletin & Hanly, 2003), and the failure to reproduce a relationship between the ESS and the Multiple Sleep Latency Test (MSLT), a test which operationalises sleep propensity, or between the ESS and measures of sleep apnea severity in patients with OSA, and the failure to robustly track improvements in the MSLT or Maintenance of Wakefulness Test with therapy (Chervin, 2003; Chervin & Aldrich, 1999).

Stanford Sleepiness Scale (SSS) (Hoddes et al., 1973). This is a single item, 7 response scale that can be applied repeatedly within a single day to measure changes in sleepiness. It is a measure of state sleepiness, in relation to the behaviours that accompany sleepiness. The SSS has been shown to be sensitive to sleepiness related to sleep deprivation (Hoddes et al., 1973) and narcolepsy (Aguirre & Broughton, 1987). It may not be uni-dimensional: factor analysis of the item descriptors has indicated three dimensions, "alertness/sleepiness", "loss of control" and a "cognitive factor" (MacLean, Criollo, Fekken, & Knowles, 1990; MacLean, Saskin, Fekken, & Knowles, 1989).

Karolinska Sleepiness Scale (KSS). This is a single item, nine-response scale in which respondents are asked to indicate their state over the *preceding 10 minutes* (T. Akerstedt &

Gillberg, 1990; Gillberg, Kecklund, & Akerstedt, 1994). Only 5 of the items contain verbal labels. This is another measure of state sleepiness, and is similar to the Stanford Sleepiness Scale. It allows repeated intra-day measurement. The KSS has been validated against neurocognitive performance and EEG measures of sleepiness (Kaida et al., 2006).

Visual analogue scales and others. Generic scales can be used to indicate the degree of any phenomenon, including sleepiness. Examples include a Visual Analogue Scale for continuous rating of the degree of sleepiness (Casagrande, Violani, Curcio, & Bertini, 1997; Folstein & Luria, 1973; Monk, 1989), or an interval scale expressing the strength of sleepiness (Borg, 1990; van den Berg, Neely, Nilsson, Knutsson, & Landstrom, 2005). Care must be taken in choosing the wording of the item descriptions.

Profile of Mood States (POMS) (McNair, Lorr, & Droppleman, 1971). This questionnaire includes six mood factors: tension-anxiety, depression-dejection, anger-hostility, vigour, fatigue and confusion. A series of 75 adjectives are rated on a 5-point scale. Fatigue, vigour, and to a lesser extent, confusion, subscales are sensitive to sleep loss (D. F. Dinges et al., 1997).

Activation-Deactivation Adjective Check List (ADACL) (Thayer, 1967)**.** Each of a series of adjectives is rated on a 4-point scale.

Sleep-wake Activity Inventory (SWAI), a 59-item questionnaire that contains one sub-scale that measures sleepiness based on frequency of dozing (Rosenthal, Roehrs, & Roth, 1993). Like the ESS, it also measures sleep propensity, but it is not widely used, probably because of its length.

Hyperarousal scale (Pavlova et al., 2001; Regestein, Dambrosia, Hallett, Murawski, & Paine, 1993). This scale has been proposed as a measure of wake drive (De Valck & Cluydts, 2003), but

it has not yet been specifically studied in regard to this. It was developed for insomnia research, and contains 26 items, each with 4 Likert-type responses. Some items show content validity with the wake drive, including "I have trouble falling asleep", "I cannot take naps even if I try", however there are others relevant to insomnia, that relate to mood, personality and coping strategies (e.g. "I get tearful easily", "I take a long time to make decisions", "I keep thinking about things long after they happened."). This is probably a trait measure.

Indirect symptomatic measures

Generic or disease-specific quality of life questionnaires may also measure the effect of sleepiness on daytime function or overall health status. The two examples mentioned here reflect trait sleepiness.

Functional Outcomes of Sleep Questionnaire (Weaver et al., 1997). This 30-item questionnaire explores the extent to which sleepiness or sleep disruption affects five aspects of daily activities: general productivity, social outcome, activity level, vigilance, and intimate relationships and sexual activity. This questionnaire is described further in page 98.

Medical Outcomes Study SF-36 (Ware & Sherbourne, 1992). Subscales of this generic quality of life instrument have been shown to be sensitive to treatment of disorders causing sleepiness such as sleep apnea (Jenkinson, Davies, Mullins, & Stradling, 1999).

Performance measures

A variety of performance tasks are sensitive to the effects of interventions designed to increase sleepiness such as sleep loss (D. F. Dinges & Kribbs, 1991). A convenient classification of the performance tests used to evaluate sleepiness divides the tests broadly into psychomotor or cognitive tasks (Curcio et al., 2001). Psychomotor tasks are tasks of reaction time, tracking or tapping, while the cognitive tasks measure more complex functions including attention, memory and executive function. Executive function can be further divided into nine components (Harrison & Horne, 2000):

- Appreciating a complex situation while avoiding distractions.
- Keeping track of events, and developing and updating strategies (working memory).
- Thinking laterally and being innovative.
- Assessing risk, anticipating a range of consequences.
- Maintaining interest in an outcome.
- Controlling mood and uninhibited behaviour.
- Showing insight into one's own performance.
- Remembering "when" rather than "what", temporal memory.
- Communicating effectively, language processing.

Examples of performance tasks are described in pages 87-92. There are several issues to note in using these tests. Firstly, the tests may measure overlapping components of performance, and there are also other schemes to classify the domains of cognitive and executive function (For

example Beebe & Gozal, 2002). Related to this, different metrics from a single test may provide different types of information. At their simplest, reaction time tests measure motor speed and hence response slowing, a consequence of sleep deprivation. Longer test sessions tap the subject's ability to sustain attention, and increase the ability to detect lapses when response times to individual stimuli are markedly prolonged, or false positives when the subject triggers a response inappropriately when no stimulus is presented (Doran et al., 2001). Secondly, the tests measure the condition the subject is in at the time of testing: a combination of state and trait influences (Johns, 1998).

Thirdly, there are many other factors that also influence performance on these tests, apart from the level of sleepiness. Optimal performance in complex tasks also relies upon intact function in more basic tasks such as sensory perception, motor response, coordination, and vigilance (Edwin Verstraeten & Cluydts, 2004). Performance on the tests are also affected by motivation (Horne & Pettitt, 1985), situational or environmental influences, and variations in test administration (Curcio et al., 2001). Increasing the duration of a monotonous test increases its sensitivity to performance decrement from sleep loss (R. T. Wilkinson, 1961), but there may be increased boredom and reduced motivation. Conversely, a stimulating or engaging task increases motivation. While increased motivation may allow the test subject to compensate for the effects of some sleep loss (Bonnet & Arand, 2005), a task with very high cognitive demand may also have a greater ability to detect performance decrements from sleep loss (A. S. Smit, P. A. T. M. Eling, & A. M. L. Coenen, 2004). Such a cognitively-demanding task may also heighten accumulation of homeostatic sleep drive (R. Huber et al., 2007; A. S. Smit, P. A. Eling, & A. M. Coenen, 2004). Thus, it is difficult to predict the effects of different task durations and task demands.

Fourthly, the test subject will improve performance on a test with repeated administrations. Practice effect can be balanced using a cross-over experimental design, or minimised by performing practice tests until performance reaches a plateau (H. P. A. Van Dongen & Dinges, 2000). Some complicated tasks, such as puzzle solving tests, cannot be repeated as they rely on a novel presentation, because a subject remembering the answer might spuriously show improved performance. Different versions of such tests will need to be developed and validated.

Physiological responses

Tests have been developed to measure a variety of physiological responses that usually accompany the onset of sleep and the period preceding it. Electroencephalographic changes are discussed later, in page 30. The period between wake and sleep is marked by changes in autonomic tone, manifested as alterations in pupil diameter, cardiovascular measures, skin conductance, and body temperature (Pivik, 1991). One of the most extensively studied measures of autonomic response, pupillometry, shows significant correlation with MSLT (Spearman's r -0.57) (McLaren, Hauri, Lin, & Harris, 2002). Although a relatively non-intrusive technique, the need for equipment to detect pupil size such as an infrared camera, and the need to conduct the test in a darkened room pose difficulties for field measurement.

Eyeball and eyelid movements have also been studied as markers of the drowsy state. It has long been appreciated that slow eye movements appear during drowsiness (Aserinsky & Kleitman, 1955). Slow eye movements detected on the electrooculogram (EOG) have been shown to increase during sleep deprivation (Cajochen, Khalsa, Wyatt, Czeisler, & Dijk, 1999), and have been related to performance decrements at work (Torsvall & Akerstedt, 1987) and on prolonged

tasks performed at night (Torsvall & Akerstedt, 1988), and reduced sleep latency on the Multiple Sleep Latency Test after sleep restriction (De Gennaro, Devoto, Lucidi, & Violani, 2005). The latter two studies also showed that EOG-recorded blinking frequency is also reduced with increasing sleepiness (De Gennaro, Devoto et al., 2005; Torsvall & Akerstedt, 1988). The degree of eyelid closure increases with increasing sleepiness, as detected by camera (D F Dinges, Mallis, Maislin, & Powell, 1998; Johns, 2003; Wierwille, 1999) or inferred from the electrooculogram (T. Akerstedt, Peters, Anund, & Kecklund, 2005; Papadelis et al., 2007). In the latter real-world driving study, increased eyelid closure preceded serious errors (Papadelis et al., 2007).

Sleep propensity and wake propensity

Laboratory tests of sleep latency have also been proposed as an objective test of sleepiness. The **Multiple Sleep Latency Test (MSLT)** (M. A. Carskadon et al., 1986; W. C. Dement & Carskadon, 1982) measures sleep latency in subjects asked to sleep in a quiet, dark room for 4-5 periods across a morning and afternoon. It predicts an individual's vulnerability to fall asleep in a low-stimulus environment (Mary A. Carskadon & Dement, 1987). Average sleep latency values less than 5 minutes are generally regarded as severe or pathological, while values above 10 minutes are considered to be normal (Thorpy, 1992). These thresholds were defined arbitrarily, rather than empirically. As a measure of sleep propensity, it differs from ESS in that the latter presents a range of situations, and asks the subject to rate their probability of dozing. Although the MSLT has high test-retest reliability (Pearson's r = 0.97 over 4-14 months), it is time-consuming and costly to administer, and hence relatively impractical to be used for frequent

testing (Zwyghuizen-Doorenbos, Roehrs, Schaefer, & Roth, 1988). Reducing the number of nap periods reduces reliability of the test (Zwyghuizen-Doorenbos et al., 1988). There is also a problem with distinguishing amongst severe levels of sleepiness, in which case a test such as the Maintenance of Wakefulness test (below) might show better utility (Sugerman & Walsh, 1989).

The **Maintenance of Wakefulness Test (MWT)** (Mitler, Gujavarty, & Browman, 1982) requires subjects to resist sleeping over several twenty or forty-minute periods across a day. A lower limit of normal of 10.9 minutes was calculated from a mean value of 18.1 minutes and standard deviation 3.6 minutes if a 40-minute test was truncated at 20 minutes (Doghramji et al., 1997). It should be noted that many healthy individuals had 20-minute latencies (in the study, the mean sleep latency for the 40-minute test was 35.2 minutes), and hence the distribution of the test scores is non-Gaussian – this ceiling effect makes the calculation of normal ranges based on the normal distribution problematic. Overall, normative data is limited on this test (S. Banks, Barnes et al., 2004; Mitler, Doghramji, & Shapiro, 2000). There is also disagreement on what criteria determine the onset of sleep on this test: for example the onset of microsleeps, or the need for 3 consecutive epochs of stage 1 sleep.

The lack of a strong correlation (Pearson's $r = 0.41$, and a 30% discordance rate) between MSLT and MWT performed on the same day and a factor analysis showing two factors, "sleepiness" and "alertness", accounting for 91% of variance supports the notion that the processes initiating sleep are distinct from the processes responsible for maintaining alertness (R. B. Sangal et al., 1992). This discrepancy also lent support to the presence of optional sleep within the theory of core and optional sleep (Horne, 1991).

The **Oxford Sleep Resistance (OSLER) Test** was designed as a simple alternative to the MWT (Bennett, Stradling, & Davies, 1997). Subjects are asked to press a button in response to a red

light which flashes approximately every 3 seconds across several 40-minute periods. Lapses on the OSLER have been related to EEG-scored microsleeps (Priest, Brichard, Aubert, Liistro, & Rodenstein, 2001), and also sleep latency on the MWT (Krieger, Ayappa, Norman, Rapoport, & Walsleben, 2004). Though presenting some practical advantages, the test has not been widely adopted. There are ceiling effects with the OSLER and MWT if many individuals do not lapse or fall asleep within the duration of the testing period (Bliwise, 2001; Doghramji et al., 1997).

Often regarded as the gold standard tests for sleepiness, the MSLT and MWT have been suggested in relation to the evaluation of driver safety in guidelines put forward by regulatory authorities (Austroads & National Road Transport Commission, 2003), even though there is only data relating the MWT to simulated, but not real-world driving performance (Sagaspe et al., 2007). The tests are also cumbersome and expensive to administer, with the requirement for several nap or test periods to reduce measurement error, and the propensity to be affected by prior activity (Bonnet & Arand, 1998), and motivational factors.

Why would we need to measure sleepiness?

The significant and catastrophic consequences of falling asleep while driving and working are well recognised (T. Akerstedt, 1995; W. C. Dement & Mitler, 1993; D.F. Dinges, 1995), as are the presence of more subtle but potentially functionally important impairments in reaction times and cognition with increasing sleepiness (Williamson & Feyer, 2000). Sleepiness is a common symptom experienced at some times by all individuals, however there would be utility in being able to detect levels of sleepiness that might put the individual or the wider community at risk (Bliwise, 2001).

A measurement tool for sleepiness might be useful to assess an individual's safety to work or drive, to identify those with sleep disorders who might need treatment, and to monitor the effect of treating sleepiness. As a research tool, it would have applications in the development and assessment of countermeasures for sleepiness, and perhaps provide a better understanding of the underlying mechanisms controlling sleep and wake.

Problems with current measures of sleepiness

There is no direct measure of sleepiness: the theorised "sleep factor" has yet to be discovered. Problems with the use of the currently available tests have been discussed earlier. Currently employed measures relate to different nuances of sleepiness, and thus correlations between them are mostly moderate at best. There can be no single gold standard measure: multiple dimensions of sleepiness exist, and the choice of measure also depends on the purpose. There is also the issue of generalisability to other situations and times, which offer different situational sleep propensities (Johns, 1998).

Different measures of sleepiness do not correlate very well with each other. This could be due to their relation to differing definitions of sleepiness, but also differences in the way the measures are quantified and scored, and measurement error (Pivik, 1991). Similar physiological tests like the MSLT and MWT show poor correspondence, perhaps due to the different instructions given to subjects (R. B. Sangal et al., 1992). Questionnaire instruments often correlate with other questionnaires, but studies relating questionnaires to other measures of sleepiness have shown inconsistent results. The Stanford Sleepiness Scale showed no relation to pupillometry during normal wakefulness (Lavie, 1979). While the initial validation studies for the Epworth

Sleepiness Scale showed moderate correlations with the MSLT, others have not replicated this (Chervin & Aldrich, 1999).

Questionnaire scales may contain more than one domain or construct, and even more simple performance tasks may provide several metrics that are differentially sensitive to sleepiness.

Many of the tests have been evaluated in total sleep deprivation experiments. These protocols provide a large variation in the level of sleepiness, and allow for within-subjects analyses, hence maximising the power to detect a relationship between the candidate measure and other markers of sleepiness. The presence of significant inter-individual variability (H. P. Van Dongen, Baynard, Maislin, & Dinges, 2004; R. T. Wilkinson, 1961) in the extent to which subjects manifest performance impairment from sleep loss leads to a further loss of power in between-subject comparisons, particularly when the variability in sleepiness in non-sleep deprived individuals may be lower. There may also be inter-individual variability in which individuals at the same level of sleepiness may express impairment in different tests (Leproult et al., 2003).

Thus, in order to validate a new measure of sleepiness, it is not sufficient to compare it with a single measure of sleepiness, for example the MSLT. Construct validity needs to be demonstrated across range of measures, and ideally there should be consistency of findings across a variety of types of tests. Similarly, the finding that a test relates poorly to the MSLT will not necessarily imply that the test is a poor measure of sleepiness.

1.3 Sleepiness and the electroencephalogram

EEG as a measure of sleepiness

Changes on the electroencephalogram (EEG) accompanying the onset of sleep were reported within a decade of the discovery of the EEG (Davis, Davis, Loomis, Harvey, & Hobart, 1938; Loomis, Harvey, & Hobart, 1937). This physiological measure of brain electrical activity remains the primary method used to distinguish wake from sleep, and to stage sleep (W. Dement & Kleitman, 1957; Rechtschaffen & Kales, 1968). This staging method has been recently revised, but still relies on visual recognition of patterns of EEG waveforms (Silber et al., 2007). Some laboratory tests of sleepiness such as the Multiple Sleep Latency Test (M. A. Carskadon et al., 1986) or the Maintenance of Wakefulness Test (Mitler et al., 1982) use EEG recordings, but score time to the discrete onset of sleep, rather than quantifying degrees of transition to sleep.

Visual changes with the EEG in the period between wake and stage 2 sleep have been described in detail, including the dropping-out of the alpha rhythm, EEG slowing, and reduction in amplitude (Santamaria & Chiappa, 1987). Analogous to the staging within sleep, they have been classified into two (Davis et al., 1938), four (Foulkes & Vogel, 1965), or even nine stages (Tanaka, Hayashi, & Hori, 1996). The presence of characteristic patterns on the EEG with as the individual approaches the onset of sleep supports the face validity of the EEG as a measure of sleepiness.

The widespread availability of equipment to measure a limited number of EEG leads within sleep laboratories, and the ability to make continuous measurements while subjects are undertaking other tasks without interruption (Sallinen et al., 2004) make this a practical and accessible measure. Aside from visual scoring, the EEG can be quantified, with the most

common method being power spectral analysis, which employs a Fourier transform to express the power of the EEG waveform as a function of frequency, thus making the EEG measure amenable to numerical analysis. Other EEG related measures have also been studied in relation to sleepiness, including alternatives to, or variants of, the Fourier Transform such as period analysis (Daniel, 1967), period-amplitude analysis (Feinberg, Maloney, & Campbell, 2000; Uchida, Feinberg, March, Atsumi, & Maloney, 1999), time-frequency decomposition (Chervin et al., 2004), or measures of entropy (Papadelis et al., 2007). Some of these techniques have advantages in providing a greater time resolution, or the ability to provide separate information on the amplitude and frequency of the waveforms. Other EEG-related procedures have also been studied, such as event-related potentials in wake (Broughton, 1982), and in sleep (Bastuji & Garcia-Larrea, 1999).

Apart from face validity and practicality, several other important factors, discussed below, support the EEG as a potential measure of sleepiness:

- EEG power shows circadian variation during wake and sleep.
- EEG power varies with increasing time awake: with waking recordings, or during naps or recovery sleep.
- Wake EEG relates to other measures of sleepiness: symptoms, performance measures, physiological responses, sleep or wake propensity.
- Sleep EEG relates to sleepiness measures on the following day.

A table summarising the literature concerning EEG power measures as markers of sleepiness is provided on page 36.

EEG power shows circadian variation during wake and sleep

The wake EEG is known to fluctuate with time of day (Gundel & Witthoft, 1983; Thompson & Harding, 1968). Circadian variations in theta and high alpha (10.25 – 13.0 Hz) power have been shown during sleep deprivation, using mathematical modelling to separate homeostatic from circadian effects (Aeschbach et al., 1999). The circadian fluctuation in EEG activity was also demonstrated in the absence of sleep deprivation, with the change in EEG correlating with performance (Lafrance, Paquet, & Dumont, 2002).

The EEG within NREM sleep also shows circadian modulation. Using a 28-hour day to separate homeostatic from circadian effects, circadian modulation of NREM EEG power was found in the 13.75-15.5 Hz and 12.25 – 13.0 Hz bands, with the rhythms for the two frequency bands being out of phase with each other (D.-J. Dijk, Shanahan, Duffy, Ronda, & Czeisler, 1997). An experiment using a 90-minute day protocol also demonstrated circadian variation in delta, sigma, and beta power for NREM sleep EEG (Niggemyer, Begley, Monk, & Buysse, 2004).

EEG power shows homeostatic modulation

Sleep deprivation experiments examining waking eyes-open EEG have found theta (6.5 – 9 Hz) (Cajochen, Brunner, Krauchi, Graw, & Wirz-Justice, 1995; Taillard et al., 2006), frontal delta-theta (1-7 Hz) (Cajochen, Knoblauch, Krauchi, Renz, & Wirz-Justice, 2001), or alpha, theta and

beta (Aeschbach et al., 1999) to increase with increasing time awake. The frequency range to most consistently show this response during wake is the theta band.

Delta power during NREM sleep is thought to be an important marker of the homeostatic regulation of sleep (Borbely, 1982; Feinberg, 1974). It increases in the recovery sleep after sleep deprivation (Borbely, Baumann, Brandeis, Strauch, & Lehmann, 1981), and sleep delta and theta power increase with increasing time spent awake (D. J. Dijk, Beersma, & Daan, 1987). This effect was also shown with mathematical modelling of sleep data from a 90-minute day protocol (Niggemyer et al., 2004).

Wake EEG relates to symptomatically-determined sleepiness, within and between subjects.

There is extensive evidence from the literature relating symptomatic sleepiness to EEG power, particularly within the alpha and theta frequency bands. The majority of these studies have examined changes within subjects during a prolonged and monotonous task, as the degree of sleepiness changes with sleep deprivation, or due to the circadian fluctuation in sleepiness.

The effect of sleepiness on alpha power depends on whether the eyes are open or closed. Alpha power in the eyes open condition increases with greater rated sleepiness (T. Akerstedt & Gillberg, 1990; T. Akerstedt, Kecklund, & Knutsson, 1991; Cajochen et al., 1995; Torsvall & Akerstedt, 1987), while in the eyes closed condition, alpha power decreases with greater rated sleepiness (Aeschbach et al., 1999; Leproult et al., 2003; Strijkstra, Beersma, Drayer, Halbesma, & Daan, 2003).

Theta power increases with increased reported sleepiness (T. Akerstedt & Gillberg, 1990; Cajochen et al., 1995; Cajochen, Khalsa et al., 1999; Cajochen et al., 2001; Lal & Craig, 2002; Torsvall & Akerstedt, 1987). Delta power also increases with increased sleepiness (Cajochen, Khalsa et al., 1999; Cajochen et al., 2001; Lal & Craig, 2002; Leproult et al., 2003).

Wake EEG relates to test-related sleepiness: neurocognitive performance and the MSLT

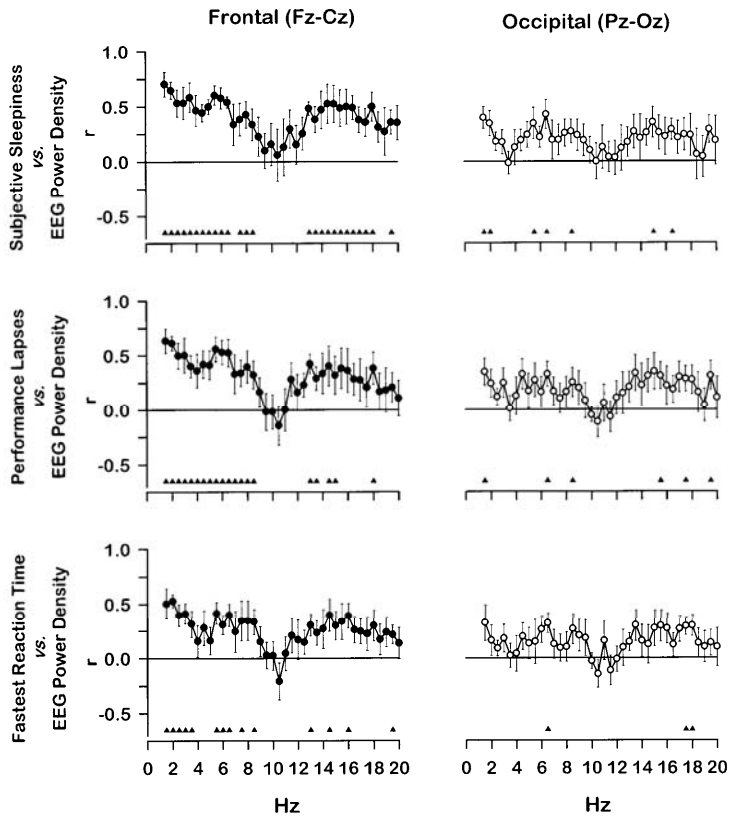


Fig. 4. Pearson product-moment correlations between EEG power density during wakefulness in each frequency bin between 1 and 20 Hz (Fz-Cz and Pz-Oz) and KSS, performance lapses (reaction times > 500 ms), and 10% fastest reaction times (n = 9, means ± SE). ▲, significant r values ($P < 0.05$, t-test).

Figure 1.4 Relationships between symptomatic sleepiness, PVT performance, and EEG power by frequency bin (Cajochen, Khalsa et al., 1999)

Figure 1.4 shows results from a sleep deprivation experiment. Ten healthy subjects were assessed every 2 hours during 32 hours sustained wakefulness. Pearson correlation coefficients (*r*) between wake eyes-open EEG power density in 1-Hz frequency bins and three sleepiness-related outcomes (Karolinska Sleepiness Scale rated sleepiness, Psychomotor Vigilance Task Lapses and mean fastest 10% reaction times) are shown on the vertical axes of the plots, and frequency bins for EEG power are shown on the horizontal axes (Cajochen, Khalsa et al., 1999). It can be seen that the relationships between EEG and sleepiness related measures extends over a wide range of frequencies. It also appears that frontal (Fz-Cz, left-sided graphs) EEG power tracked sleepiness more closely than Occipital (Pz-Oz, right-sided graphs) power.

EEG power has also been related to EOG measures as measures of sleepiness. Moderate to strong correlations were noted between the two measurements across several levels of drowsiness between alert to stage 1 sleep. The study employed analysis methods intended to minimise cross-contamination of signals in the two types of recordings (Hyoki, Shigeta, Tsuno, Kawamuro, & Kinoshita, 1998).

Sleep EEG relates to sleepiness measures on the following day

In cross-sectional studies, sleep beta power was associated with higher reported sleepiness according to the ESS (Wichniak et al., 2003), while lower sleep theta power was associated with greater sleep propensity on the MSLT (Wichniak et al., 2003). Low frequency delta (0.5-.10 Hz) activity in the first cycle of NREM sleep correlated with tests of prefrontal cortical function such as the Wisconsin Card Sorting Test and the Tower of London (Anderson & Horne, 2003).

These results could be interpreted as suggesting that the EEG marks the process that restores sleep, marks the proportion of undissipated sleepiness, or indicates the degree of disruption to sleep.

Summary table of EEG literature

In summary, the EEG has been extensively studied as a measure of sleepiness. The wake and sleep EEG is affected by processes known to alter levels of sleepiness such as sleep deprivation, and the time of day. It has also been correlated to other measures of sleepiness including self-report questionnaires, measures of cognitive slowing including reaction time, vigilance, memory and executive function, as well as the sleep latency on the MSLT or MWT.

A summary of the EEG literature pertaining to sleepiness is included in Table 1.2 below. The literature is grouped by frequency bands, and then by recording condition (wake eyes open (EO), wake eyes closed (EC), or sleep). Literature concerning the EEG in obstructive sleep apnea has also been included here for completeness, although it will be discussed in a later section.

Condition	Frequency Band	Findings
Wake EO	Delta	After reducing sleep by 2 hours, EO delta and theta rose with increasing fatigue during a 2-hour driving simulator task (Lal & Craig, 2002).
Wake EC	Delta	EC delta increased (and EC alpha decreased) with 27 hours sleep deprivation (Leproult et al., 2003). EEG slowing (delta-theta divided by alpha-beta) occurred in frontal leads in wake and REM in OSA as opposed to controls (Morisson et al., 1998). Waking EC

Condition	Frequency Band	Findings
		absolute delta power also increased in OSA. Delta-theta fell after CPAP in EC wake and in REM.
Wake EO	Theta	Power in the 6.25-9.0 Hz band increased during 40 hours of sleep deprivation resembling saturating exponential function. Power correlated with VAS-scored fatigue (Cajochen et al., 1995). Eyes open delta and theta during a 40 hour sustained wakefulness constant routine protocol shows a steady increase with a trough in the evening centred 6 hours before temperature trough. Theta was modulated by process S, but also showed circadian modulation (Aeschbach et al., 1997). Continuously recorded eyes open theta (5.25 – 6.0 Hz) showed cosine and saturating exponential effects during 40 hours of sustained wakefulness (Aeschbach et al., 1999). Theta trough coincided in the trough in sleep propensity (amount of sleep in a 10 minute nap period) (Aeschbach et al., 1999) i.e., low EO theta coincided with high alertness. Increasing EO theta activity during extended waking (40 hours) correlated with increased SWA in the first NREM sleep cycle. Theta increased in a linear pattern during wake. This study suggested wake theta is linked functionally to sleep SWA (Finelli et al., 2000). Theta during a performance test increased after 28 hours sleep deprivation (Smulders, Kenemans, Jonkman, & Kok, 1997). An early study on the ambulatory EEG to measure sleepiness showed increased alpha and theta power in the eyes open state with increased sleepiness (T. Akerstedt & Gillberg, 1990). During a vigilance task, EEG theta activity and reaction time increased while blood flow in the medial thalamus as well as in several cortical regions decreased (Paus et al., 1997). EO theta increased within 60 minutes of melatonin given at 1300 or 1800 (Cajochen et al., 1996). Caffeine reduces theta and alpha activity during wake

Condition	Frequency Band	Findings
		(Dimpfel, Schober, & Spuler, 1993).
Wake EC	Theta	Performance on reaction time tasks worsens with increased EC theta in amphetamine-dependent subjects (Newton et al., 2004). Theta power (and also EC beta-2) increases after mental effort. Performance on a clock test worsened with higher theta power (A. S. Smit et al., 2004). Note in this paper some metrics behaved in the opposite direction: errors vs. RT.
Wake EO	Alpha	A study with ambulatory EEG to measure sleepiness showed increased alpha and theta power in the eyes open state with increased sleepiness (T. Akerstedt & Gillberg, 1990). Waking hourly mean alpha power density (but not theta) increased across a night shift and correlated with subjective sleepiness measured by the SSS (r = 0.24) (T. Akerstedt et al., 1991). Predicted alertness is low with higher alpha density (T. Akerstedt & Folkard, 1995). Alpha power increased across night shift in train drivers (Torsvall & Akerstedt, 1987). Subjective sleepiness correlated with (presumed eyes open) EEG alpha burst activity in truck drivers during evening and night driving (Kecklund & Akerstedt, 1993). EO Alpha during a seven-lap 2-4 hour simulated drive increased with each lap, and it correlated with some measures of driving performance: the index of running of the road on the left, particularly in the 'lighted motorway' condition. The latter effect may only have been demonstrated because of 4 outliers (Campagne, Pebayle, & Muzet, 2004). Eyes open alpha showed circadian (but not homeostatic) modulation across a 40 hour sustained wakefulness constant routine protocol (Aeschbach et al., 1997). The alpha attenuation coefficient (mean EC alpha divided by mean EO alpha) fell (i.e. EC alpha falls or EO alpha rises, or both) in the early hours of the morning, concomitant with fall in MSLT (more sleepy) during 40

Condition	Frequency Band	Findings
		hour sleep deprivation (Stampi, Stone, & Michimori, 1995). Thus a higher AAC implies higher alertness. In OSA subjects, AAC fell with increasing MWT (contrary to expectations) (Sforza, Grandin, Jouny, Rochat, & Ibanez, 2002). EO alpha divided by EC alpha during a driving simulator task was lower after a nap in sleep deprived truck drivers (Macchi, Boulos, Ranney, Simmons, & Campbell, 2002). This metric is analogous to the inverse of the AAC, and is consistent with predicted higher alertness after a nap.
Wake EC	Alpha	Beta and alpha (9.77-25.15 Hz) was increased after sleep deprivation (Corsi-Cabrera et al., 1992). EC alpha:theta ratio reduced after 2 nights of auditory-induced sleep fragmentation (Cote, Milner, Osip, Ray, & Baxter, 2003). Eyes closed waking alpha power is high with low subjective sleepiness (higher alertness) (Strijkstra et al., 2003). Circadian variation of high frequency eyes closed alpha activity related to subjective alertness during 40 hours sleep deprivation (Aeschbach et al., 1999). EC delta increased and EC alpha decreased with 27 hours sleep deprivation. There was a negative relationship between Global Vigour and EC high alpha (10.5-12.5 Hz) EEG power. There was no relationship between alpha and reaction times (Leproult et al., 2003). Modafinil abolished and reversed the steady fall in eyes closed awake alpha power over prolonged (60 hours) sleep deprivation in healthy volunteers (Chapotot, Pigeau, Canini, Bourdon, & Buguet, 2003). Amphetamine caused a more prominent fall in alpha. Delta not affected.
Wake EO	Beta	–
Wake EC	Beta	Beta (and alpha) (9.77-25.15 Hz) was higher after sleep deprivation (Corsi-Cabrera et al., 1992).
Sleep	Delta	Delta waves in the sleep EEG were an indicator of sleep

Condition	Frequency Band	Findings
		intensity (Blake & Gerard, 1937; Borbely, 1982).
		Daytime naps reduced delta during the following night's sleep (Campbell & Feinberg, 2005).
		Reduced SWS related to poorer reaction time in normal subjects (Jurado, Luna-Villegas, & Buela-Casal, 1989).
		Delta and theta activity increases during sleep after 40 hours sleep deprivation, with the changes most pronounced in the frontal regions (Cajochen, Foy, & Dijk, 1999).
		Sleep delta increased in recovery sleep after 40 hours sleep deprivation. The increase was highest in the frontal region (Munch et al., 2004).
		Sleep delta correlates negatively to spindle range (13.25-15.0) power in NREM (Aeschbach & Borbely, 1993).
		During recovery sleep after 40 hours sleep deprivation, spindle activity varied inversely with slow-wave activity in NREM, except at the very beginning and end of each NREM cycle (D. J. Dijk, Hayes, & Czeisler, 1993).
		Slow wave activity (0.75-4.5 Hz) declined exponentially towards a horizontal asymptote in baseline sleep and after 36 hours sleep deprivation (D. J. Dijk, Brunner, & Borbely, 1990).
		Auditory tone induced slow-wave **sleep disruption** led to rebound increased delta and theta activity afterwards, but it did not increase sleep duration (D. J. Dijk & Beersma, 1989).
		Increasing theta (EO) activity during extended waking (40 hours) correlated with increased SWA in the first NREM sleep cycle (Finelli et al., 2000).
		Low frequency delta (0.5-1.0 Hz) activity during the first cycle of NREM sleep correlated with daytime performance on tests of prefrontal cortex function: Wisconsin Card Sorting Test, Tower of London, etc (Anderson & Horne, 2003). Greater sleep delta was associated with better performance.
		Sleep SWA increased in the right parietal area after a motor learning task. Decreased directional error correlated with the peak increase in SWA (Reto Huber,

Condition	Frequency Band	Findings
		Felice Ghilardi, Massimini, & Tononi, 2004). More sleep SWA was associated with better learning. Subjects with sleep maintenance insomnia showed a lower increase in SWA during sleep after 21 hours of sleep deprivation compared with controls, suggesting they had accumulated a lower sleep pressure (Besset, Villemin, Tafti, & Billiard, 1998). Insomniac older (> 60 years) subjects with slower reaction times showed reduced slow wave power (2-4 Hz) during sleep (Crenshaw & Edinger, 1999). In OSA, SWA in the first NREM cycle correlated with the MSLT (more SWA, less sleepy), and SWA rises in the 1st and 2nd NREM cycle after CPAP (Heinzer et al., 2001). OSA subjects had lower sleep delta than UARS patients and controls (Guilleminault et al., 2001). OSA subjects had slower exponential decay of SWA across the night compared with controls. SWA was not significantly lower in OSA vs. controls (Ondze, Espa, Dauvilliers, Billiard, & Besset, 2003). After CPAP, delta increased in stage 2 sleep (Wang, Chen, Bian, & He, 2002). Visually scored sleep slow wave sequences are reduced in the first NREM cycle in 8 OSA subjects compared with age/sex matched controls (Himanen, Joutsen, & Virkkala, 2004).
Sleep	Theta	Sleep theta power correlated with the MSLT (Spearman r = 0.55) (Wichniak et al., 2003). Those with MSLT < 10 had a trend to reduced sleep Delta. Delta and theta activity increases during sleep post 40 hours sleep deprivation, with the changes most pronounced in the frontal regions (Cajochen, Foy et al., 1999). Auditory tone induced slow-wave **sleep disruption** led to rebound increased delta and theta activity afterwards, but it did not increase sleep duration (D. J. Dijk & Beersma, 1989).

Condition	Frequency Band	Findings
Sleep	Alpha	–
Sleep	Beta	Beta power during sleep increased in subjects with high ESS (Wichniak et al., 2003). Subjects with insomnia had more spectral power in 26-30 Hz in sleep than normals (Freedman, 1986).
Sleep	Sigma	Sleep sigma corresponds to the presence of intermittent sleep spindles (D. J. Dijk et al., 1993). Delta correlates negatively to spindle range (13.25-15.0) power in NREM (Aeschbach & Borbely, 1993). During recovery sleep post 40 hours sleep deprivation, spindle activity varied inversely with slow-wave activity in NREM (Except at the very beginning and end of each NREM cycle) (D. J. Dijk et al., 1993).

Table 1.2 Summary of EEG literature pertaining to sleepiness

Limitations of quantitative EEG in measuring sleepiness

> *"Until one knows definitely the parts of the brain giving rise to the different types and the factors determining their appearance or disappearance, a logical system of nomenclature is impossible* (Loomis, Harvey, & Hobart, 1936)*"*

There are several limitations concerning the use of quantitative measures of the EEG as a measure of sleepiness. The measurement is cumbersome, prone to artefact, and different authors have used different metrics and different definitions of the frequency bands. The artefact

problem can be partially compensated for by procedures to subtract the signal caused by eye blinks (Croft, Chandler, Barry, Cooper, & Clarke, 2005), or muscle artefact (Shwedyk, Balasubramanian, & Scott, 1977).

The concept of EEG power in a frequency band is still an empirical one, based on the frequencies of waveforms observed with visual scoring. Reported findings can depend on the range of frequencies defining the bands (Darchia, Campbell, Tan, & Feinberg, 2007). Rhythms in the alpha frequency range may arise from more than one process or area within the brain (Klimesch, 1999). No direct interpretation can be made on underlying brain processes, aside from there being a more or less electrically active cortex, and perhaps the thalamocortical interactions that may underlie characteristic waveforms such as a sleep spindle or a K-complex. Fluctuations in EEG power that have been noted as a marker of sleepiness may not actually be generated by the brain. Eyes-open alpha power may increase in the sleepy state due to increased intermittent eye closure, and delta power may be increased from the greater frequency of eye blinks (Torsvall & Akerstedt, 1985).

Inter-individual variation is also an issue with EEG measurements. It has been established that some characteristics of the EEG are trait-like (Hoptman & Davidson, 1998), and there are inter-individual differences in the relationship between EEG and sleepiness-related measures (Torsvall & Akerstedt, 1988).

The EEG mathematical model

A mathematical model (EEG model, also referred to as the EEG inverse model) has been developed that examines interactions of groups of neurons in the cortex and thalamus (Robinson

et al., 2005; Robinson et al., 1997). A less mathematically-oriented description of this model is extant (Rowe, Robinson, & Gordon, 2005). A brief description of the model follows, paraphrased from these sources.

Neuronal firing: dendritic inputs to axonal response. In a cortical pyramidal cell, action potentials from excitatory, inhibitory or subcortical neurons arrive at the dendritic tree, with the total effect being the sum of the product of pulse densities, number of synapses, and post-synaptic response for each population of neurons. The total dendritic impulse affects the membrane potential of the cell body V_a, after modulation by dendritic rate constants α and β, which parameterise the rise and decay rate of the impulse response. As the EEG is assumed to be due to small perturbations about a steady state, the sigmoidal relationship of membrane action potential to firing rate is simplified as a linear response with slope ρ. The gain, G, is then ρNs, where N is the number, and s the response strengths of the synapses.

Signal propagation and EEG generation. EEG arises mainly from pyramidal cells in the cortex. Other structures, such as the thalamus, exert their influence indirectly via the pyramidal cells. Stellate cells (the local interconnections) have much shorter axonal range than the pyramidal cells, their axonal lengths and delays are ignored. Axons from pyramidal cells decrease in synaptic density as axonal range increases, so terminal density decreases at exponential rate. Inputs to the cortex are from sensory signals arising from the periphery and the activity from the corticothalamic pathways. Delays in signal propagation and the additional nuclei in the pathways contribute to resonance in the system.

Cortical and thalamic connections (Figure 1.5). The cortex is treated as a two-dimensional continuum. Connections between neurons consist of short-range (local) inhibitory and excitatory connections, and long range excitatory connections by pyramidal cells (corticocortical and

corticothalamic projections). Projections to thalamocortical relay (TC relay) cells are matched by reciprocal projections to the cortex, forming corticothalamic loops. There are also projections to inhibitory neurons of the thalamic reticular nucleus (TRN) which then project to TC relay nuclei.

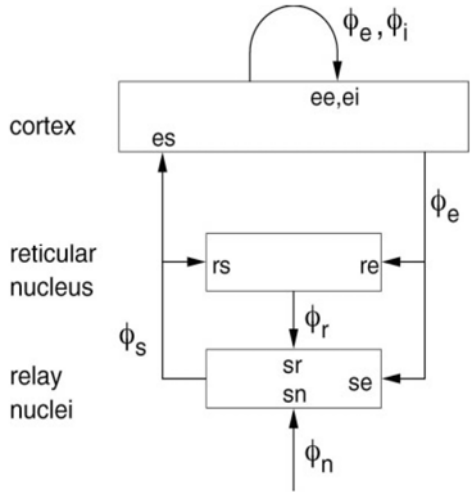


Figure 1.5 Representation of the interactions between groups of neurons in the cerebral cortex, thalamic reticular nucleus and thalamic relay nuclei described by the mathematical model

Cortical neurons receive excitatory and inhibitory inputs from other neurons within the cortex. (φ_e, φ_i) Ascending input from the periphery (φ_n) passes to the cerebral cortex (φ_n) via the thalamic relay nuclei. This relay of the signal is inhibited by an input from the reticular nucleus. (φ_r) There are descending inputs to the thalamus from the cortex (φ_e) (Robinson, Rennie, Rowe, & O'Connor, 2004).

Figure 1.5 shows the pathways between cortex and thalamus included in the model (Robinson et al., 2004). The direct corticothalamic loop is comprised of neural activity from excitatory corticothalamic efferents φ_e that synapse with thalamic relay nuclei, and then a return limb to the cortex, φ_s. The total gain in this pathway is represented by G_{ese} ($G_{ese} = G_{es}*G_{se}$). This quantity is positive because both G_{es} and G_{se} are gains of excitatory neurons.

There is a second pathway involving corticothalamic signals that pass through the inhibitory thalamic reticular nucleus (φ_r), before returning to the cortex via the thalamic relay nuclei. Total

gain is represented by $G_{esre} = G_{es}G_{sr}G_{re}$. A third loop is between the thalamic reticular nucleus and the thalamic relay nuclei, with total gain G_{srs} ($G_{srs} = G_{sr}G_{rs}$). The sensory signals from the periphery (φ_n) is treated as white noise.

Electromyogram (EMG) artefact from firing of scalp, facial, jaw and neck muscles can affect the EEG spectra, particularly above 25 Hz. The effect of the EMG is separately modelled (Shwedyk et al., 1977; van Boxtel, 2001), so the EEG power spectrum can be analysed independently of the effect of EMG artefact.

Model parameters. Model parameter values are constrained to 'physiologically valid' values as determined by review of the literature, so as to improve convergence of the fitting algorithm (Rowe, Robinson, & Rennie, 2004). The parameters within the model are listed in Table 1.3. Some parameters can be assumed to be relatively fixed, such as the cranial filtering parameter k_0r_e, axonal range r_e. The dendritic rate constant β is set at 4α, based on data from 100 normal subjects. Nine EEG model parameters are fitted from experimental data. There are also three derived or transformed parameters, X, Y and Z, which are calculated from the fitted models. These three parameters allow a simpler conceptualisation of the system in three dimensions. Preliminary data suggest these transformed parameters to have greater test-retest reliability than the raw gain parameters (van Albada, Rennie, & Robinson, 2005). By convention, gain parameters from inhibitory interactions are given a negative sign. Thus a higher absolute magnitude of a gain parameter would represent a greater inhibitory influence.

Parameter	Description	Mean value	Notes

Parameter	Description	Mean value	Notes
γ_e	Cortical damping (v_e/r_e)	140	
α	Dendritic decay rate	75	
β	Dendritic rise rate		$\beta = 4\alpha$
t_0	Conduction delay through thalamic nuclei and projections	0.084	
G_{ee}	Excitatory gain, pyramidal cells	5.8	$G_{ee} = G_{ie}$
G_{ei}	Local intracortical gain, stellate cells	-7.5	$G_{ei} = G_{ii}$
G_{ese}	Cortico-thalamocortical gain via thalamic relay nuclei	5.4	
G_{esre}	Cortico-thalamocortical gain via thalamic reticular nucleus	-3.3	
G_{srs}	Intrathalamic gain	-0.50	
$k_0 r_e$	Volume conduction filter parameter	3.0	Constant, not fitted.
r_e	Characteristic pyramidal axon length	0.08	Constant, not fitted.
P_0 (also l_{norm})	Overall power normalization	2.49	Calculated from data.
X	Net cortical excitation.		$x = G_{ee}/(1 - G_{ei})$,
Y	Net cortico-thalamic excitation		$y = \frac{G_{ese} + G_{esre}}{(1 - G_{srs})(1 - G_{ei})}$,
Z	Internal feedback between thalamic reticular nucleus and thalamocortical relay nuclei.		$z = -G_{srs}\alpha\beta/(\alpha + \beta)^2$
EMG model			
A	Power normalization	0.5 $\mu V^2 Hz^{-1}$	
f_{pk}	Spectra peak frequency	40 Hz	

Parameter	Description	Mean value	Notes
δ	Asymptotic slope	2	

Table 1.3 Parameters of the EEG model

Mean values are awake eyes closed values from 100 healthy volunteers (Rowe, Robinson, & Rennie, 2004).

Validation of the model. Normative values of the EEG model parameters have been published, from eyes open and eyes closed resting EEG recordings taken from 100 normal volunteers (Rowe, Robinson, & Rennie, 2004). There has also been work on the test-retest reliability of the parameters (van Albada et al., 2005). Apart from the waking state, the model has also been fitted to EEG from recordings representing stages of non-rapid eye-movement sleep (Robinson, Rennie, & Rowe, 2002), and has been used to generate EEG waveforms resembling various types of seizure activity (Robinson et al., 2002), and event-related potentials (C. J. Rennie, Robinson, & Wright, 2002). Predictions from the model have suggested that corticothalamic delay (the t_0 parameter) is the dominant contributor to the alpha frequency (Robinson, Whitehouse, & Rennie, 2003).

The EEG model has also been applied to the study of clinical conditions such as attention deficit hyperactivity disorder (ADHD) (Rowe, Robinson, & Gordon, 2005; Rowe, Robinson, Lazzaro et al., 2005), and memory loss (Alexander et al., 2006). In the first study, fifty-four subjects with diagnosed ADHD were studied off-therapy, and compared with 54 age and sex matched controls (Rowe, Robinson, Lazzaro et al., 2005). The waking eyes-closed EEG was fitted to the EEG model. The parameter α was reduced ($p < 0.001$) in the ADHD group, implying longer dendritic response times. Gain parameters G_{ii} ($p = 0.05$) and G_{srs} ($p < 0.05$) were increased in magnitude in the ADHD group (that is, the absolute value of the parameter was greater). The latter implies a

greater degree of inhibitory intrathalamic gain. The study also found that power in the 7 Hz bin was increased in the ADHD group, relative to controls. The second study studied 11 ADHD patients before and after treatment with stimulant therapy (dexamphetamine or methylphenidate) (Rowe, Robinson, & Gordon, 2005). After medication therapy, there was a trend to reduced theta and delta EEG power after medication therapy, a trend to reduced magnitude of G_{srs} ($p = 0.06$) and G_{ei} (not significant, p-value was not quoted), while there was a significant reduction in G_{ee} ($p < 0.05$).

Based on the theorised interaction between thalamus and cortex being important in the onset and deepening of sleep (Steriade, 2003), we might expect the intrathalamic gain (G_{srs} and parameter Z) to increase in magnitude with increasing sleepiness during wake, and also to increase with greater depth of NREM sleep, as this might represent greater influence of the thalamic reticular nucleus upon the thalamocortical relay cells, leading to their hyperpolarisation. As amphetamines are often used as wake-promoters in disorders of excessive somnolence such as narcolepsy, the finding that amphetamine was associated with reduced magnitude of G_{srs} would be consistent with this proposal. With greater sleepiness, and deepening sleep, we might also expect gain within the intracortical networks to be reduced, hence G_{ee}, and parameter X should fall. Conversely, with increased effort required to stay awake or to maintain satisfactory performance, cortical gain may be increased as a compensatory measure. The finding of reduced cortical gain with amphetamine therapy appears to support the second proposal.

The third study recruited 79 healthy subjects aged 52-88 with self-reported memory complaints, and compared them with 79 controls from a normative dataset. Eyes-closed resting EEG spectra from lead Cz was fitted to the EEG model. The group with memory complaints had reduced parameter α ($p = 0.02$), inhibitory local cortical gain (G_{ei}, $p = 0.02$) and intrathalamic inhibitory

gain (G_{srs}, p = 0.003) compared with controls. Within the memory complaints group, high intrathalamic gain was associated with better performance on a maze task, a measure executive function. Puzzlingly the result from this study with respect to G_{srs} appears inconsistent with the inference drawn earlier from the amphetamine study.

Physiological measures may need to be observed in context, and possibly also with carefully planned experimental manipulation, in order that the primary process and the secondary antagonistic process may be identified.

Assessing the EEG model as a measure of sleepiness

Validity refers to the extent to which a measurement measures what it is designed to measure. The key requirements to establish the validity of a measure are summarised in Table 1.4 below (Kirshner & Guyatt, 1985; Terri E. Weaver, 2001). Examples of experimental designs to validate the EEG model in measuring sleepiness are included in the last column of the table.

An important issue in deciding what is needed to validate the test lies in determining the purpose of the measure of sleepiness: whether it is to be applied to **discriminate** between groups, or whether to **evaluate** changes in the level of sleepiness within individuals (Kirshner & Guyatt, 1985). Part of the evaluation of the new measure will involve comparing its behaviour with other established measures of sleepiness. Participants could be sought with different levels of sleepiness, or the same individuals might be measured at different periods when we would expect their level of sleepiness to be different. Sleep deprivation protocols have been commonly employed for the latter comparison, as the intervention is capable of inducing marked changes in the level of sleepiness.

While the theorised interactions between cortex and thalamus cannot be directly measured, drug or experimental manipulation of states expected to change these underlying cortico-thalamic interactions may provide a means to evaluate the theoretical basis of the EEG model. The increasing depth of NREM sleep, and also the transition from drowsiness to sleep, are proposed to follow inhibition of the thalamic relay nuclei mediated by the thalamic reticular nuclei (Steriade & Amzica, 1998). Hence demonstration of theoretically-consistent changes in EEG model parameters related to the cortico-thalamic interactions will support the face validity of the model and its possible role as a measurement of sleepiness.

The ultimate component in external validity lies in its applicability in the real world, for example the ability to determine that someone is alert enough for work or driving, or determining if a treatment has worked sufficiently.

Psychometric characteristic	Explanation	Proposed method of assessment
Reliability		
Internal consistency	The degree to which one item correlates with another and the scale as a whole. It depends on the number of items in an index, as well as the set of correlations among them.	Not relevant to EEG model parameters, which are individual items.
Test-retest reliability	The measure does not vary significantly between repeat administrations.	In healthy individuals and in those reporting sleepiness, demonstrate consistency of the measurement across repeated measurements. Note that sleepiness can change in an individual – we would also want the measure to reflect true fluctuations e.g. with time of day. The reliability in healthy individuals is the subject of another doctoral thesis (van Albada et al., 2005).
Validity		
Content or face validity	Expert confirmation that items on measure reflect areas expected to respond to change in health status.	There is face validity of the original EEG signal as a measure of sleepiness. The theorised interaction between thalamus and cortex also lends face validity to the mathematical model.

Psychometric characteristic	Explanation	Proposed method of assessment
Construct validity	The measure relates well to other measures of the phenomenon.	Examine correlations between EEG model parameters and other measures of sleepiness. (page 123) Note the existence of different types of measures covering different nuances of sleepiness. Establish the extent to which the EEG model parameters linked to sleepiness are state vs. trait characteristics. Similar to establishing the reliability of the measure, evaluate stability of the measure across repeated measurements. Within subjects, examine if changes within subject in the measure bear an expected relationship to other variables measuring sleepiness in intervention studies of sleep deprivation, wake promoting drugs in narcolepsy, and the treatment of OSA.
Discriminant validity	The measure differentiates between two groups known to differ on only the outcome of interest.	Distinguish those with significant and pathological sleepiness from those without: subjects who test positive should demonstrate meaningful impairment such as reduced quality of life, cognitive impairment, or increased car accident risk. Distinguish disease groups with significant sleepiness from those without, e.g. narcolepsy and OSA. This is approach can be problematic because there are those with OSA without sleepiness. Within occupational group, detect those too sleepy to work or drive. Collect normative data (Rowe, Robinson, & Rennie, 2004).
Responsiveness	The test is able to detect a clinically important difference.	MCID should be the increase in the sleepiness measure after 14-18 hours awake, given correspondence with blood alcohol levels close to the legally-accepted limit for driving in most countries (Williamson & Feyer, 2000).

Table 1.4 Requirements to establish the validity of the EEG model as a measure of sleepiness

1.4 OSA as a disease model of sleepiness

Sleepiness is a key symptom in OSA

Excessive sleepiness is a key symptom of obstructive sleep apnea (OSA), and a major source of the morbidity arising from it (Guilleminault, Stoohs, & Duncan, 1991; Young et al., 1993). Sleepiness is thought to be due to a combination of sleep fragmentation and chronic intermittent hypoxia (Beebe & Gozal, 2002). There is also possibly some non-reversible cerebral damage from exposure to the repetitive apnea events, as suggested by imaging studies, and the persistence of sleepiness in some despite treatment (Morisson et al., 2001; Sforza & Krieger, 1992).

Various neurocognitive impairments have been observed with OSA, particularly with regard to tests of vigilance, memory and executive function (Aloia, Arnedt, Davis, Riggs, & Byrd, 2004; Brown, 2005; Engleman & Joffe, 1999). Many of these coincide with the types of neurocognitive impairments thought to be associated with sleepiness. There is no reason to suspect that the type of sleepiness seen in OSA would represent a different phenomenon to the concepts associated with sleepiness above (E. Verstraeten, Cluydts, Pevernagie, & Hoffmann, 2004). The delayed recovery in some tests of executive function up to four months after treatment of sleep apnea may suggest that sleep loss may not be the sole reason for the impairment seen in OSA (Bearpark, Grunstein, Touyz, Channon, & Sullivan, 1988).

It should be noted that excessive sleepiness is not universal amongst sufferers of OSA (Black, 2003). It would not be appropriate to employ a measure of sleepiness to distinguish OSA patients from those without OSA, however differing levels of sleepiness might be appreciated within a group.

Whatever the nature of the relationship between pure sleepiness and neurocognitive impairment, all these concepts are relevant to daytime function in an individual with OSA. Sleepiness is closely related to the interaction between daytime function and the injury that OSA imposes on the brain or organism.

Why would OSA be a good disease model of sleepiness?

Many aspects of OSA make it a good disease model of sleepiness. The symptom of sleepiness is important in OSA, and the ability to measure sleepiness has relevance. OSA is common (Young et al., 1993), and a population of otherwise healthy subjects is potentially available to undertake various testing procedures. The condition is chronic and stable over the intermediate time period, allowing time for a variety of different tests to be conducted, and there are established methods to diagnose OSA and measure the degree of OSA.

The disease is reversible with treatments such as CPAP, which can abolish the sleep apnea events that define the disease, hence alleviating the cause of the sleepiness.

There are precedents in the literature. Other measures such as the Epworth Sleepiness Scale (ESS) and Functional Outcomes of Sleep Questionnaire (FOSQ) have used OSA patient populations in validation. There have also been studies in OSA patients showing they have lower sleep delta activity (Guilleminault et al., 2001; Himanen et al., 2004), and relating low sleep delta power to greater sleepiness-related measures during the day (Heinzer et al., 2001). Wake EEG changes suggesting increased sleepiness showed reversal after treatment of the OSA

with continuous positive airway pressure therapy (Morisson et al., 2001; Morisson et al., 1998) (See page 124 for a review of EEG studies of OSA.).

There are some potential limitations to the use of OSA as a model of sleepiness. As mentioned above not all OSA sufferers report sleepiness. Moreover, while treatment of OSA reduces sleepiness, some report residual sleepiness. There is also the possibility of other effects of OSA on our outcomes that are not mediated by sleepiness, including the presence of nocturnal intermittent hypoxia, and the possibility of irreversible brain injury from chronic exposure to sleep apnea events. It is not known the extent to which the chronicity of the disease might allow the individual to adapt to any OSA-mediated brain injury or sleep disruption. Lastly, although there are established methods to diagnose and measure the severity of OSA, strong relationships between daytime functional outcomes and polysomnographic measures of OSA have not been consistently demonstrated (For example Adams, Strauss, Schluchter, & Redline, 2001; Cheshire, Engleman, Deary, Shapiro, & Douglas, 1992; Kim, Dinges, & Young, 2007 in press).

Thus, although there may not be complete congruence between the symptoms and cognitive impairments experienced in OSA and the abstract quality sleepiness, it forms a convenient model in which sleepiness can be studied. Other models include sleep deprivation and other disease states such as narcolepsy.

1.5 Summary and study aims

Sleepiness is a common problem, and an important source of morbidity, yet there is no gold standard measure of sleepiness. Currently-available measures have their limitations (page 28), including the presence of different nuances or dimensions of sleepiness, a lack of standardisation in procedures for the tests, the limited validation undertaken for many tests, and the derivation of many tests as research rather than clinical tools. A better measure of sleepiness is needed, and such a tool would have applications both in research, and clinically in the assessment and management of sleep disorders such as obstructive sleep apnea.

The common sleep disorder OSA represents a clinically-relevant setting in which to develop and test a potential measurement of sleepiness (page 53). Any proposed new measurement of sleepiness in OSA would need to be validated against sleepiness-related outcomes relevant to patients with the condition. There are difficulties with the measurement of these outcomes in OSA, with no widely-accepted, single battery of assessments (page 54, and discussed further in page 60). The assessments used need to be sensitive to the effects of OSA, and moreover methods need to be employed to deal with the multiple outcome variables obtained, and allow separation of different nuances of sleepiness from other effects of OSA.

The EEG has been studied, particularly in sleep deprivation research, as a physiologically-based marker of sleepiness (pages 30, 42). Quantitative EEG analysis methods studied to date have not been validated for clinical use to measure sleepiness. A new method of EEG analysis has potential advantages of a theoretical basis, hence allowing inferences to be made about underlying brain processes. This method has not yet been examined as a measure of sleepiness, and its validity of this method needs to be established (page 50). Additionally, in the absence of a

readily measurable biological substrate for sleepiness (or to support the theoretical basis of the measurement), the new measurement may to be referenced to naturally-occurring or experimentally-induced changes in physiological processes.

The aim of this study is to improve the measurement of sleepiness. We will assess the mathematical modelling of the EEG power spectra ("the measure") as a means of measuring sleepiness, using a clinical group of patients with sleep apnea.

Specific aims are to:

1. Confirm a new neurocognitive battery is sensitive to the effects of sleep apnea.

2. Assess symptoms and daytime function in a sleep centre population with varying severities of sleep apnea.

3. Aggregate outcomes into domains by factor analysis to reduce number of statistical comparisons to be made.

4. Measure waking EEG in this population, and extract parameters from analysis of the EEG.

5. Relate the EEG measures to functional outcomes relevant to sleepiness.

6. Describe the time course of EEG model parameters with increasing depth of non-rapid eye movement sleep.

2 A new neurocognitive battery is sensitive to the effects of OSA

2.1 Chapter abstract

Obstructive sleep apnea (OSA) is expected to impair vigilance and executive functioning, owing to the sensitivity of the prefrontal cortex to the effects of sleep fragmentation and intermittent hypoxia. Studies examining the pattern of cognitive dysfunction show variable results, with the heterogeneity in part due to small sample sizes in current studies and little consistency of the tests used. This problem has prompted the call for the use of a comprehensive testing battery (Decary, Rouleau, & Montplaisir, 2000), and as a result several are being established. We examined a group of fifty subjects from the Brain Resource International Database (BRID) predicted to have OSA on the basis of the Multivariable Apnea Prediction Index, and compared them with 200 matched controls. On electrophysiological tests, the OSA group showed reduced eyes open alpha power, increased auditory oddball N100 and P200 amplitude, but reduced N200 and P300 amplitude. The latency to P300 was not significantly different between groups, but latencies to N200 and P200 were prolonged in the OSA group. Performance testing of executive function found verbal interference and switching of attention was impaired in the OSA group. The study demonstrated that a group of subjects with likely OSA identified by a diagnostic algorithm based on apnea symptoms and demographic factors manifested deficits in information processing and executive function.

The text of this chapter has previously been published as Wong, K. K., Grunstein, R. R., Bartlett, D. J., & Gordon, E. (2006). Brain function in obstructive sleep apnea: results from the Brain Resource International Database. Journal of Integrative Neuroscience, 5(1), 111-121.

2.2 Introduction

Obstructive sleep apnea (OSA) is characterised by repetitive episodes of upper airway obstruction that occur during sleep. OSA is thought to impair brain function through the combined effects of intermittent hypoxemia and sleep fragmentation from repetitive apneas (Beebe & Gozal, 2002).

Studies examining the effect of OSA on brain function have yielded inconsistent results in terms of the abnormalities demonstrated in neurocognitive performance testing. In general, deficits have been observed in vigilance and various components of executive functioning (Aloia et al., 2004; Beebe, Groesz, Wells, Nichols, & McGee, 2003; Fulda & Schulz, 2003). One proposed mechanism for this pattern of deficit relates to the prefrontal cortex, the area crucial for executive function, being more vulnerable to the effects of sleep fragmentation and intermittent hypoxia because it is highly metabolically active (Harrison & Horne, 2000). Functional magnetic resonance imaging (fMRI) has also shown that the prefrontal cortex is sensitive to the effects of sleep deprivation (Drummond et al., 1999).

The reliability of neurocognitive tests can be affected by factors such as the variety of tests across laboratories, timing of the testing, and the motivation of the subject. These factors can adversely affect the sensitivity of the tests within a study, and consistency of findings across studies. It has been proposed that standardised testing batteries (that assess a profile of cognitive function) should be used to address this problem (Decary et al., 2000). In degenerative neurological or psychiatric illness such as the dementias, Parkinson's disease and depression, event-related potentials (ERP) have been used as a sensitive tool to distinguish groups with cognitive impairment and disturbed speed of information processing from unaffected controls.

They are partially generated in subcortical structures, and can indicate cognitive dysfunction not evaluated by the neurocognitive tests (Kotterba et al., 1998). Changes on ERP, such as the P300 wave, have been interpreted as reflecting problems with cognitive processing independent of attention. Many studies in OSA have shown prolonged P300 latency, implying deficits in processing the context of information (Inoue, Nanba, Kojima, Mitani, & Arai, 2001; Kotterba et al., 1998; R. B. Sangal & Sangal, 1997a, 1997b; Walsleben, Squires, & Rothenberger, 1989).

Aim and hypotheses

The purpose of this study is to explore the pattern of neurocognitive dysfunction in obstructive sleep apnea in a new neurocognitive testing battery. We selected cases and matched controls from the normative dataset of the Brain Resource International Database (BRID) of healthy volunteers who have undergone a standardised protocol which probes multiple aspects of brain function (Evian Gordon, 2003). A prediction algorithm validated in a sleep clinic population was employed to select the subjects (Maislin et al., 1995).

The hypotheses were that

1. the battery of tests employed by the BRID would be sensitive to detecting neurocognitive impairment in OSA;
2. performance in tasks requiring vigilance and executive functioning would be impaired in those with OSA, and
3. evoked response potential metrics probing cognitive processing would be diminished and delayed in the OSA group compared with their matched controls.

2.3 Methods

Participants

All subjects aged 18-70 years from the BRID with likely obstructive sleep apnea (OSA) as determined by a Multivariable Apnea Prediction Index (MAPI) of greater than 0.50 were selected for inclusion in the OSA group (Maislin et al., 1995). The MAPI is calculated from gender, age, body mass index, and the average score from three self-administered questions regarding symptoms of sleep apnea (snorting or gasping, loud snoring, and breathing stops, choking or struggling for breath, comprising "Index 1" of the questionnaire). Validated in clinical populations, the MAPI reflects the probability of being diagnosed with obstructive sleep apnea. A MAPI threshold of 0.50 has a sensitivity of 0.88 and specificity 0.55 in diagnosing OSA (Maislin et al., 1995).

For each subject in the OSA group, we found 4 closest matches for gender and age from the database who scored zero for the index 1 questions, and had MAPI < 0.50, who then comprised the control group.

Data from the BRID had been acquired from 6 laboratories in Australia, North America and Europe in a standardised manner. Volunteers were recruited to form a large normative sample (over 2,000 subjects) spanning a wide age range. The normative database excluded those with history of psychiatric, neurological or other serious medical disorders. All subjects provided written informed consent to participate in the database.

Testing procedure

The testing protocol has been previously described (E. Gordon, Cooper, Rennie, Hermens, & Williams, 2005; Kemp et al., 2005; L. M. Williams et al., 2005) and comprises of psycho-physiological tests, a cognitive test battery, and a series of questionnaires. The methods of the tests reported in this study are briefly described below. Subjects refrained from caffeine and smoking for at least 2 hours prior to testing.

Questionnaire

Subjects completed questionnaires including demographic factors, the MAPI questionnaire described above (Maislin et al., 1995), and the Depression Anxiety Stress Scale (DASS) (S. H. Lovibond & Lovibond, 1995).

Resting Electroencephalography (EEG)

Two minutes each of eyes open and eyes closed EEG was collected in the resting state. EEG collected from frontal (Fz, F3, F4, F7, F8) and posterior (Pz, P3, P4, T5 and T6) regions with an electrode cap (Nuamps, Compumedics, Abbotsford VIC, Australia) was analysed. Eye movements were recorded with electrodes placed 1.5 cm lateral to the outer canthus of each eye, 3 mm above the middle of the left eyebrow, and 1.5 cm below the middle of the left bottom eyelid. The sampling rate was 500 Hz, with a low pass filter attenuating 40 dB per decade above 100 Hz. The signals were re-referenced offline to linked mastoid electrodes, and were corrected for eye movement artefact (Gratton, Coles, & Donchin, 1983). Average power spectra were obtained by Fast Fourier Transform (FFT) using consecutive 2-second periods with a Welch

window. The FFT spectrum was grouped into delta (0.5-3.5 Hz), theta (4.0-7.5 Hz), alpha (8-12 Hz), and beta (12-35 Hz) frequency bands.

Auditory Event-Related Potentials (ERP)

ERPs were recorded in a standard auditory oddball paradigm. A series of high and low tones were presented at 75dB, lasting 50ms, with inter-stimulus interval 1s. Rise and fall times of the tones was 5 ms. Subjects were asked to press buttons held in both hands in response to the target tones (1000 Hz), but not to respond to the background tones (500 Hz). The duration of the task was 6 minutes. Components of the ERP analysed were target stimulus latency and amplitude for P300b, N200, P200, N100; and background stimulus latency and amplitude for P200 and N100.

Cognitive test battery

Subjects performed a battery twelve tasks in front of a touch-screen computer. The tasks analysed in this study are described in Table 2.1. These tasks were chosen to represent cognitive domains potentially vulnerable to OSA, with particular emphasis on vigilance and executive function.

Performance metrics derived from the chosen tests were the metrics recommended by the testing manual of the neurocognitive test battery (Brain Resource Company, 2005).

Test	Description	Key metrics
Choice reaction time	One of four circles lights up and the subject has to press the lit circle as quickly as possible	Average reaction time.
Timing test	A circle appears on the screen for 1 to 12 seconds, then the subject is required to indicate the correct duration.	Proportional bias.
Span of visual	The subject is required to press a series	Length of the longest sequence

Test	Description	Key metrics
memory	of squares on the screen in the order in which they previously lit up	correctly identified twice.
Digit span (forward and backward)	The subject is presented with a sequence of digits and has to repeat them in either forward or backward order.	Length of the longest sequence correctly recalled in forward or reverse order.
Verbal interference	The subject is required to name the 'ink' colour a word is written in, and not read the actual word.	The number of correct responses in naming the colour or reading the text of the displayed word.
Switching of attention	Numbers and letters are connected in various sequences.	Total time to connect the sequence of numbers or numbers and letters.
Memory recall and recognition	The subject has to recall a set of words after various time intervals and later recognise the words from list of repeated and new words.	Total score trials 1-4; score trial 7 (delayed recall); Recognition accuracy (number of words from a prompt list that were correctly recognised).
Maze test	A dot-based maze is presented on the screen. Using a directional button box, the subject is required to discover (by trial and error) a hidden path through the maze and remember it.	Number of trials completed.

Table 2.1. Description of the neurocognitive tests and the key metrics analysed.

Statistical Analyses

Differences between groups in the three subscales of the DASS were analysed by univariate analysis of variance. Analysis of covariance (ANCOVA) was used to analyse the effect of group on the cognitive test outcomes, with age and years of education as covariates.

Eyes open and then eyes closed EEG power spectra in delta, theta, alpha and beta bands were analysed using multivariate analysis of covariance (MANCOVA) with group, brain region (anterior or posterior), and age as predictors. Univariate analyses for individual frequency bands were explored if there was a significant effect of group ($p < 0.05$). Alpha peak frequency and

amplitude in both eyes open and eyes closed conditions (4 outcome variables) were also analysed with a similar MANCOVA. The 12 evoked response potential outcomes were analysed with a similar MANCOVA. In view of the multiple outcome variables available for the psychophysiology measures, a two-stage statistical test was employed: an initial overall F test, followed by univariate tests with the alpha level adjusted by the Bonferroni method.

2.4 Results

From the dataset of all 1961 subjects in the BRID in September 2004, fifty subjects met the inclusion criteria and were classified in the OSA group. Two-hundred of the closest matched subjects from the same database comprised the control group. Table 2.2 shows demographic and sleep questionnaire information used to classify subjects into the OSA or control groups. As expected, the OSA group had a higher BMI than the control group (37.7 vs. 24.9, $p < 0.00001$, 95% CI of difference 11.2 – 14.4).

	OSA	Control
M:F ratio (% males)	35:15 (70%)	140:60 (70%)
Age (yrs)	51.59 (13.42) range 19.7-69.99	50.53 (12.83) range 19.7 – 69.80
BMI (kg/m2)	37.66 (9.11)	24.86 (3.40)
MAP index	0.69 (0.12)	0.20 (0.12)
Index 1 of MAP	1.04 (1.36)	0 (0)
Years of education	13.6 (3.76)	13.9 (3.27)
Reported sleep duration (h)	6.46 (2.0)	6.8 (1.4)
Coffees / week	9.9 (5.2)	10.4 (5.4)

Table 2.2. Demographic and sleep questionnaire responses in the two groups

Results are presented as mean values with standard deviations in brackets.

DASS questionnaire

In a sample in which significant depression has been excluded, overall scores were low for the subscales of the DASS, although the values for the OSA group tended to be higher in all three subscales although this apparent difference was not significant (Table 2.3).

	OSA	Control	p
Depression	1.80 (2.05)	1.27 (1.83)	0.096
Anxiety	1.09 (1.32)	0.77 (1.23)	0.13
Stress	2.45 (2.60)	2.31 (2.52)	0.76

Table 2.3. Univariate analyses for DASS subscales: effect of group alone

Waking resting electroencephalography (EEG)

MANCOVA tests with group, age and scalp region as predictors were significant for waking eyes closed EEG band power ($F_{4,44}$ = 5.45, p = 0.0012), waking eyes open band power ($F_{4,44}$ = 4.12, p = 0.006), and for alpha peak frequency and amplitude ($F_{4,43}$ = 5.47, p = 0.0012).

Univariate analyses of the EEG bands, using a Bonferroni-adjusted significance level of 0.004, showed that eyes closed alpha peak amplitude was significantly lower in OSA than controls (mean (standard error) = 18.7 (4.97) vs. 34.8 (2.47), p = 0.004). Eyes closed alpha power was also lower in the OSA group than the control group (101.4 (17.68) vs. 171.2 (8.79), p = 0.0006).

Auditory event-related potentials (ERP)

The twelve ERP outcomes were analysed by MANCOVA with group, age, and scalp region as predictors. The effect of group was highly significant ($F_{12,35}$ = 9.93, p < 0.00001). Univariate analyses were performed for the twelve ERP outcomes, using a Bonferroni-adjusted alpha level of 0.004. The results, and the estimated mean values for each group adjusting for age and scalp region, are given in Table 2.4 below.

ERP metric			OSA	Control	Univariate p
Target	P300b	Latency	349.5 (4.0)	360.0 (2.0)	0.03
		Amplitude	5.68 (0.44)	7.53 (0.23)	0.0003*
	N200	Latency	256.2 (2.96)	239.1 (1.52)	<0.0001*
		Amplitude	-0.78 (0.38)	-2.22 (0.19)	0.0015*
	P200	Latency	198.4 (2.12)	186.0 (1.09)	<0.0001*
		Amplitude	4.50 (0.33)	1.85 (0.17)	<0.0001*
	N100	Latency	115.6 (1.39)	118.5 (0.71)	0.08
		Amplitude	-5.56 (0.23)	-4.56 (0.12)	0.0003*
Background	P200	Latency	227.5 (2.82)	227.5 (1.45)	0.72
		Amplitude	2.67 (0.15)	2.41 (0.08)	0.10
	N100	Latency	118.0 (1.39)	119.4 (0.71)	0.36
		Amplitude	-3.70 (0.17)	-3.18 (0.09)	0.013

Table 2.4. Results of Oddball ERP metrics adjusted for age and scalp region.

Mean and standard error values are shown

*Univariate tests below Bonferroni-adjusted significance level of 0.004.

For ERP in response to oddball targets, OSA subjects show significantly reduced P300b amplitude, prolonged N200 latency, reduced N200 amplitude, prolonged P200 target latency, increased P200 target amplitude, and an increased N100 amplitude. P300b target latency was not significantly different between groups. With oddball background stimuli ERP, there were no significant differences between groups after correcting for multiple comparisons.

Cognitive test battery (Cognitive Performance Profile)

Results from the cognitive test battery are shown in Table 2.5. The OSA group showed a significant reduction in performance compared with the Control group with respect to the Verbal Interference and the Switching of Attention (mixed numbers and letters) tests. The OSA group showed a marginally significant reduction in performance in the Reverse Digit Span test.

	OSA	Control	p
Span of visual memory (longest span of jumps remembered correctly)	5.21 (1.04)	5.25 (1.15)	0.9
Timing test: proportional bias	-0.024 (0.17)	-0.071 (0.20)	0.16
Maze trials completed	9.56 (4.27)	9.15 (3.83)	0.56
Maze total number of errors	57.7 (44.1)	58.1 (71.0)	0.99
Choice reaction time, average reaction time	782.4 (189.3)	786.6 (387.3)	0.86
Forward digit span score	6.76 (2.29)	7.10 (2.26)	0.37
Reverse digit span score	3.21 (2.34)	4.02 (2.57)	0.07
Memory Recall Test: Score Trial 7 (delayed recall)	6.35 (2.46)	6.98 (2.02)	0.17
Memory Recall Test: Total Score Trials 1-4	31.0 (5.4)	31.5 (4.7)	0.55
Memory Recognition Test: Recognition Accuracy	10.92 (1.37)	11.14 (1.13)	0.27
Verbal Interference Test: Score (text)	15.43 (3.97)	16.62 (3.52)	0.04
Verbal Interference: Score (colour)	7.84 (4.10)	9.27 (4.26)	0.04
Switching of Attention: time to completion (digits)	26989 (2676)	24963 (9436)	0.22
Switching of Attention: time to completion (mixed)	60475 (19606)	53537 (19155)	0.03

Table 2.5. Cognitive tests: ANCOVA with group, age and years of education as predictors.

2.5 Discussion

In this study, we used a validated diagnostic algorithm to select a group with possible obstructive sleep apnea, and compared them with a control group who were not likely to have OSA. Tests of executive function requiring attention despite interference and set-switching such as the verbal interference (Stroop), switching of attention (Trails B), and reverse digit span tasks were negatively affected by OSA. No between group differences were demonstrated with tasks of simple reaction time (choice reaction time), and short and intermediate term memory. The pattern of cognitive impairment demonstrated is broadly consistent with other studies in OSA, and reflects deficits in the prefrontal cortex. Resting eyes-closed alpha EEG power was reduced in the OSA group. Although P300 latency was not altered, amplitudes of later components (N200, P300) were reduced with OSA, while those of earlier components (N100, P200) were increased.

In contrast to the majority of ERP studies in OSA, the P300 latency was not shown to be prolonged. This delay may emerge in a study sample with unambiguous OSA based on sleep studies (rather than the current group, who were identified on the basis of a screening questionnaire). This group based on this screening for OSA, showed increased N100-P200 amplitude, bur reduced N200-P300 amplitude and delayed P200-N200 latency. One possibility is that the subject's are compensating for their OSA by over-processing the sensory stimuli (reflected in enhanced N100-P200). However, they are not able to fully compensate, and thus show dysfunction in processing task relevant stimuli, reflected in decreased N200-P300. Similar findings have been shown in one OSA study (Rumbach, Krieger, & Kurtz, 1991), and studies of experimental sleep fragmentation (Cote et al., 2003; Kingshott, Cosway, Deary, & Douglas,

2000). Interestingly, the 47 patients with severe OSA (mean apnea-hypopnea index 58.9 events per hour) also demonstrated a reduced N200-P300 peak-to-peak amplitude and increased N100-P200 amplitude relative to 40 age-matched controls, findings that are entirely compatible with our findings. While the P300 latency is thought to reflect the speed of cognitive processing (McCarthy & Donchin, 1981), the reduction in P300 amplitude observed in the OSA group might imply reduced attention and context processing. Another substantiative factor is our finding of reduced resting alpha power in the eyes closed condition in the OSA group, suggesting increased levels of physiological sleepiness even though performance tasks probing vigilance such as the choice reaction time were not abnormal. The three-minute duration of this reaction time task may also not be sufficient to detect subtle performance impairment.

With regard to statistical power, the number of subjects in the OSA group was relatively small, however it represented all subjects in the nearly 2000-subject database meeting the inclusion criteria, and is consistent with the expected 2-4% prevalence of sleep apnea syndrome in the general population (Young et al., 1993). We optimised statistical power by employing a 1:4 ratio of subjects to controls. Conversely, many outcomes were tested and the electrophysiological outcomes also included multiple variables reflecting different conditions of testing and electrode positions. The risk of type I error was mitigated by our selection of a limited number of outcomes from the wide battery that has been acquired for the database, and the application of an overall F test to individual tests with multiple sub-outcomes. In addition, a conservative Bonferroni adjustment of the significance level was applied to these sub-outcomes.

A limitation of the study is the lack of objective polysomnographic data in the diagnosis of obstructive sleep apnea and thus we have no objective measure of disease severity to relate with the outcomes studied. While the diagnostic algorithm has been validated in clinical populations

with a higher background prevalence of OSA, when applied to a community-based normative sample, it may select a higher proportion of subjects without sleep apnea (false positives) or subjects with milder disease. The diagnostic algorithm included the nocturnal breathing symptoms of OSA, age, gender and body mass index, the latter factors representing the key risk factors for the disease. The algorithm does not include sleepiness or other daytime functional symptoms: these would have biased subject selection toward finding effects in the outcomes analysed. Nevertheless, it is apparent from our data that the diagnostic algorithm was able to select a group which exhibited impairments in neuropsychological functioning. Additionally, with the lack of objective measurements of sleep apnea severity it might be considered that the difference in performance between groups might be mediated by the difference in BMI rather than the presence of OSA in the study group. While obesity independent of sleep apnea has been reported to be associated with sleepiness (Vgontzas et al., 1998), the association between neurocognitive performance and obesity has not been well characterised in studies controlling for the presence of OSA, and the strong links between OSA and obesity make this area difficult to investigate.

Another potential implication from our findings relates to the presence of participants within the normative database with obstructive sleep apnea and detectable neurocognitive impairments. While this may affect the interpretation of studies in which clinically sampled patients with OSA are compared with reference values from the normative data set, it is not likely to have a large effect as only a small proportion of this data set (5%) could be classified into the OSA group. Narrower inclusion criteria for the normative sample may make this sample less representative of the population, and reduce the utility of the database in exploring cognitive effects of milder states of pathology.

This study demonstrates the utility of a large neuropsychological outcomes database in the evaluation of hypotheses about the patterns of impairment in various diseases including sleep apnea. Such studies facilitate the planning of subsequent studies designed to both confirm the findings, and to explain the mechanisms of the deficits demonstrated. In regard to our study of sleep apnea, the current study confirms the sensitivity of the testing protocol in evaluating the outcomes of interest, provides data for estimation of the required sample sizes, and prompts the search for mechanisms by which the impairment is mediated. Work remains to be done in clarifying relative contribution of sleep disruption, hypoxia and other factors to the impairment of brain function in obstructive sleep apnea. Studies employing a standardised battery of tests facilitate such work in allowing the aggregation of results across studies, and the comparison of findings between disease groups.

2.6 Conclusion

In summary, a group of subjects predicted to have a high likelihood of OSA by a diagnostic algorithm based on apnea symptoms and demographic factors, demonstrated deficits in executive functioning and cognitive processing in a standardized battery of tests. These results support the utility of databases of such standardized tests in exploring the patterns of cognitive dysfunction, and the availability of large brain function databases facilitates the exploration and gathering of knowledge of possible deficits in various other diseases.

The study supports the notion that the testing battery is sensitive to the effects of OSA, and could be applied in prospective studies in clinical OSA populations.

2.7 Implications

In this comparison of 2 groups of subjects drawn from a normative database, deficits in some aspects of executive function (verbal interference/Stroop, switching of attention (Trails B) and reverse digit span) were identified in the group with likely obstructive sleep apnea. No single variable demonstrated large between group effects. This supports the need for a more broad-based assessment rather than the use of a limited set of tests.

Nevertheless, using a comprehensive testing battery raises statistical concerns regarding multiple comparisons, given the large number of outcome variables and variant metrics for each test. It is imperative to follow a priori plans to limit the number of statistical tests applied: specifying the metrics to be included in analysis, application of omnibus tests such as the MANOVA and analysis methods by which the multiple outcome variables can be aggregated.

EEG power changes and alterations in ERP amplitudes provide encouragement that neurophysiological measures may provide greater sensitivity than standard neurocognitive tests in assaying the injury to the brain associated with OSA. For our intended aim to assess quantitative EEG as a measure of sleepiness, ERP measures should not be included in the outcomes as they are form of EEG measure, and are not independent of the types of measures we wish to test.

3 Sleepiness and function in a clinical OSA population

3.1 Chapter abstract

Aims: (1) Characterize sleepiness and daytime function in a sample presenting to a sleep centre for possible obstructive sleep apnea (OSA). (2) Summarize the components of daytime function into a limited number of outcome variables.

Methods: Cross-sectional study. Subjects presenting to sleep laboratory for suspected OSA were administered polysomnography, resting electroencephalograph (EEG), symptom questionnaires, and a neurocognitive performance battery. Outcome variables of daytime function were aggregated by exploratory factor analysis. Univariate regression and Pearson correlations examined predictors of daytime function: age, gender, BMI, and OSA severity.

Results: N=123, males =101 (82.1%), age 45.1(11.6) years, BMI 32.9(6.4) kg/m^2, AHI 31.2 (25.9). Factor analysis: Thirty-eight outcome variables were entered into the factor analysis. Five factors explained 43% of the variance: Factor 1 "Subjective sleepiness and functional impact", Factor 2 "Mood and anxiety", Factor 3 "Memory and Learning", Factor 4 "Driving", and Factor 5 "Executive functioning". Internal consistency within factors was moderate to good (Cronbach's alpha 0.55-0.90). Demographic variables and measures of OSA severity were not strongly associated with daytime function.

Conclusions: In this sleep centre population, standard measures of OSA severity did not correlate well with daytime function.

3.2 Introduction

Sleepiness is an important outcome of obstructive sleep apnea (OSA). As mentioned in the previous chapter, other functional impairments have been described in relation to OSA, with these domains of impairment including memory, vigilance, and executive functioning closely tied with the concept of sleepiness on several levels, in terms of the sleep disruption from the sleep apnea events, the intermittent hypoxic injury to cortical systems, and the effect of sleep state instability (sleepiness) on an individual's ability to perform in neurocognitive testing. A method for measuring sleepiness in OSA would have clinical applications, including quantifying the impact on daytime function, measuring response to treatment, and perhaps as a guide to priority if resources to commence or maintain a person on therapy were limited.

A population with varying degrees of the phenomenon of sleepiness can be used to evaluate a proposed measure of sleepiness. In the absence of a gold standard measure of this phenomenon, indirect comparisons must be made with other measures of sleepiness, or phenomena that we would expect to be related to sleepiness. Furthermore, with a disease model of sleepiness, the finding that a measure of sleepiness covaried with disease severity would add support for the validity of the tool in quantifying the level of sleepiness.

While there may not be complete congruence between the symptoms and cognitive impairments experienced in OSA and the abstract quality sleepiness, it is a convenient model in which sleepiness can be studied. In this study we examine sleepiness and functional outcomes in a clinical population in which the severity of obstructive sleep apnea has been characterized by polysomnographic testing. The latter is the currently-accepted gold-standard method to diagnose the condition. The functional evaluation will involve the Brain Resource International Database

Battery, in addition to questionnaire and neurocognitive tests commonly employed in sleep and sleep apnea research.

A wide range of tests is recommended to more fully evaluate the functional and neurocognitive impact of sleep apnea (Decary et al., 2000) and also to capture the different nuances of sleepiness (Curcio et al., 2001; Johns, 1998). The use of multiple outcomes leads to a problem of multiple statistical comparisons, and the potential for inflation of the type I error (spuriously finding a significant finding when the finding had occurred due to chance). This problem needs to be overcome. A responsive set of tests for sleepiness should report results in the same direction for many of the sub-outcomes that it includes: there should be consistency of findings. The general effect on a battery of tests can be examined by a total score (if present), or statistical methods to examine all or broad groups of outcome variables simultaneously.

There are statistical methods that could be used when there are multiple related outcome variables to evaluate differences in the combination of outcome variables that might exist between groups. An example of such a technique is the multivariate analysis of variance (MANOVA) in which a single summary dependent variable is generated internally by the procedure, which is then compared between the groups (Pallant, 2004). If such between-group differences are shown to exist, the problem of multiple comparisons must still be overcome if the individual variables are examined to find the ones that accounted for the between-group differences.

Alternatively, similar outcomes might be grouped together into a smaller number of variable sets or domains. Aggregate scores for these domains combine the effects of the component variables, and can be treated as single variables in subsequent analyses. These aggregate scores have the potential to improve the precision in the measurement, by smoothing out fluctuations in scores

due to measurement error. Variables can be grouped together by prior theoretical knowledge about the types of phenomena the individual tests probe, and this process can be guided by statistical methods to confirm the similarities between variables in a domain, and differences between variables belonging to separate domains.

Factor analysis seeks linear combinations of multiple variables, called factors, which represent underlying fundamental quantities of which the observed variables are expressions (Venables & Ripley, 2002). The method can be used to show commonalities between several outcome variables, and composite scores can be generated for a limited number of factors that are orthogonal (independent) of each other, perhaps reflecting distinct nuances of the quality we are measuring. The variables comprising the factors can be examined for what they are thought to measure, and the overall quality that each factor is likely to reflect can be deduced from commonalities in the variables.

Factor analysis is widely used in behavioural research, and in the validation of many questionnaires including instruments widely used in sleep research (Johns, 1991; Weaver et al., 1997). This technique has been used in sleep research to aggregate outcomes from neurocognitive performance tasks and questionnaires (Adams et al., 2001; Frey, Badia, & Wright, 2004; Kraemer et al., 2000), and is the method employed in this study.

Aims

The broad purpose of the study was to provide an improved characterization of the extent of sleepiness, neurocognitive and functional impairment in a clinical population with varying degrees of sleep apnea. Specific aims of this study are:

1. To study the degree of sleepiness and functional impairment using questionnaire and neurocognitive tools, over a broad spread or range (a variety) of domains, within an OSA clinical population.
2. To combine the outcomes by factor analysis, thus reducing the number of outcome variables to a manageable size.
3. To relate the outcomes to PSG measures of severity.

Hypotheses

The hypotheses are that:

1. Comprehensive batteries provide thorough characterization of function,
2. Factor analytic methods improve the detection of consistent variations in functional domains (and hence reducing the risk of type I error), and
3. The chosen battery of tests will be sensitive to detect the effect of OSA.

3.3 Methods

Participants.

Sources of recruitment.

From August 2003 to February 2006, participants presenting to the sleep centre at Royal Prince Alfred Hospital (Camperdown, NSW) with possible sleep apnea were asked to participate.

Inclusion criteria were as follows:

1. Subjects aged 18 to 70 referred to polysomnogram to investigate suspected obstructive sleep apnoea.

3. Male or female.

4. Holding a current driver's licence.

5. Able to give informed consent.

6. Fluent in English (verbal and written) to allow neurocognitive testing.

Exclusion criteria were as follows:

1. History of psychiatric, head injury or neurological disorder: including previous stroke or use of psychotropic medications.

2. Known severe and uncontrolled medical conditions e.g. cardiac failure, hypertension, hypercapnia etc.

3. Binge alcohol exposure or alcohol consumption > 30 g/day.

4. Treatment with medications potentially affecting cognitive testing or EEG recordings such as opiates, benzodiazepines, anti-depressants, anticonvulsants, antipsychotic drugs.

5. Concurrent participation in another research protocol.

6. For participants undergoing the MRI sub study, there were specific exclusion criteria relating to MRI scanning such as recent metal implants, ocular foreign bodies, severe claustrophobia, or excessive weight (scanner bore is 60 cm in diameter).

In order to obtain a representative sample and reduce the risk of selection bias, the initial phase of recruitment was directed at consecutive attendees to the sleep laboratory for diagnostic sleep studies. For this initial consecutive recruitment phase, participants were approached about the study before the result of their polysomnogram was known. As testing usually occurred at the time of their attendance for their PSG, they were usually approached before the PSG was performed. The presence of subjective sleepiness was also not part of the inclusion criteria, so that all degrees of reported sleepiness would be included in the sample.

Practical issues with the need to schedule the components of the tests led to recruitment focusing on those booked for a diagnostic sleep study on Sunday evening, and in some cases Saturday evening if a testing appointment could be made on Sunday morning at the Brain Resource Company. Difficulties in coordinating all tests, or time commitments on the part of the participants, led to omission or postponement of test components in some cases, and the substitution of simpler tests in other cases (for example, the IntegNeuro kiosk battery replaced

the more comprehensive battery of tests undertaken at the Brain Resource Company on some occasions).

After the first 76 participants (at ID number 084), the study population was enriched by including healthy asymptomatic participants without OSA, and participants who already had their diagnostic polysomnogram with known very severe OSA (AHI > 50). This was aimed at obtaining a sample with a broad range of OSA severities, and 50% of recruited participants with AHI ≥ 30. Potential healthy controls were recruited by advertisement and from a volunteer database. They were pre-screened by the Multivariable Apnea Prediction Index (Maislin et al., 1995) and recruited for testing if their predicted probability of OSA was 0.50 or less. Healthy controls were excluded if they were found to have OSA (AHI ≥ 10) on PSG.

Tests

Polysomnography (PSG)

Polysomnography was performed at the sleep unit at Royal Prince Alfred Hospital (Camperdown, NSW) using Compumedics E series acquisition hardware (Compumedics, Melbourne, VIC). The set up was similar to what is routinely used in that laboratory for research (Grunstein, Ho, & Sullivan, 1991) and clinical diagnostic sleep studies, but with the addition of EEG leads in the Fz, Cz, Pz, and Oz regions, and left lateral, right lateral and right supraocular electrooculogram (EOG) leads. All EEG leads were sampled at a rate of 256 Hz, with a high-pass filter applied at 0.3 Hz and a low-pass filter at 50 Hz.

Sleep staging and manual scoring of arousal and respiratory events was performed using standard criteria (American Academy of Sleep Medicine Task Force, 1999; Rechtschaffen & Kales, 1968; Sleep Disorders Atlas Task Force of the American Sleep Disorders Association, 1992). In brief, an apnea was defined by a complete cessation of airflow detected by nasal pressure cannula for 10 seconds or more, and a hypopnea was defined by a *discernable* reduction (with no numerical criteria explicitly stated) from baseline of the amplitude of the nasal pressure signal for 10 seconds or more, and accompanied by either an EEG arousal or an oxygen desaturation of 3% or more.

Waking EEG recording

As part of the preparation for the diagnostic sleep study, the following leads were placed according to the International 10-20 specification (American Electroencephalographic Society, 1991): Fz, Cz, Pz, Oz, C3, C4, A1, A2, Fp1. The ground electrode was placed just below A2. All EEG acquisition at the sleep laboratory used the same Compumedics E series hardware that was used for the polysomnogram.

A five-minute resting recording of the electroencephalogram (EEG) (T. Akerstedt & Gillberg, 1990) was performed using the leads described above (Fz, Cz, Pz, Oz). Subjects were instructed to sit upright in bed and relaxed in a quiet room with the usual ambient indoor lighting. They were asked to fix their gaze at a circular dot placed 1-3 metres away at eye level on a wall, keeping as relaxed as possible while staying awake.

The research assistant started the recording after checking signal integrity, that impedances were below 10 kOhms, and after ensuring that the recording was not contaminated with movement or

electromyography (EMG) artefact. After 2 minutes had elapsed, subjects were instructed to close their eyes, and the recording was terminated after another 2 minutes.

The analysis of the EEG recordings will be described in the following chapter.

EEG recording during performance tasks

Waking EEG recording was also performed concurrently with the Tower of London and the AusEd Driving Task (described below). They were collected to allow analysis of time on task effects on the EEG, and to relate these EEG measures to performance across the task. These recordings were not analysed as part of this study.

Tower of London Task

The Tower of London Task (Krikorian, Bartok, & Gay, 1994; Shallice, 1982) was included in the battery as a test of planning and problem solving ("analysis and synthesis" as termed by Beebe & Gozal, 2002). These skills are components of executive function. It was added to supplement the tasks included within the Brain Resource International Database (BRID) battery, to augment the assessment of executive function.

To our knowledge the test has not been applied to subjects with OSA, however, it has been shown to measure function in the prefrontal lobe, a region thought to be vulnerable to the effects of OSA (Beebe & Gozal, 2002). It was initially shown to be sensitive to damage in the left frontal lobe (Shallice, 1982), and in healthy subjects, performance correlated with left prefrontal activation on single-photon emission computed tomography (number of moves above the minimum, $r = -0.83$, $p < 0.02$) (R. G. Morris, Ahmed, Syed, & Toone, 1993). Task performance (reciprocal of time to completion) has also been shown to correlate with posterior temporal and

parietal EEG asymmetries in healthy volunteers ($r = -0.36$, $p < 0.05$) (Hoptman & Davidson, 1998).

The subject is instructed to generate a sequence of moves that will cause an arrangement of balls on pegs to match a goal arrangement. They are instructed to perform this in as few moves as possible, and as quickly as possible. A sequence of arrangements with varying difficulty is presented. The task probes the prefrontal lobe in its ability to plan and hence problem-solve in a novel situation.

A computerised version of the task was administered (Colorado assessment tests, version 1.2) (Colorado Springs, Colorado, USA). In this implementation, subjects moved coloured balls between pegs with the use of the computer mouse (See Figure 3.1). The software automated the process of administering the test and scoring the results. The software provided written and spoken instructions at the start of the testing session, and feedback was given on performance at the end of each trial. A brief practice session lasting approximately 3 minutes preceded the actual test, which was of 8 to 12 minutes duration. Individual trials from the practice session were not repeated in the actual test.

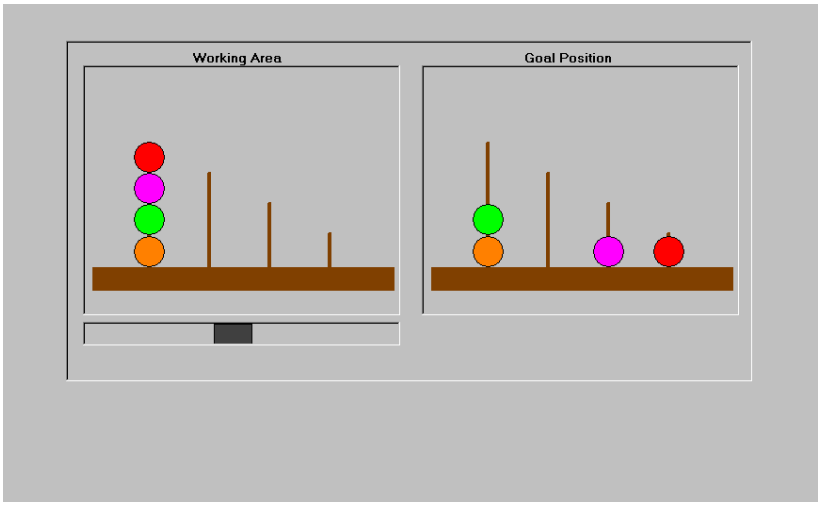


Figure 3.1 Screen display of Tower of London Task

We examined two metrics from the Tower of London task, the total number of moves above the minimum required to solve the problem (*t.excess*), and the average time taken per trial to solve the problem (*t.avtrial*).

AusEd simulator task

The AusEd Driving task (Woolcock Institute of Medical Research, Sydney, Australia) (Desai et al., 2007) simulates driving on a country road at night. After satisfactory understanding of the instructions and controls was achieved on a 5-minute practice run (supervised by the researcher who repeated the practice task if needed), the 30-minute actual test was run.

The test course comprised alternating 2-minute winding (chicane) and 5-minute straight periods. The subjects were asked to drive in the centre of the left hand lane, and to maintain their speed between 60 and 80 km/h. A speedometer was depicted on the top left hand corner of the display. Ten trucks were presented at random intervals during the task. Upon their appearance and as

quickly as possible, the subjects were asked to remove their foot from the accelerator pedal, depress the brake pedal, and then return their foot to the accelerator to continue driving.

Eight outcome variables are reported by the analysis software provided with the driving task. (Table 3.1) Metrics of interest in this task were steering deviation (STDVM), speed deviation (SPDEV), mean reaction time to braking (RTMN), and the number of crashes (CRASH) across the entire 30-minute task, after ignoring data from the first 6 minutes. This was done to reduce the effect of subject acclimatization to the task.

The AusEd task has been shown to be sensitive to impairment in performance associated with sleep loss, circadian effects, alcohol and obstructive sleep apnea. Mean reaction time and speed deviation metrics from the simulator have been shown to be sensitive to the effects of 30 hours sleep deprivation in OSA and healthy subjects (Desai, Marks, Jankelson, & Grunstein, 2006). Weak but significant correlations have been shown between the AHI and Aused performance (r=0.38 for number of crashes, r=0.22 for steering deviation) (Desai et al., 2007). Using a 70-minute form of the test, the addition of alcohol to a mean blood concentration of 0.037 g/dL to one night of sleep restriction to 5 hours time in bed led to a deterioration in steering deviation and mean reaction time to breaks, when compared to the sleep restriction condition alone (S. Banks, Catcheside, Lack, Grunstein, & McEvoy, 2004). In the same experiment, the steering deviation, reaction time to brakes and the number of crashes negatively correlated with sleep latency on a single trial of a 40-minute Maintenance of Wakefulness Test (Siobhan Banks, Catcheside, Lack, Grunstein, & McEvoy, 2005).

Variable name	Description
*a.STDVC**	steering deviation from the centre of left lane
a.STDVM	steering deviation from the median lane position
a.SPDEV	speed deviation outside 60-80 km/h
a.RTMN	mean reaction time (braking in response to trucks)
a.RTSD	standard deviation of reaction time
a.RTMD	median reaction time
a.CRASH	Number of crashes
a.LAPSES	number of reaction times greater than 3.0 seconds

Table 3.1 Outcome variables from AusEd Driving Task

*in this study all variable names for the AusEd task begin with the prefix "*a.*".

Brain Resource International Database Battery

Components

The Brain Resource International Database (BRID) Battery comprises questionnaires, electrophysiological measures, neurocognitive tasks (Cognitive Performance Profile Battery) and magnetic resonance imaging (Clark et al., 2006; E. Gordon et al., 2005; Paul et al., 2005; L. M. Williams et al., 2005). Test outcomes can be referenced to a normative data set of over 2000 healthy volunteers who were drawn from the community. For this study, we looked at the neurocognitive outcomes from the battery. The set of neurocognitive tests in the BRID battery is referred to as the Cognitive Performance Profile.

The questionnaire included in the BRID Battery contains questions relating to demographics, sensory impairment, handedness, learning difficulties, psychiatric history, neurologic history, drug use, emotional intelligence, early life stress and trauma (Brain Resource Company, 2005). It

also includes a modified version of the Depression Anxiety Stress Questionnaire (S. H. Lovibond & Lovibond, 1995), and the Multivariable Apnea Prediction Index (Maislin et al., 1995). For this study, the BRID questionnaire was replaced in favour of a collection of sleep-specific questionnaires exploring symptomatic sleepiness, sleep-related quality of life, mood, anxiety and stress (described below).

Outcomes used from neurocognitive (Cognitive Performance Profile) battery

The Cognitive Performance Profile battery is administered with a touch-screen computer (Clark et al., 2006; Paul et al., 2005). Instructions are played to the subject via a headset, and responses can be made by touching the display or, at some parts of the battery, making verbal responses through a microphone attached to the headset. The total battery requires approximately one hour.

Descriptions of the components have been published (Clark et al., 2006; Paul et al., 2005), and technical aspects of data acquisition and analysis are documented in detail in the testing manual accompanying the battery (Brain Resource Company, 2005). The tests included are briefly described below, and the outcomes included in the analysis are described in the table (Table 3.2, page 97), along with the variable names.

Memory recall and recognition (Rey Auditory Verbal Learning and Memory Task) (Schmidt, 1996). A list of 12 words is read to the subject four times. The subject is asked to recall the words after each reading (recall trials 1-4). A second distracter list is presented that has to be memorised. 20 minutes later, the subject is again asked to recall the original list (delayed recall). Subsequently, the subject is shown a series of words, and asked if they belonged to the original set of words (memory recognition).

Digit Span (Wechsler, 1981). After hearing a sequence of digits (3-9 digits), the subject enters the digits on a numeric keypad on the screen (Forward Digit Span). The Reverse Digit Span is identical, except that the numbers have to be entered in reverse order. These are tests of working memory.

Span of visual memory (Dot Location Task, a variation on the Corsi Blocks Test) (Roth & Crosson, 1985). Nine squares on the display light up in random order. After hearing a tone, the subject is asked to press the squares in the order that they lit up. This is also a test of working memory.

Sustained Attention Task. A series of letters (B, C, D, G) are presented for 200 ms each. The subject is asked to press a button if the same letter appeared twice in a row.

Switching of Attention (Trailmaking Test Parts A and B) (Reitan, 1955). For Part A, the subject is asked to touch a sequence of 25 numbers scattered across the screen, in ascending order. This is a test of working memory. In Part B, the subject touches alternating numbers (1-13) and letters (A-L). The second part probes the executive function skill of set-shifting.

Motor tapping (Finger Tapping test): the subject is asked to tap a circle with the index finger, as many times as possible within thirty seconds. It is used here as a test of manual dexterity (Halstead, 1947), but slowing has been shown with drowsiness accompanying sleep onset (Casagrande, De Gennaro, Violani, Braibanti, & Bertini, 1997). The task is performed with each hand, but data from the dominant hand is included in the analysis.

Choice Reaction Time (Four Choice Reaction Time Test). One of 4 circles on the display lights up, and the subject has to touch that circle as quickly as possible (R. Wilkinson & Houghton, 1975). This is a test of simple reaction time.

Timing (Time Estimation task). A circle appears on the screen for 1-12 seconds. The subject is asked to indicate the duration that the circle was visible for. The test measures visual attention and subjective sense of time intervals, and has been associated with frontal cortical volume (Gunstad, Cohen, Paul, Luyster, & Gordon, 2006).

Verbal interference (Also called Word Interference, based on the Stroop task). First the subject is asked to indicate the colour that the written word spells. Secondly, the subject is asked to name the ink colour that the word is written in. Subject responses are made by pressing one of four words, permanently displayed in black on the bottom of the screen (red, yellow, green and blue) (Stroop, 1935). The second part of the task examines the executive function ability to suppress automatic well-learned responses.

Spot the real word. (Spot the Word Test) (Baddeley, Emslie, & Nimmo-Smith, 1993). The subject is presented with pairs of words on the screen, and asked to pick the word from the non-word. This tests vocabulary and hence is a marker for pre-morbid intelligence.

Word generation: Letter Fluency Test (Controlled Oral Word Association or FAS Test) (Benton & Hamsher, 1989). The participant is asked to recall as many words as possible beginning with each of the letters F, A and S. One minute is allowed for each letter. Responses are recorded by the microphone and then manually scored. This is a test of verbal fluency.

Word generation: animals. In one minute, the participant is asked to recall the names of as many animals as possible starting with a selected letter.

Maze task (Walsh, 1991). The subject is required to discover (by trial and error) a hidden path through a maze and remember it. The task is completed when the subject traces the path

through the maze correctly twice in a row, or after 8 minutes have elapsed. This task assesses high level mental functions such as planning, foresight and self-monitoring. The naïve presentation of such a task is associated with frontal lobe activation (Van Horn et al., 1998).

Table 3.2 below is an excerpt from the testing manual describing the tasks from the Cognitive Performance Profile. Main outcomes from each of the tasks were those recommended by the manual.

Test	**Analogous test**	**Metric**	**BRID variable name**	**Our variable name**	**Notes**
Memory Recall	Rey Auditory Verbal Learning and Memory Task	Total Score Trials 1-4	*memtot14*	*b.memtot14**	Total number of words recalled over trials 1-4.
Memory Recall		Score trial 7 (delayed recall)	*memrec7*	*b.memrec7*	Number of words recalled after the trial at 25 minutes.
Memory Recognition		Recognition accuracy	*memrecco*	*b.memrecco*	Words correctly identified as from original set, when prompted.
(Forward) Digit Span	Digit Span – WAIS III	Score (forward)	*digitot*	*b.digitot*	Longest sequence correctly completed.
Reverse Digit Span		Score (backward)	*rdigitot*	*b.rdigitot*	Longest sequence correctly completed
Span of visual memory	Dot Location Task	Score	*Spvm*	*b.spvm*	Longest sequence (span of jumps) correctly completed

Test	Analogous test	Metric	BRID variable name	Our variable name	Notes
Sustained attention (also called Working Memory)		False alarm rate; false positives.	*Wmfp*	*b.wmfp*	Number of incorrect responses
Sustained attention		Missed targets; false negatives.	*wmfn*	*b.wmfn*	Number of targets that the subject did not respond to.
Switching of attention	Trailmaking test A	Time to completion (digits)	*Swoadur1*	*b.swoadur1*	Time to complete the test successfully.
Switching of attention	Trailmaking test B	Time to completion (mixed)	*Swoadur2*	*b.swoadur2*	Time to complete the test successfully.
Motor Tapping	Finger tapping test	Number of taps (dominant hand)	*tapdomn*	*b.tapdomn*	Number of taps over 30 seconds.
Reaction time	Four Choice Reaction Time Test	Choice reaction time	*Ch_avrt*	*b.ch.avrt*	Average speed of response.
Timing	Time Estimation Task	Proportional bias	*t_prbias*	*b.t.prbias*	The value of the average difference between the actual length of the stimulus (ls) and the subjects estimate(lu) weighted by the length of the stimulus.
Verbal interference	Stroop Test	Score (text)	*Vi_sco1*	*b.vi.sco1*	Number of words correctly identified.
Verbal interference	Stroop Test	Score (colour)	*Vi_sco2*	*b.vi.sco2*	Number of colours correctly identified.
Spot the real word	Spot the Word Test	Score	*spotscor*	*b.spotscor*	Number of words correctly recognised.
Word generation	Controlled Oral Word Association	FAS score	*fas*	*b.fas*	Number of words recalled across the 3 letters.

Test	Analogous test	Metric	BRID variable name	Our variable name	Notes
Word generation	Animal Fluency	Animal score	*Animal*	*b.animal*	Number of words recalled.
Maze		Trials completed	*mazetrls*	*b.mazetrls*	Number of trials completed before the end of the task.
Maze		Number of errors	*mazeerr*	*b.mazeerr*	Total number of off-path moves.

Table 3.2 Variables from Cognitive Performance Profile Battery

Procedure for BRID Battery

Wherever possible, the full battery of tests (including electrophysiology and questionnaires) was administered. This was conducted during an approximately 3-hour session at the Brain Resource Company (Ultimo, NSW), usually on the morning following the polysomnogram. During periods when a full testing appointment could not be made (either from lack of availability of appointments or the participant being unable to attend the session) the Cognitive Performance Profile alone was administered in the sleep unit at a time convenient to the participant. This Cognitive Performance Profile component required approximately 60 minutes to perform on a computer workstation (IntegNeuro Kiosk, Brain Resource Company, Ultimo, NSW) that was identical to the hardware used when the full battery was acquired. Sixty seven participants underwent the full battery of tests, and 50 had the Cognitive Performance Profile alone.

Questionnaires used

Questionnaires relating to sleepiness and its impact, and potential cofactors that might influence symptoms of sleepiness were included in the study.

Epworth Sleepiness Scale (ESS)

Self-reported sleepiness was rated with the Epworth Sleepiness Scale (ESS), an 8-item scale extensively used in sleep research that has been validated in healthy volunteers and clinical sleep cohorts including sleep apnea (Johns, 1991). The subject was asked to indicate the likelihood of dozing during each of 8 situations, with responses being scored 0 (would never doze) to 3 (high chance of dozing). The total score (out of 24) has been shown to comprise a single factor, the propensity to fall asleep. The measure has been shown to be sensitive to CPAP treatment of moderate and severe OSA (Marshall et al., 2006; Patel, White, Malhotra, Stanchina, & Ayas, 2003), but shows weak correlations with other measures of sleep propensity such as the MWT in narcolepsy (r = -0.29) (R. Bart Sangal et al., 1999) and the MSLT in OSA (r = -0.03, r = -0.15) (Chervin & Aldrich, 1999; Olson, Cole, & Ambrogetti, 1998). The variable included in the analysis was the ESS total score (*ESS*).

Functional Outcomes of Sleep Questionnaire (FOSQ)

The Functional Outcomes of Sleep Questionnaire (FOSQ) (Weaver et al., 1997) explores the extent to which sleepiness or sleep disruption affects five aspects of daily activities: general productivity (items 1-4 and 8-11), social outcome (items 12 & 13), activity level (items 5, 14-16, 22-26), vigilance (items 6, 7, 17-21), and intimate relationships and sexual activity (items 27-30). The questionnaire states in its instructions to the respondent that "the purpose of this questionnaire is to find out if you generally have difficulty carrying out certain activities because you are too sleepy or tired". The respondent was asked to indicate 1-4 (depending on item, smaller values represent greater difficulty) the degree of difficulty experienced in performing the

named activities, and allowed to indicate if they did not engage in the activity described (response treated as missing).

The September 1996 version of the questionnaire was used. The arithmetic mean of the non-missing responses formed the subscale results (potential range 1-4), and the total score sum of the subscale scores (range 5-20).

Previous experience at our centre with the instrument has suggested that there was likely to be many missing responses to the items on the intimate relationships subscale of the questionnaire. This subscale has also been shown to be less responsive to treatment of OSA than the other subscales of the questionnaire (Terri E. Weaver, 2001). The individual subscale scores will be included in the analysis (*Fosq.genprod, Fosq.social, Fosq.activity, Fosq.vigilance*), excluding the intimate relationships subscale score.

Depression Anxiety Stress Scales (DASS)

The 21-item version of the Depression Anxiety Stress Scales (DASS) was included as a self-reported measure of depression, anxiety and stress (Henry & Crawford, 2005; S. H. Lovibond & Lovibond, 1995). The instrument is widely used, as it has been placed in the public domain.

The Depression scale assesses dysphoria, hopelessness, devaluation of life, self-deprecation, lack of interest/involvement, anhedonia, and inertia. The Anxiety scale assesses autonomic arousal, skeletal muscle effects, situational anxiety, and subjective experience of anxious affect. The Stress scale is sensitive to levels of chronic non-specific arousal. It assesses difficulty relaxing, nervous arousal, and being easily upset/agitated, irritable/over-reactive and impatient (P. Lovibond, 2007).

Subjects are asked to use 4-point severity/frequency scales to rate the extent to which they have experienced each state over the past week. The 21-item version of the scale comprises 7 items per scale. Scores for Depression (items 3, 5, 10, 13, 16, 17, and 21), Anxiety (items 2, 4, 7, 9, 15, 19, and 20) and Stress (items 1, 6, 8, 11, 12, 14, and 18) are calculated by summing the scores for the relevant items, and multiplying by 2 to give a range of (0-42) comparable to the standard 42-item questionnaire.

The questionnaire was included to collect information on mood states as they may be affected by the effect of OSA, and also as a potential confounder for symptoms of sleepiness and fatigue experienced by the study sample. The individual subscales were included in the analysis (*Dass.depression, Dass.anxiety, Dass.stress*).

State-Trait Anxiety Inventory (STAI)

This instrument comprises 20 items to measure state anxiety (that component that arises from the appraisal of threatening demands or dangers, *Stai.s*), and a further 20 items concerning trait anxiety (reflecting the stable individual tendency to respond with state anxiety, *Stai.t*) (Spielberger, Gorusch, & Lushene, 1970).

Karolinska Sleepiness Scale (KSS)

The Karolinska Sleepiness Scale (KSS) is a single item Likert-type scale that asks the respondent to indicate how he/she feels at that moment (T. Akerstedt & Gillberg, 1990). There are nine possible responses, from “1: very alert” to “9: very sleepy; great effort to keep awake; fighting sleep”.

It is a commonly used instrument to indicate state (as opposed to trait) sleepiness: the phenomenon of sleepiness that may vary with the immediate, short-term, situation. The KSS was not included at the commencement of recruitment, but included later to allow comparison between subjective state sleepiness and the EEG measures. This variable was not included in the analysis.

Other tests

Subsets of participants from the general cross-sectional study also underwent proton magnetic resonance spectroscopy, near-infra-red spectroscopy concurrently with the PSG, and blood collection to identify markers of metabolic syndrome, and to store for later genetic analyses. They form part of sub-protocols, and will not be described further as the data will not be analysed in this study.

Procedure

Screening and recruitment

Potential participants who were booked for a diagnostic sleep study were contacted by telephone prior to the date of their diagnostic PSG. The study rationale and procedures were explained, and eligibility for the study determined. Those verbally consenting to the study were asked to attend their appointment for the PSG at the usual time.

The researcher made bookings for the appointments at the Brain Resource Company, and in a subset, MRI scanning.

Laboratory visit

Testing was performed on the evening prior to the diagnostic sleep study, and the morning following the study. Participants were asked to refrain from caffeine from 09:00 on the morning of the sleep study, until all testing was completed. They arrived at the sleep unit at 16:30. This was the routine time at which patients who were booked for a sleep study were advised to arrive.

The study procedure (explained prior to their arrival to the laboratory) was explained again, and written consent was obtained. Specifics of the testing schedule were described.

Sensors used in the PSG were attached on the subject, with the addition of extra EEG leads as described above. After checking the quality of the signals, the waking EEG recording was commenced (5 minutes), followed by the Tower of London Task (3 minute practice run and 8-12 minute test) and the AusEd Driving Task (5 minute practice run and 30 minute test).

After the testing, dinner was provided (routine for the sleep unit). The participant was asked to complete a booklet of questionnaires.

The PSG recording commenced at approximately 21:30 and terminates at 06:00. After breakfast, they were asked to attend MRI scanning at the Prince of Wales Medical Research Institute (07:30 – 08:30), and neurocognitive testing at the Brain Resource Company (09:00-12:00).

Transport between testing locations was arranged by taxi, and all participants were provided with detailed directions and maps.

Quality

To ensure standardization of procedures, testing was conducted by the project officer (KW) and 4 research assistants who were trained in the testing procedures, and also worked as sleep

technologists who prepared the subjects for their PSG. Study procedures were documented in a printed manual, and source notes documented any planned variations in the protocol, and were used during data collection to document any issues.

Variations in procedure

When an appointment for testing at the Brain Resource Company could not be arranged, the participants underwent testing on the Cognitive Performance Profile at the sleep unit. This was usually performed on the morning following the sleep study.

MRI scanning was performed in a subset of the patients. The scanner became available for the study in September 2004 (from participant ID 060), and there were periods of downtime required for hardware and software upgrades. Some subjects were not able to be scheduled when the scanner was available owing to contraindications to scanning such as the presence of ocular or cranial metallic foreign bodies, severe claustrophobia, or excessive weight.

Sources of funding, ethics approval, and trial registration

The study was funded by a project grant from the National Health Medical Research Council (253792). The procedures were approved by the ethics committees of the University of Sydney (3616), the Sydney South West Area Health Service (Eastern Zone) (X02-0334, X06-0299), Northern Sydney Health (0402-046M), and the Prince of Wales Medical Research Institute (2003.1). The study was registered with the Australian Clinical Trials Registry as an observational study (ACTRN012605000089639).

Data analysis

Statistical analysis methods

Descriptive data

Demographic data and general characteristics of the study population were described. Means and standard deviations were used to describe continuous data, and numbers and percentages used to describe categorical data.

Factor analysis

Exploratory factor analysis was performed with factor extraction by maximum likelihood, and with variance maximization (varimax) rotation, using the *factanal* function in R (R Development Core Team, 2007).

The outcome variables (see Table 3.3 above) were standardised so that they were denominated in the same scale. Variables from the BRID Cognitive Performance Profile were standardised against reference values from the age, gender and education-level matched normative sample: the difference between the raw scores and the normal values was divided by the standard deviation of the score in the normative sample. Other variables were standardised by subtracting raw scores from the mean value for the overall sample, and dividing this by the sample standard deviation.

Subjects with incomplete data for the 38 variables of interest were excluded from the factor analysis (n=50 available). Models with 1 to 7 factors were explored. The number of factors to include in the model was determined primarily on the examination of a scree plot for inflections,

on the interpretability of the factor structure, and thirdly, on the recommendation that eigenvalues above 1 should be retained (Kaiser's criterion).

Variable names			
ESS	*a.stdvc*	*b.memtot14*	*b.ch.avtr*
Fosq.genprod	*a.stdvm*	*b.memrec7*	*b.vi.sco1*
Fosq.social	*a.spdev*	*b.memrecco*	*b.vi.sco2*
Fosq.activity	*a.rtmn*	*b.digitot*	*b.spotscor*
Fosq.vigilance	*a.rtsd*	*b.rdigitot*	*b.fas*
Dass.depression	*a.rtmd*	*b.spvm*	*b.animal*
Dass.anxiety	*a.crash*	*b.wmfp*	*b.mazetrls*
Dass.stress	*a.lapses*	*b.wmfn*	*b.mazeerr*
Stai.s	*t.excess*	*b.swoadur*	
Stai.t	*t.avtrial*	*b.tapdomn*	

Table 3.3 Outcome variables included in factor analysis

Loadings for individual variables on each of the factors were examined, and variables loading strongly to the factors were grouped to them. Where there were multiple methods of calculating the same score (egg mean and median reaction time, prior knowledge of the better metric or the

one with the highest loading was chosen). Aggregate scores for each of the factors were calculated as the arithmetic mean of the standardised scores for each component variable. To allow addition of the scores within factors, the sign of all scores were adjusted such that a more positive score implied better performance, improved mood, or less sleepiness.

Internal consistency of items within each factor was examined with a Pearson's correlation matrix of items comprising the factor, and Cronbach's alpha, with the aid of a function written in R to perform the calculations and print the results (Magill, 2001). Correlations between factors was examined with Pearson's *r*.

Relation of outcomes grouped into factors with measures of OSA

The derived outcome factors were individually analysed as outcomes in univariate (Pearson's correlation coefficient *r* reported) and multivariate linear regression, with demographic/biometric, PSG and EEG variables as predictors.

3.4 Results

Study population

From August 2003 to November 2005, 123 participants were recruited for the study. The participant characteristics are shown in Table 3.4 below.

Characteristic	Mean (SD)	N (%)
Males		101 (82.1%)
Age, years	45.1 (11.6)	
BMI, kg/m^2	32.9 (6.4)	
AHI, events/hour	31.2 (25.9)	
< 5 events/hour		12 (9.8%)
5-15 events/hour		31 (25.4%)
15-30 events/hour		28 (23.0%)
30-60 events/hour		31(25.4%)
>60		20 (16.4%)
Arousal index, events/hour	30.2 (22.4)	
Minimum Saturation, %	81.6 (12.1)	
Proportion of TST below 90% saturation, %	7.4 (16.8)	
Sleep Efficiency, %	81.5 (12.5)	
ESS	10.5 (4.9)	
>= 10, N (%)		66 (57.4%)

Table 3.4 Characteristics of study population

Seventy-two participants had waking EEG collected. Of these, 70 had questionnaire, AusEd & Tower of London and cognitive performance profile data. 51/70 went to BRC for testing with the full BRID Battery, 19 had IntegNeuro kiosk testing (Cognitive Performance Profile component of the BRID Battery) in the sleep laboratory. The distribution of OSA severity among the group that had the IntegNeuro kiosk testing (13%, 28%, 22%, 37% for AHI 0-5, 5-15, 15-30, 30 or

higher) was similar to the group that had the full testing (11%, 19%, 30%, 41% for AHI 0-5, 5-15, 15-30, 30 or higher).

The study sample included participants with a wide range of OSA severity, including some with AHI values in the normal range (<5 events/hour), and 41.8% of participants with severe OSA (AHI 30 events/hour or higher). Fifty-seven percent of participants reported excessive sleepiness on the Epworth Sleepiness Scale (ESS 10 or greater).

Results of Factor Analysis

Fifty participants (AHI mean (SD) 29.5 (26.5)) had complete data for the 38 variables considered for factor analysis. Factor analytic models with 1, 2, 3, 4, 5, 6 and 7 models were considered (Table 3.5). Chi-squared statistics provide a measure of how well the model fits the data, with high values implying a poorer fit to the data. Figure 3.2 below shows a scree plot of eigenvalues (the variance in a set of variables explained by a factor or component, i.e. the sum of squared factor loadings) from the model containing 7 factors. Although eigenvalues remained above 1.0 with up to 7 factors in the model, there was an inflection in the plot suggesting the optimal number of factors to be five. Factor loadings (see below) for the models with 4, 5 and 6 factors were also examined to confirm that the model with 5 factors was the most readily interpretable.

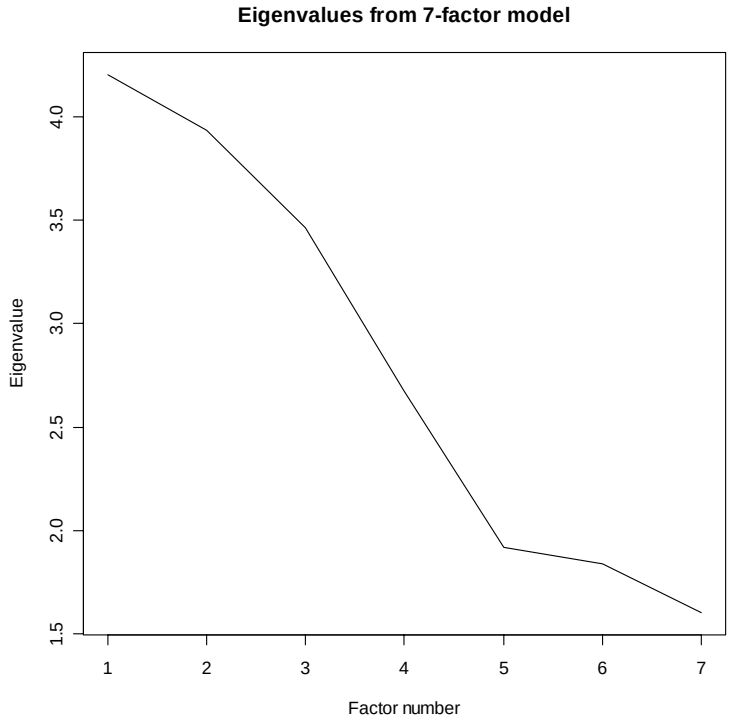


Figure 3.2 Scree Plot from Factor Analytic Model

Model	No. of factors	Proportion of variance	χ^2	df	p
1	1	0.165	1053.1	665	<0.0001
2	2	0.265	931.23	628	<0.0001
3	3	0.331	789.14	592	<0.0001
4	4	0.388	680.92	557	0.0002
5	5	0.432	591.17	523	0.02
6	6	0.479	528.45	490	0.11
7	7	0.517	476.3	458	0.27

Table 3.5 Statistics for factor analytic models containing 1-7 factors

The model with five factors explained 43% of total variance in data. Factor loadings, sorted in descending order within each of the five factors, is shown in Table 3.6 below. The loadings

represent correlation coefficients of each item with the factor. Loadings above 0.40-0.50 are usually considered high, and values below 0.30 are considered low. Loadings below 0.10 have been suppressed, and the list of variables has been sorted according to the magnitude of loadings and the five factors. Crossloading, where an item has loading greater than 0.32 on two or more factors (Costello & Osborne, 2005), does not appear excessive. It is seen with the components of the DASS questionnaire and memory recall over the first 4 trials.

Loadings:

	Factor1	Factor2	Factor3	Factor4	Factor5
ESS	**-0.645**	0.175	0.190		
fosq.genprod	**0.845**	-0.108			
fosq.social	**0.775**	-0.252		0.151	
fosq.activity	**0.885**	-0.157			
fosq.vigilance	**0.881**				
dass.depression	-0.405	**0.805**	0.116	-0.188	
dass.anxiety	-0.354	**0.591**			
dass.stress	-0.376	**0.640**			
stai.s	-0.273	**0.710**	-0.234		
stai.t	-0.378	**0.755**		-0.123	
b.memtot14z	0.269	**-0.522**	0.427		-0.365
b.wmfpz	0.140	**-0.625**	0.138		-0.222
b.memrec7z		-0.190	**0.710**		-0.117
b.memreccoz		0.174	**0.540**	-0.122	-0.116
b.swoadur2z		-0.119	**0.568**	-0.117	
b.spotscorz			**0.807**		
b.fasz	-0.121		**0.516**		0.148
b.animalz			**0.692**		0.190
a.rtmn	0.149		-0.117	**0.974**	
a.rtmd	0.170			**0.968**	-0.133
a.rtsd			-0.103	**0.234**	0.888
a.stdvc	0.107	-0.254		**0.327**	
a.stdvm	-0.113	-0.111		**0.359**	
a.spdev	-0.196	-0.283	-0.108	**0.378**	
a.crash	-0.128		-0.225	**0.370**	0.174
a.lapses	0.103		0.106	**0.274**	0.228
t.excess			-0.350	-0.107	**0.298**
t.avtrial	0.157	-0.144	-0.113		
b.digitotz			0.380	-0.232	
b.rdigitotz		0.267			
b.spvmz	-0.204			-0.291	**-0.146**
b.wmfnz	-0.135			-0.214	**-0.197**
b.tapdomnz			0.260		
b.ch.avrtz		0.249		-0.225	**-0.113**
b.vi.sco1z	-0.271		0.335	-0.113	**-0.157**
b.vi.sco2z		0.264	0.365	-0.128	**0.392**
b.mazetrlsz		0.240			**-0.173**
b.mazeerrz	0.172	-0.113			

Table 3.6. Factor loadings for five-factor model

Table 3.7 below gives the communalities (h^2) for the individual variables entered in the 5-factor model. The communality is the proportion of a variable's variance that is explained by a factor structure. Item communalities are often between 0.40 and 0.70 in social science research, and values below 0.40 may suggest that the factor may not be related to the other items, or that an additional factor should be considered (Costello & Osborne, 2005).

ESS	0.488	t.avtrial	0.065
fosq.genprod	0.735	b.memtot14z	0.663
fosq.social	0.694	b.memrec7z	0.557
fosq.activity	0.809	b.memreccoz	0.350
fosq.vigilance	0.791	b.digitotz	0.206
dass.depression	0.861	b.rdigitotz	0.085
dass.anxiety	0.493	b.spvmz	0.152
dass.stress	0.558	b.wmfpz	0.485
stai.s	0.637	b.wmfnz	0.110
stai.t	0.740	b.swoadur2z	0.357
a.stdvc	0.194	b.tapdomnz	0.084
a.stdvm	0.157	b.ch.avrtz	0.134
a.spdev	0.280	b.vi.sco1z	0.224
a.rtmn	0.995	b.vi.sco2z	0.373
a.rtsd	0.863	b.spotscorz	0.662
a.rtmd	0.995	b.fasz	0.317
a.crash	0.234	b.animalz	0.532
a.lapses	0.149	b.mazetrlsz	0.095
t.excess	0.224	b.mazeerrz	0.053

Table 3.7 Communalities for individual variables in 5-factor model

Factor 1

The variables from the Functional Outcomes of Sleep Questionnaire and the Epworth Sleepiness Scale, both directly probing sleepiness and daytime consequences of sleepiness, loaded strongly with the **first factor**. The first factor was termed “Subjective sleepiness and functional impact”, and comprised the total score of the ESS, and the individual subscales of the FOSQ. It accounted for 11.6% of total variance.

Components of the DASS questionnaire show some cross-loading with this factor, however they loaded more strongly to the second factor, and were expected to probe a similar construct to the State-Trait Anxiety Inventory.

Cronbach's alpha reflects the internal consistency of the variables within a factor, and is a measure of how well each individual item in a factor correlates with the sum of the remaining items. Cronbach's alpha for factor 1 is 0.55. With the ESS variable removed, alpha was high at 0.89, as the factor then consisted of 4 subscales of a single questionnaire (FOSQ). For questionnaire instruments the ideal value for the alpha lies between 0.70 and 0.90, with very high values suggesting the presence of item redundancy (Streiner & Norman, 1989). Although Cronbach's alpha suggests the internal consistency of the variables in factor 1 to be less than ideal, this factor was retained as the combination of the two symptomatic measures of sleepiness made sense, and the factor explained the greatest proportion of the variance in the data.

Factor 2

The components of the Depression Anxiety Stress Scale (DASS) and the Spielberger State Trait Anxiety Index (STAI) loaded strongly on the second factor. Memory recall after 4 trials (intermediate memory) and false positives on the working memory task also loaded moderately strongly on this factor, but they also loaded moderately to factors 3 and 5 respectively. These two variables were moved to factors 3 and 5. The average of the non-missing values for the standardised component scores for the DASS and STAI questionnaires comprised the **second factor**, "Mood and anxiety". This factor accounted for 10.2% of the total variance. Cronbach's alpha was 0.89, indicating high internal consistency for factor 2.

Factor 3

Factor 3 contained variables from the BRID battery relating to memory recall and recognition, and vocabulary (word recognition and word generation). The factor accounted for 9.4% of variance, and was termed "Memory and Learning". Internal consistency was moderate within this factor (Cronbach's alpha = 0.69).

Factor 4

The outcome variables from the AusEd Task loaded to **factor 4**. This factor "Driving" accounted for 7.8% of variance. Many outcome variables on this task are similar. The mean reaction time to brakes (*a.rtmn*) was chosen in favour of the median reaction time, and the Steering deviation from the median lane position chosen instead of the steering deviation from the central lane position. There was moderate internal consistency for the items in this factor (Cronbach's alpha = 0.69).

Factor 5

Factor 5 accounted for 4.2% of total variance, but contained tests relevant to executive functioning: problem solving (the Tower of London, the maze task), behavioural inhibition (the visual interference or Stroop tasks) visual and verbal working memory. Internal consistency for this factor was low (Cronbach's alpha 0.55).

Formation of five summary outcome variables

Table 3.8 summarises the items comprising the five summary variables, the proportion of variance in the data explained by the factors, and the internal consistency of items within the

factors. The mean of non-missing values of the standardised variables were used to generate aggregate scores for the five summary factors. The sign of the standardised score was adjusted such that a more positive score implied better performance, or less sleepiness.

Factor	Component variables	Proportion of variance	Cronbach's alpha within factor
1. Subjective sleepiness and functional impact	ESS and FOSQ questionnaires	0.116	0.553
2. Mood and anxiety	DASS and STAI inventory components	0.102	0.897
3. Memory & learning	Delayed memory recall, Memory recognition accuracy, spot the real word and animal naming	0.090	0.690
4. Driving	AusEd Driving Task components	0.072	0.691
5. Executive functioning	Tower of London number of trials, Digit span, Reverse digit span, Span of visual memory, Working memory false negatives, Verbal interference (Stroop), Switching of attention (Trails B), and Choice reaction time.	0.062	0.550

Table 3.8 Composition of the five factors derived from factor analysis

Table 3.9 below shows Pearson's correlations between the factor scores within the study sample. Low correlations between factors are desired, given the factor analysis aims to yield factors that are orthogonal (independent) of each other. There are moderate to high correlations between the factors 1 and 2. The two factors are comprised of questionnaire data and it is not surprising to find correlations between responses to subjective instruments. There are low to moderate correlations between factors 3, 4 and 5, the factors which comprise tasks from the battery of neurocognitive tests used.

	Factor 2	Factor 3	Factor 4	Factor 5
Factor 1	0.60***	-0.02	0.01	-0.18
Factor 2		-0.02	-0.08	-0.12
Factor 3			0.30**	0.26
Factor 4				0.35**

Table 3.9 Correlations between factor scores

** $p < 0.01$, *** $p < 0.001$

Relating factor scores to sleep apnea severity and demographic factors

To examine if demographic factors or sleep apnea severity predict daytime symptomatic and performance outcomes, the factor scores generated in the previous section were related to the demographic variables such as age, body mass index, and with conventional polysomnographic measures (total respiratory disturbance index, minimum saturation, the proportion of sleep time spent below 90% saturation, the arousal index, and sleep efficiency). A correlation matrix is presented below (Table 3.10).

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Age	0.00	0.09	0.00	-0.09	-0.01
BMI	-0.04	0.07	0.03	0.12	0.05
totRDI	-0.17	-0.14	0.08	0.16	0.25*
MinSat	0.23*	0.09	0.03	-0.06	-0.20*
Proptimebelow90	-0.15	-0.12	-0.04	0.04	0.20
AI	-0.07	-0.13	0.04	0.15	0.26*
SleepEfficiency	-0.07	0.03	-0.02	0.05	-0.05

Table 3.10 Correlations between demographic and PSG variables, and factor outcomes

* $p < 0.05$

Minimum saturations correlated weakly with symptomatic sleepiness (factor 1, r = 0.23, p = 0.02, n = 106). It should be noted that weak correlations were also demonstrated between PSG variables and executive function (factor 5), with the direction of the effect unexpectedly implying better performance in executive function tests with higher severity of OSA. Arousal index (AI, r = 0.26, p = 0.04, n = 63) and total RDI (totRDI, r = 0.25, p = 0.04, n = 66) correlated weakly with factor 5. Minimum saturation also correlated weakly with factor 5 (r = - 0.20, p = 0.11, n = 65). These findings were unchanged after correcting for age with multiple linear regression.

3.5 Discussion

With the assistance of exploratory factor analysis and background theoretical knowledge about the questionnaires and tests, we were able to group the variables into domains that had a sensible theoretical interpretation, and hence reduce the number of outcome variables from 38 to five. This study has used multiple outcomes from a comprehensive battery in which multiple domains of cognitive function have been probed (Decary et al., 2000). Not surprisingly, given the often conflicting findings in the literature, we did not show strong and consistent relationships between our summary factors of neurocognitive function and measures of sleep apnea severity (Aloia et al., 2004).

The main aim of this study was not to explore relationships between conventional measures of OSA severity and daytime outcomes, but rather to aggregate the multiple outcome variables collected from the study sample, and to obtain a broad assessment of the aspects of function relevant to sleepiness in obstructive sleep apnea, in order to test a proposed measure of sleepiness. Nevertheless, our study sample was sufficient to detect small to moderate correlations: 60 subjects would have been sufficient to detect correlations r=0.35 with 80% power and a two-sided significance level of 0.05. After accounting for missing data, pairwise comparisons between PSG indices and the summary variables in our sample have involved more than 60 subjects. Internal consistency within the factors was moderate to good. Correlations between individual items were moderately high given the different circumstances probed by these scales, especially given the mean correlation for sleep latencies between test sessions of the MSLT or MWT has been reported as 0.61 (R. B. Sangal et al., 1992).

The factor analysis has grouped a moderate collection of questionnaires and tests into five domains representing outcomes relevant to patients with OSA. There may be a clinical application in summarising results for OSA patients undergoing comprehensive testing, and may reduce unwarranted concern induced by isolated metrics falling out of the reference range. In research, these factor outcomes may provide a common reference by which to compare results from different studies employing different assessment batteries. The technique of factor analysis provides a means to verify theoretical assumptions on the domains of cognitive function assessed by the individual tests in the population studied, and serve as a useful means to reduce the number of outcome variables that need to be considered in statistical analysis.

There are several issues regarding this study that bear discussion: the small sample size, the number of factors chosen, the potential for type I error and the decision to aggregate the variables rather than considering them individually. These will be discussed below.

The first issue was in regard to the validity of the factor analysis given the small sample size. Owing to the presence of missing data, the sample size of 50 available for the factor analysis was much lower than the 380 recommended subjects for our 38-item analysis, by an often cited 'rule of thumb' of unknown origin. A more rigorous review of robust techniques for factor analysis (Costello & Osborne, 2005) examined the likelihood of obtaining correct factor structures with subsamples of a large data set (n=24599). They found that even with item to subject ratios of 20:1, the correct factor solutions were produced only 70% of the time. Unfortunately, such sample sizes would be an order of magnitude greater than achieved in our study, and indeed greater than the vast majority of studies completed in sleep apnea research. Otherwise, our methods were in accordance with the recommendations of the review with regard to the factor extraction method and the choice of number of factors to retain (the model with the 'cleanest'

factor structure, item loadings above 0.30, no or few item crossloadings, and no factors with fewer than 3 items). Ultimately, the review authors emphasised that the nature of exploratory factor analysis is exploratory, and that theories and hypotheses need to be tested with subsequent studies. Our intention to correlate these aggregated outcome variables with other measures that were distinct from these variables may contribute to the validation of our methods, and any errors made in variable aggregation would be expected to bias against finding a relationship. The study sample used in factor analysis could have been enriched by the addition of data from the normative dataset from the Brain Resource International Database. Similar measures have been adopted by others (Adams et al., 2001). The dataset contains testing information from more than 2000 volunteers, and the factor analysis could have been used to aggregate the variables. To our knowledge, a factor analysis has not yet been performed on the items from the BRID normative data set. Such an analysis would guide the aggregation of the variables from that battery, but we would not have comparable information on the questionnaire and test outcomes specific to our study. There was also a concern that incorrect factor solutions might arise when a pathological sample is mixed with a normal sample: two variables sharing a high degree of variance in normal participants and thus appearing to measure a single construct can dissociate into two distinct constructs, but only in certain homogeneous patient populations (Delis, Jacobson, Bondi, Hamilton, & Salmon, 2003).

The five factors derived in our analysis have a sensible interpretation, but they represent a coarser classification of the domains of cognitive function. For example, there was no further delineation of the subtypes of executive function. The number of factors was kept intentionally small to minimise the number of outcome variables that would be examined in our subsequent study. In the case of executive function, many tests probe more than one aspect, and the

relatively small number of tests of executive function might lead to factors containing only one or two items. Others have aggregated a similar collection of tests into four (Adams et al., 2001; Kraemer et al., 2000) to seven (Frey et al., 2004) factors to investigate relationships among subjective measures and performance tasks in subjects exposed to sleep apnea, sleep deprivation or time-of-day effects.

With five factors as outcomes and many predictors of interest, the problem of multiple comparisons persists. We did not correct the significance level for multiple comparisons, and it is possible that any significant findings in this study could arise from an inflation of the type I error. Indeed this could explain the unexpected relationship between PSG variables and executive function, implying that more severe sleep apnea was associated with improved function. For subsequent analyses on these summary outcomes, careful consideration needs to be made to limit the number of potential predictors to those that are of interest, and any findings need to be confirmed prospectively with a new sample of participants.

The tests within the battery provide a comprehensive assessment with respect to the cognitive processes thought to be affected by OSA (Decary et al., 2000), and many of the components of the battery used are commonly used tasks that directly correspond to those recommended (see Table 1 of reference). It could be argued that some more subtle information about specific aspects of cognitive function might be lost with aggregation of the variables. A more test-centred analysis could have been performed with the data, and results from the OSA group could be compared with the large normative sample as done in the second chapter. That method would maximise the power in detecting cognitive deficits in our study sample. We would argue that important changes should be seen in more than one test – and this requirement is an important means to reduce the risk of type I error. Variable aggregation provides one more advantage in the

presence of missing values. Missing values from individual tests are a problem in that they may lead to other data from the participant to be excluded from analyses. Assuming the variables within a factor to be similar or related, we might be able to ignore missing values from individual variables. In addition, any measurement error from the individual outcome variables may be reduced by the averaging process.

3.6 Conclusion

Interpretable outcome variables aggregated by theoretical knowledge and factor analytic methods can be used to evaluate the degree of symptomatic sleepiness and daytime functional impairment from OSA. We will relate these outcomes to our measure of interest, the EEG.

As raised above, the findings from the factor analysis may be incorrect owing to the small sample size. Checking these factors by relating them to other independent variables may support the validity of our conclusions, but any findings will still require confirmation with a new set of prospectively-collected data. We will look at our predictors of interest in the next section.

4 Relating the EEG to daytime function and sleepiness

4.1 Chapter abstract

Aim: To assess parameters derived from a mathematical model of the EEG as a measure of sleepiness.

Design: Cross-sectional study in a clinical OSA population.

Methods: Eyes-closed resting EEG recordings were fitted to the EEG mathematical model for each subject. Outcome data from a battery of neurocognitive tasks and questionnaires were aggregated into five summary factor outcomes (chapter 3).

Analysis: Pearson's correlations.

Results: N=74 subjects with EEG recordings had at least 1 summary outcome variable available for analysis. Mean age 43.8 (SD 11.2), AHI 28.5 (22.4), ESS 10.8 (4.7). Increasing cortical excitation (Parameter X) was associated with reduced depression, anxiety and stress ($r = 0.27$, $p = 0.02$, $n = 73$). The inhibitory effect of the thalamic reticular nucleus (Parameter Z or $-G_{srs}$) increased with poorer executive function ($r = -0.32$, $p = 0.01$, $n = 59$). Parameters X (net cortical excitation) and Y (feedback between thalamic reticular nucleus and thalamocortical relay nuclei) were very strongly correlated ($r = -0.99$).

Conclusion: During wake, EEG model parameters show weak to moderate correlations to sleepiness-related outcomes in a clinical population with OSA. In particular, executive function deteriorated with greater inhibitory influence of the thalamic reticular nucleus. The findings will need to be confirmed with further research.

4.2 Introduction

As mentioned in Chapter 1, there has been extensive work on the EEG as a potential measure of sleepiness. Low frequency activity (1-7 Hz) while awake has been shown to increase, and eyes-closed alpha power decreases with increasing time awake. In contrast, during sleep, slow-wave activity (1-4 Hz) in sleep has been related to dissipation of homeostatic sleep drive. Previous studies have related quantitative EEG measures to symptoms of sleepiness and decrements in neurocognitive performance, both in the laboratory and in the field (For example Cajochen et al., 1995; Daniel, 1967; Lal & Craig, 2002; Torsvall & Akerstedt, 1987).

There has also been some work in OSA patients that supports the ability of the EEG to reflect changes in sleep pressure due to OSA and its treatment. EEG slowing, as measured by a rise in the ratio of power in the delta-theta and alpha-beta bands, was reported in the eyes-closed awake condition in OSA (Morisson et al., 1998). The power ratio fell after commencement of CPAP therapy (Morisson et al., 2001).

The repetitive respiratory events characteristic of OSA may mediate the increased sleepiness commonly reported in OSA, perhaps by interfering with the ability to dissipate sleep pressure during sleep. Slow-wave sleep (stages 3 and 4) is reduced with increasing apnea-hypopnea index (Redline et al., 2004). Delta power during sleep (Guilleminault et al., 2001) and the number of visually-scored slow waves (Himanen et al., 2004), are reduced in sleep-disordered breathing. The degree of reduction in delta power in the first sleep cycle correlated with sleepiness as measured by the Multiple Sleep Latency Test (Heinzer et al., 2001). The expected exponential decay of delta activity is attenuated in mild OSA (Ondze et al., 2003).

These OSA-related changes in the sleep EEG respond to treatment: after nine months of CPAP therapy in ten subjects with severe OSA, there was an increase in delta EEG power in the first 2 sleep cycles (Heinzer et al., 2001). The relative contributions of OSA-mediated sleep fragmentation and intermittent hypoxia on the change in sleep EEG has yet to be elucidated, although an animal study has suggested the hypoxia to be the main process (Polotsky et al., 2006).

Work involving quantitative measures of the electroencephalogram has largely been observational, rather than based on theory. Most studies make empiric observations between particular frequencies of the EEG and increasing time awake, or relate EEG to tests that reflect sleepiness. There has been work theorizing links between particular EEG rhythms and underlying brain processes, but variables derived from a recorded EEG are not directly quantified in terms of interactions between underlying brain nuclei.

A model has been developed (page 43) that links current understanding of interactions between groups of neurons to the observed (recorded) EEG (Robinson et al., 2005; Robinson et al., 1997). This may allow relation of observed EEG to its neurophysiologic basis, and the parameters reflecting strengths of interactions between nuclei can be derived from the model, given an observed EEG power spectrum. Based on the currently-held theory of cortico-thalamic interactions governing wake-sleep transitions, it might be expected that with increasing sleepiness, there would be greater influence of the thalamic reticular nucleus upon the thalamocortical relay nuclei (expressed as parameters Z (internal feedback between thalamic reticular nucleus and thalamocortical relay nuclei), and also G_{srs}), and reduced cortical excitation (parameters X (net cortical excitation) and G_{ee} (excitatory gain in pyramidal cells)).

In this study, the parameters derived from the EEG mathematical model will be evaluated as a measure of sleepiness in OSA patients, by relating them to functional outcome measures relevant to sleepiness in OSA. Specific aims are:

1. To confirm that EEG recorded with usual sleep laboratory equipment can be fitted to the model.
2. To relate EEG model parameters of interest and EEG power with sleepiness and daytime functional outcomes in OSA.

4.3 Methods

Participants

This study draws from the cross-sectional study described in chapter 3 (page 83). Participants attending the sleep laboratory to investigate possible sleep apnea were asked to participate.

Procedure

Participants completed questionnaires, performed a battery of neurocognitive assessment tasks, had a diagnostic polysomnogram, and attended a 5-minute resting recording of the EEG. All assessments were usually performed within a 24-hour period.

A five-minute resting recording of the EEG was performed using the leads Fz, Cz, Pz, and Oz. Electrooculogram signals (right supra-ocular, right and left lateral) were also recorded. EEG and polysomnographic recordings were performed using standard sleep laboratory acquisition hardware (Compumedics E series, Compumedics, Melbourne, VIC). The subjects fixed their gaze at a mark with their eyes open for the first two minutes of the recording, had a half-minute transition period to close their eyes, and then were recorded with their eyes closed for a further 2 minutes (page 63).

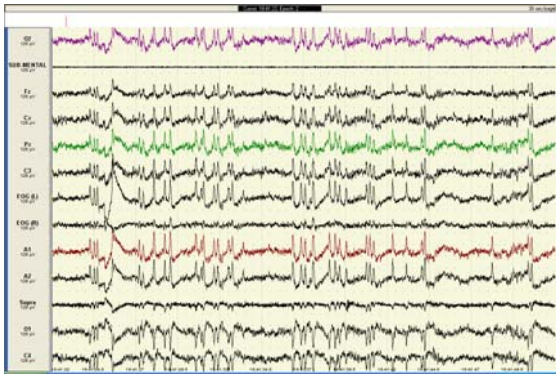
Data from questionnaires and a battery of neurocognitive tasks were aggregated into five summary outcomes, guided by factor analysis (page 104).

Data processing

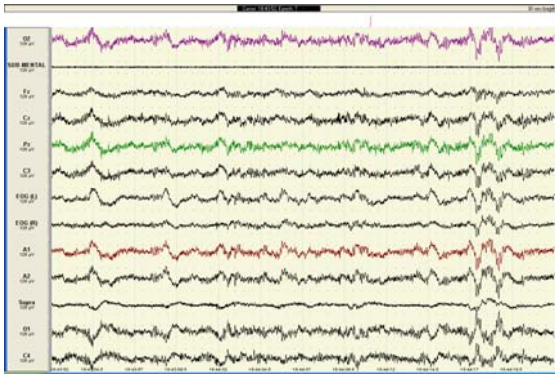
Eye movement correction and power spectral analysis

EEG data was exported to European Data Format (EDF) files for further processing. The wake EEG recordings were corrected for eye-blink artefact by a process which subtracted the waveforms observed in the EOG leads (Gratton et al., 1983). The procedure was performed by a computer program which also re-referenced the EEG to the linked mastoids (A1+A2) (C. Rennie, 2000).

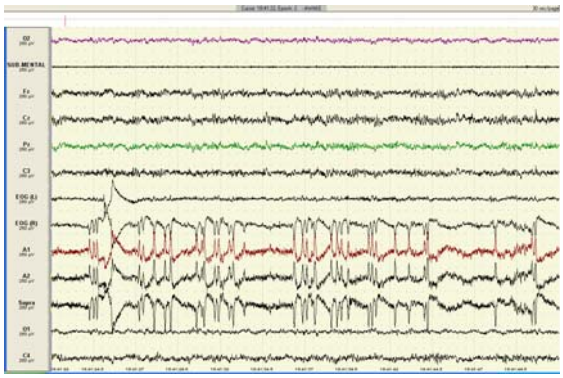
Eyes open uncorrected



Eyes closed uncorrected



Eyes open corrected



Eyes closed corrected

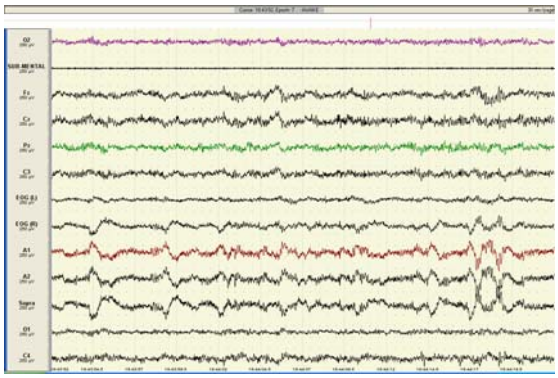


Figure 4.1 Example of resting EEG recording before and after eye-movement correction

Figure 4.1 shows a resting EEG recording from a single subject. The left column shows a 30-second period in the eyes open condition, while the right column shows the eyes closed condition. The top row shows the original, uncorrected waveforms, and the bottom row shows the result following eye movement correction. The figure illustrates that the short-duration, high amplitude eye-blinks are well corrected by the algorithm. Artefact from smaller amplitude and

lower-frequency eye movements seen in the eyes closed recording has been mostly removed, however some residual artefact persists.

The EEG derivation analysed in this study of sleepiness was Fz. The frontal and prefrontal cortex are especially vulnerable to sleep loss (Horne, 1993), and in EEG studies this area has been shown to be most sensitive to the effects of sleep loss (Cajochen, Foy et al., 1999; Cajochen et al., 2001).

The power spectra for lead Fz were generated by fast Fourier transform using non-overlapping 2-second epochs with a Welch window to reduce spectral leakage (C. Rennie, 2004a). The average power spectra were derived by calculating the arithmetic mean of individual frequency bins within the eyes open and eyes closed conditions.

EEG power was also calculated for both eyes open and closed conditions for delta (0.5-3.5 Hz), theta (4.0-7.5 Hz), alpha (8-12.5 Hz), and beta (13.0-39.5 Hz) frequency bins.

Fitting to EEG model

Spectral data for the resting *eyes closed* condition were submitted to a program that fitted observed spectra to the EEG model, by an iterative process (C. Rennie, 2004b).

Observed log-transformed power as a function of frequency is fitted as a sum of the EEG model, and a model of EMG generation. The fitting algorithm employs the Levenberg-Marquardt method (Press, 1992) to minimize a chi-squared statistic. Further details on the fitting process are provided elsewhere (Rowe, Robinson, & Rennie, 2004).

The program output provides estimates for the parameters fitted, and a chi-squared statistic indicating the goodness of fit of the model-predicted power spectrum with the observed

spectrum. The fitting algorithm requires the chi-squared value to be less than 25, with values of 12 or less indicating good fits to the observed EEG (Rowe, Robinson, Lazzaro et al., 2005).

Statistical analysis

The outcomes of interest were the five summary sleepiness-related outcome variables determined with the aid of factor analysis in chapter 3. The primary analysis explored the EEG model parameters X, Y and Z as predictors of sleepiness as described by the outcome variables. Relationships were graphically explored by scatterplot, and quantified by Pearson's correlation coefficient. EEG power in the alpha, theta and delta bands were also considered as predictors of sleepiness. To allow comparison with published literature using this EEG model, two additional parameters, G_{ee} and G_{srs} are also reported.

4.4 Results

From the 123 participants initially recruited in the cross-sectional study, data from the resting EEG recording was available from 84 subjects. Demographic and polysomnographic statistics for this subset are shown in Table 4.1.

Characteristic	Mean (SD)	N (%)
Males, N (%)		68 (81.0%)
Age, years	43.8 (11.2)	
BMI, kg/m^2	32.1 (5.8)	
AHI, events/hour	30.0 (25.8)	
Arousal index, events/hour	28.5 (22.4)	
Minimum Saturation, %	83.0 (11.1)	
Proportion of TST below 90% saturation, %	6.6 (15.8)	
Sleep Efficiency, %	82.3 (12.2)	
ESS	10.8 (4.7)	
ESS ≥ 10, N (%)		50 (61.0%)

Table 4.1 Characteristics of the study sample with wake EEG data (n=84)

Paired EEG and summary factor outcome data was available for bivariate correlations for at least 74 subjects, except for factor 5, where there were 60 subjects with available paired data).

EEG inverse model can be fitted to resting EEG recordings made with PSG hardware.

The mean chi-square value for the 84 EEG recordings fitted to the model was 3.5 (SD 2.4, range 1.1-14.6). The chi-square value was below 5 in 86% of recordings, and below 10 in 96%. These values indicate good fit of the model to the observed data.

Figure 4.2 gives four examples of EEG spectra derived from the wake resting EEG recording from lead Fz, in the eyes closed condition. Spectra are plotted on a log-log scale. The solid line shows the original observed EEG spectrum. The spectra were fitted to the EEG model, and the estimated model parameters were then used to generate predicted (or fitted) spectra, shown with the dotted line. These plots are used to check the quality of the model fits, and hence the robustness of the parameter estimates. In these examples, fitted spectra closely resemble the profiles of the observed EEG power spectra.

Observed spectra show artefact from muscle activity (EMG artefact) above approximately 30 Hz. This was particularly marked in one recording (Figure 4.2, top right). The peak at 10 Hz arises from the alpha rhythm, prominent in the resting eyes closed condition. An example of a waking eyes open power spectrum, from the same subject as depicted in the top left panel of Figure 4.2, is shown in **Error! Reference source not found.**: as expected the alpha peak is less prominent with the eyes open.

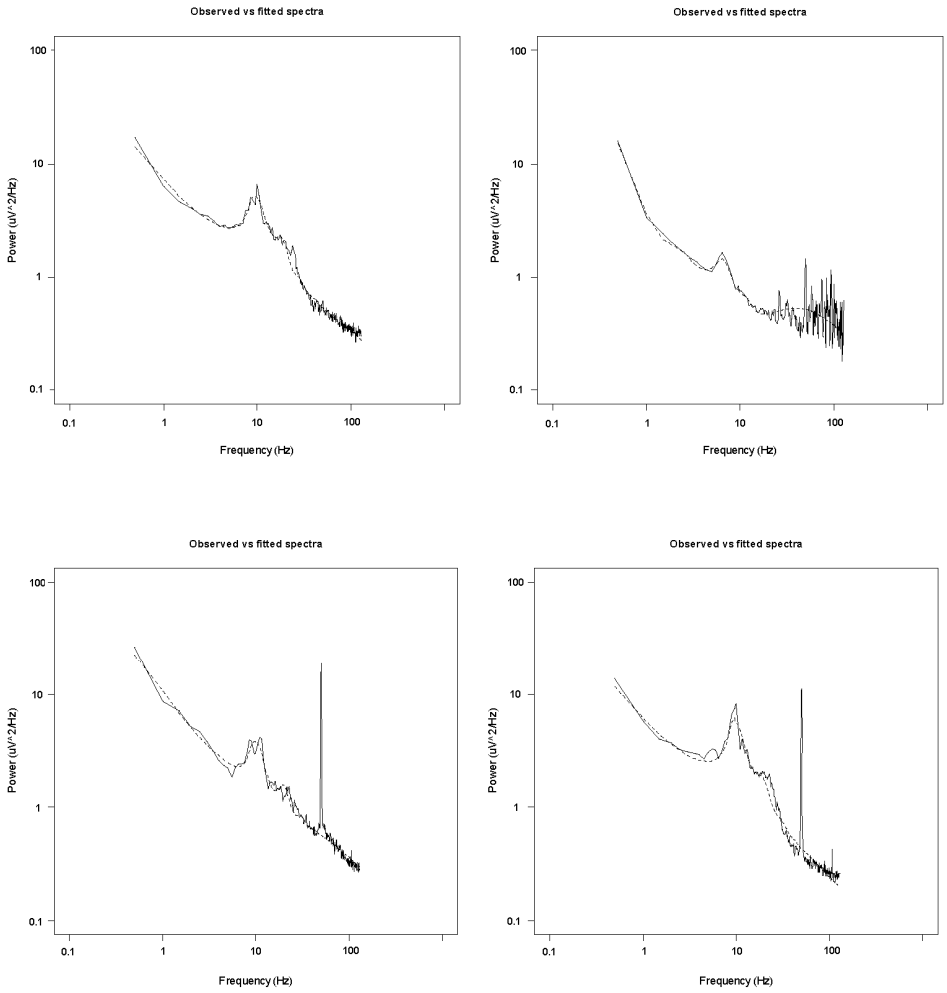


Figure 4.2 Examples of observed (solid line) vs. model fitted (dotted line) eyes-closed wake spectra

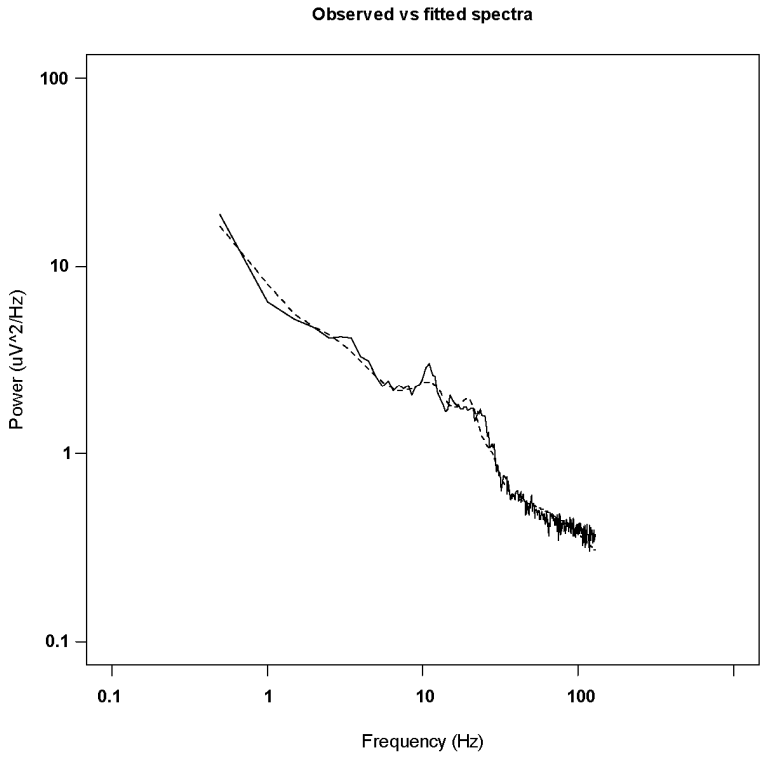


Figure 4.3 Example of EEG power spectrum from Fz, wake eyes open condition

Relation of EEG model parameters to factor summary outcomes

Figure 4.4 depicts a series of scatterplots arranged as a correlation matrix. The boxes below the diagonal plot relationships between the EEG model parameters of interest from fitting the eyes closed wake EEG spectra: parameter X (ec.X), parameter Y (ec.Y), parameter Z (ec.Z), and the five summary factor outcomes (factor1 to factor5). A *loess* smoother is applied to the scatterplots to aid visualization of relationships (or lack thereof). The boxes above the diagonal show absolute values of Pearson's correlation coefficients, with the size of the type proportional to the magnitude of the correlation coefficient.

The larger number of variables included have caused plot details to be rendered indistinctly. Nevertheless, it is apparent that there is a strong negative relationship between parameters X and Y. It also appears that relationships between model parameters and factor outcomes are not strong. A table showing the correlations between the model parameters and the five factor outcomes is provided Table 4.2.

	Parameter X	Parameter Y	Parameter Z
Factor 1	0.20	-0.22	0.09
Factor 2	0.27*	-0.30*	0.10
Factor 3	-0.10	0.12	-0.13
Factor 4	-0.13	0.15	-0.10
Factor 5	0.25	0.24	-0.32*

Table 4.2 Correlations between EEG model parameters and factor outcomes

* $p < 0.05$, ** $p < 0.01$

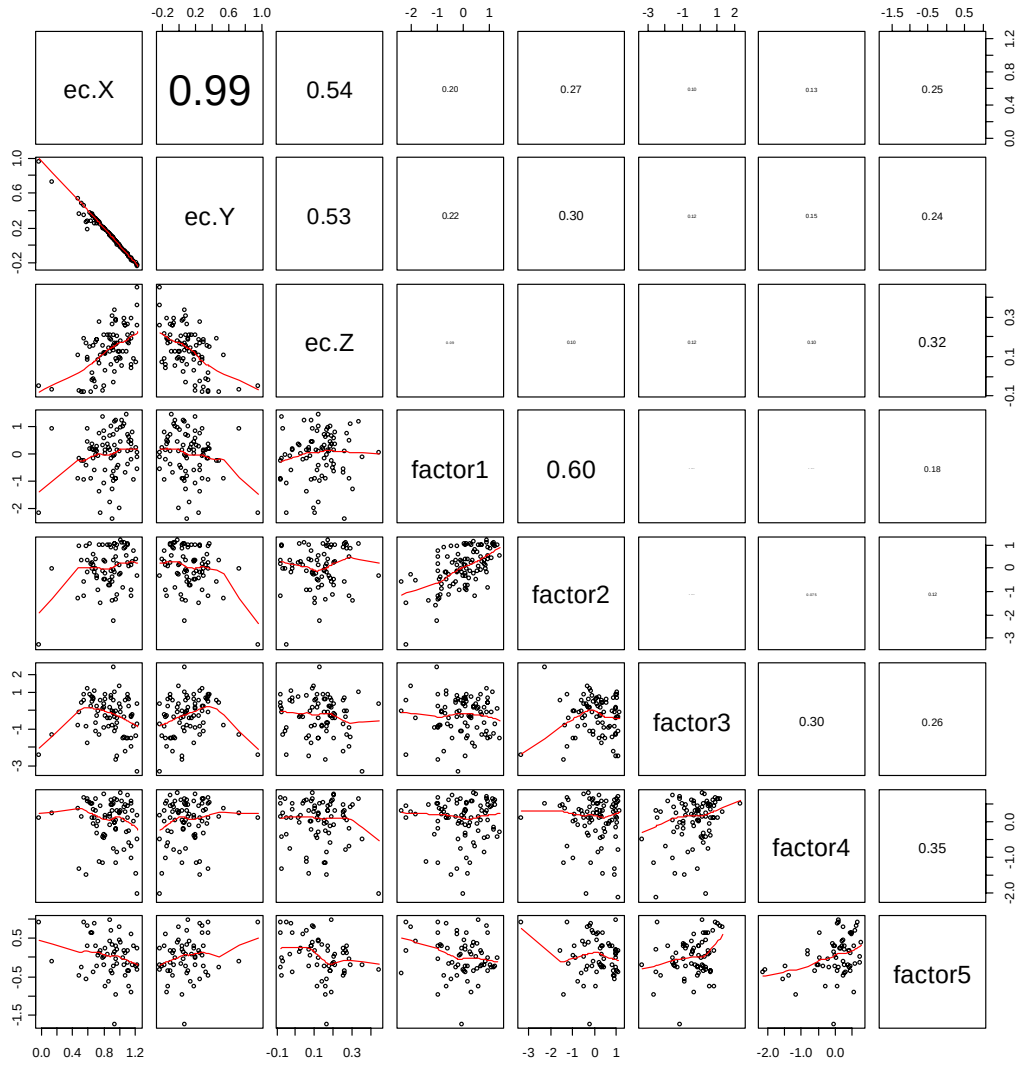


Figure 4.4 Scatterplots of showing relationships between EEG model parameters and the five summary outcomes

Parameter X

Parameter X is a measure of net cortical excitation. On inspection of the scatterplots, it should be noted that no strong trends between Parameter X and the factor outcomes are apparent.

Parameter X was positively correlated with summary Factor 2 Mood and Anxiety ($r = 0.27$, $p = 0.02$, $n = 73$). The direction of effect suggests greater cortical excitation (increasing values of parameter X) with reduced levels of depressed mood, anxiety or stress. There were two low outlying values ($X < 0.2$) for parameter X (subject ID 31 and 45, chi-squared for model fit 2.7 and 4.7). After exclusion of these two outliers the relationship remained statistically significant ($r = 0.44$, $p = 0.03$, $n = 71$).

There was a trend towards a negative correlation with Factor 5 Executive Function ($r = -0.25$, $p = 0.05$, $n = 59$). As higher values of Factor 5 imply better performance on the executive function tasks, here the direction of effect suggests poorer executive function performance with increasing net cortical excitation.

To clarify if the results were affected by the quality of the EEG model fits to the observed spectra, the correlation between Parameter X and Factor 5 was repeated, only including recordings with excellent EEG model fits (chi-squared less than 5). The relationship was unchanged ($r = -0.28$, $p = 0.1$, $n = 51$).

Parameter Y

Parameter Y is a measure of net cortico-thalamic excitation. As shown in Figure 4.4 this parameter is highly negatively correlated with parameter X. Parameter Y is negatively correlated with Factor 2 Mood and Anxiety ($r = -0.30$, $p = 0.01$, $n = 73$). Increasing parameter Y is associated with increased levels of depressed mood, anxiety and stress. There was one high outlying value for parameter Y. Removal of this potentially influential outlier (subject ID 45,

chi-square for model fit = 4.7) results in the apparent relationship being not statistically significant (r = -0.11, p = 0.3, n = 72).

Parameter Z

Parameter Z measures the internal feedback between the thalamic reticular and thalamic relay nuclei. A more negative value for Parameter Z implies a greater inhibitory effect of the thalamic reticular nucleus on the thalamocortical relay nuclei. There was a significant correlation between parameter Z and Factor 5 Executive Function (r = -0.32, p = 0.01, n = 59), with reduced performance on executive function tests accompanying increasing magnitude of parameter Z, as expected.

As a sensitivity analysis exploring the effect of quality of the model fit on the analysis results, the direction of the relationship was unchanged with excluding 8 recordings with chi-squared statistic greater than 5 (r =-0.33, p = 0.02, n = 51).

Gain parameters G_{ee} and G_{srs}

Figure 4.5 below shows the relationship between the two gain parameters of interest, G_{ee} and G_{srs}, with the transformed parameters *X* and *Z*. Parameter *X* and G_{ee} were significantly positively correlated (r = 0.37, p < 0001). There was a significant negative correlation between parameter G_{ee} and Factor 4 Driving (r = -0.23, p = 0.04), implying increased excitatory gain in the cortex (pyramidal cells) with deteriorating driving ability.

As expected from the formula from which parameter Z is calculated,[1] Z and G_{srs} are strongly inversely correlated (r = 1.00). Thus, as with parameter Z, there was a significant correlation between G_{srs} and Factor 5 executive function (r = 0.32, p = 0.01), again implying greater inhibitory effect of the thalamic reticular nucleus on the thalamocortical relay nuclei with worse executive function.

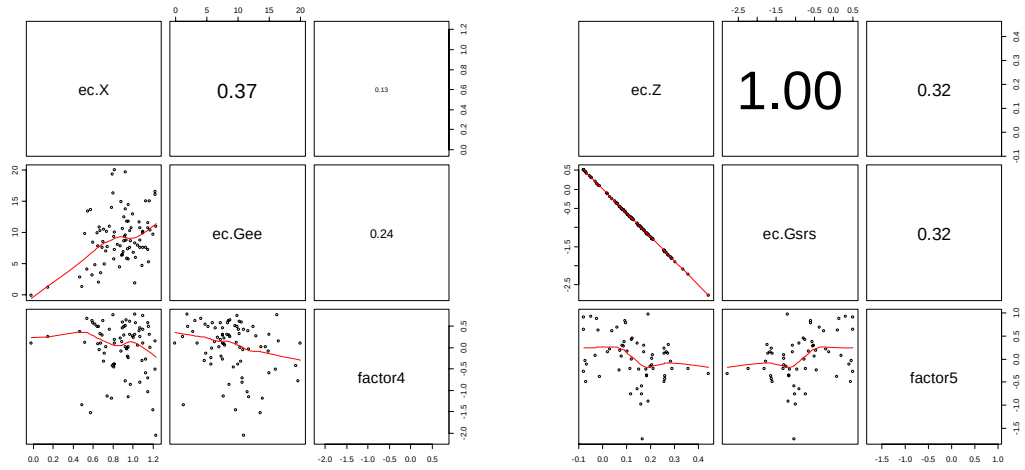


Figure 4.5 Correlations between X and Gee, Z and Gsrs

Relation of EEG delta, theta and alpha power to factor summary outcomes.

The scatterplots of eyes closed EEG delta, theta and alpha power, and the five summary factor outcomes are shown in Figure 4.6. No strong relationships are apparent. Eyes closed EEG alpha power is negatively correlated with Factor 5 Executive Function (p = -0.26, p = 0.04, n = 58,

[1] $Z = - G_{srs}\alpha\beta/(\alpha + \beta)^2$ (Equation 16, O'Connor & Robinson, 2004).

higher alpha power with worse performance). The direction of effect is a little surprising given that a higher EEG alpha in the eyes *closed* condition usually accompanies greater alertness. Equally surprising is the positive correlation between theta and alpha power as theta and delta usually rise with sleepiness, while eyes closed alpha should fall.

The apparent weak negative correlation between eyes-closed EEG theta power and Factor 4 Driving is not significant ($r = -0.18$, $p = 0.1$, $n = 78$).

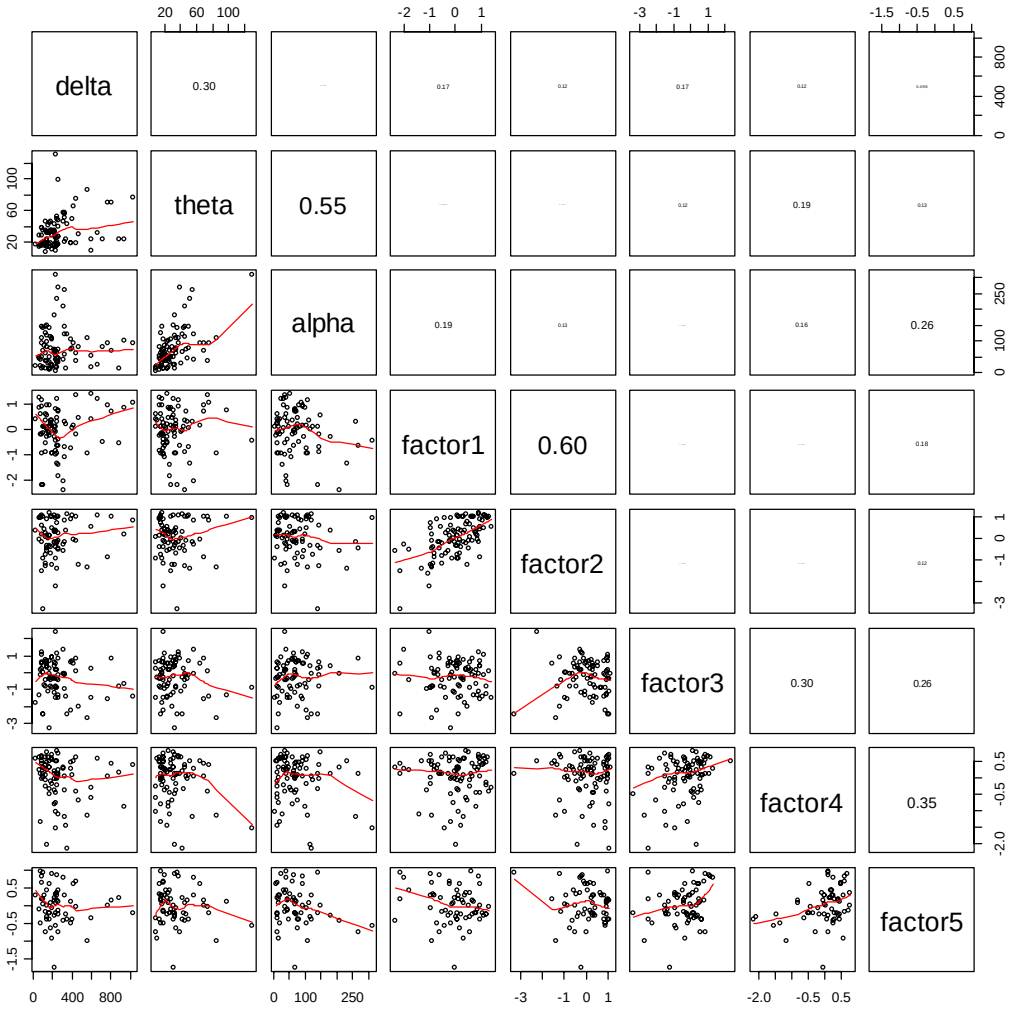


Figure 4.6 Scatterplots of showing relationships between EEG delta, theta and alpha power and the five summary outcomes

Relationship between EEG power variables and EEG model parameters

Scatterplots show moderate correlations between eyes closed delta power and the model parameters X and Y. EEG alpha power is weakly correlated with all 3 model parameters ($r \approx 0.28$). (Figure 4.7) As noted previously, parameters X and Y are strongly negatively correlated.

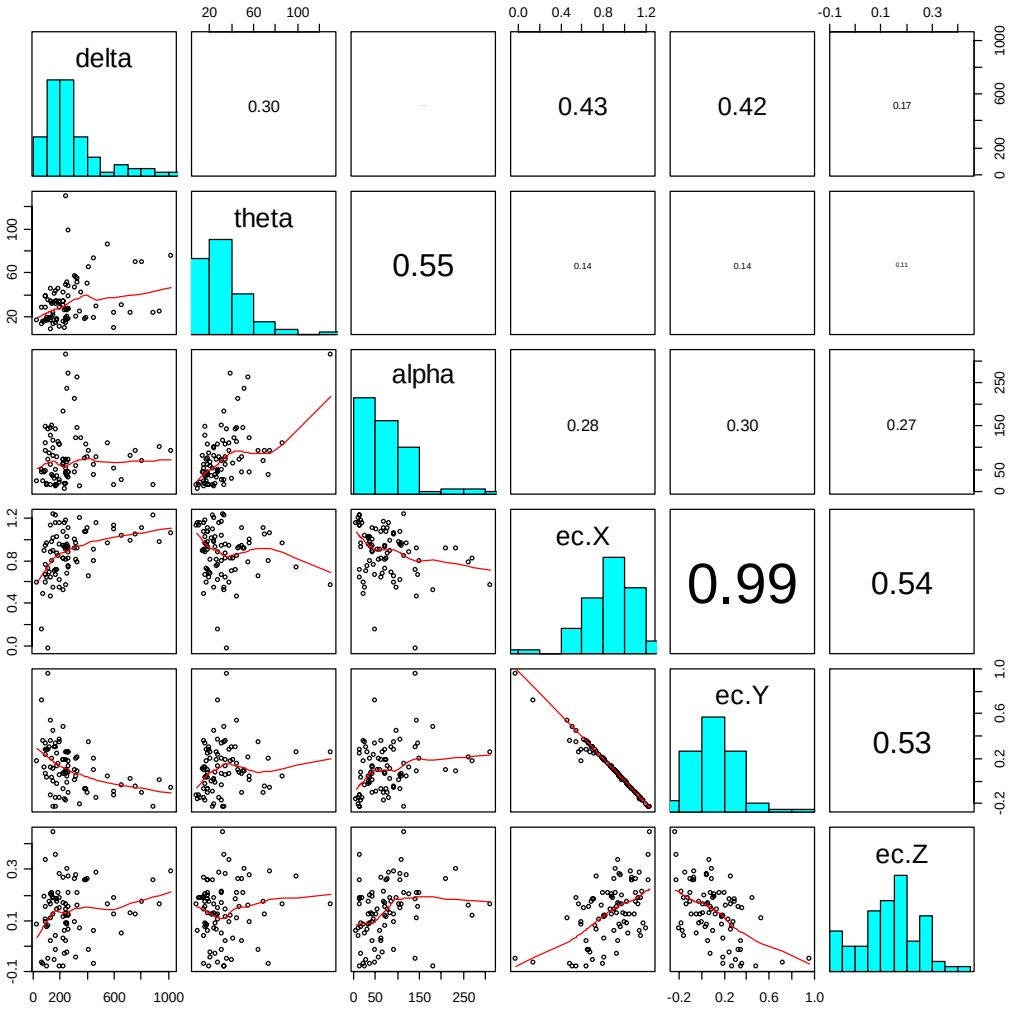


Figure 4.7 Scatterplots of showing relationships between EEG delta, theta and alpha power, and the three model parameters

The diagonal of the matrix shows histograms of the distribution of the EEG power and model parameter values.

4.5 Conclusion

During wake, EEG model parameters show a few significant but weak-to-moderate correlations to sleepiness-related outcomes in a clinical population with OSA. A weak relationship was found between eyes-closed EEG power and the summary outcome relating to executive function.

The finding that the inhibitory effect of the thalamic reticular nucleus (Parameter *Z*) increased with poorer executive function is consistent the current theory predicting reduced vigilance as a result of greater thalamic reticular nucleus-influenced hyperpolarisation of the thalamocortical relay cells (Steriade, 2000). That theory proposes that the greater hyperpolarisation of the relay nuclei would reduce alertness by alertness by reducing transmission of sensory information from the periphery, and triggering oscillations between the cortex and thalamus essential to sleep. As executive function has been shown to be vulnerable to sleep loss and reduced alertness, such an effect would be expected if the thalamic reticular nucleus were an important mediator of alertness (and hence the ability to perform on tests of executive function) in an awake subject.

A recently published study using the EEG model to fit eyes-closed EEG in lead Cz in a group of community-dwelling elderly subjects with subjective memory complaints also found high inhibitory intrathalamic gain (Parameter G_{srs}, which is negatively proportional to parameter *Z*) to be associated with lower completion times and a faster decline in errors in the executive maze task of the BRID battery (Alexander et al., 2006).

In the present study, parameters *X* and *Z* were moderately positively correlated. Increasing net cortical excitation (Parameter *X*) was associated with reduced depression, anxiety, and stress.

Eyes-closed alpha power increased with worsening performance on executive function tests. The direction of effect indicated by the relationship was unexpected and contrary to the bulk of prior

literature showing greater eyes-closed alpha with increasing alertness. Given greater OSA severity on the polysomnogram was also associated with better executive function (page 115), validity of the executive function factor (Factor 5) will need to be verified. While the unexpected result should be viewed with caution, a small number other studies in subjects with sleep-disordered breathing relating EEG to sleepiness have also produced unexpected results. A study in patients with sleep-disordered breathing showed the ratio of eyes closed to eyes open alpha power (the Alpha Attenuation Coefficient) falling, rather than rising, with longer latency on the Maintenance of Wakefulness Test (i.e. increasing alertness) (Sforza et al., 2002). A pair of studies of OSA patients before and after treatment also did not show a relationship between MSLT-quantified sleepiness and EEG slowing (delta and theta power divided by alpha and beta power) in the resting eyes closed EEG (Morisson et al., 2001; Morisson et al., 1998). The authors of the first study proposed that other factors in sleep-disordered breathing (such as the hypoxia) might lead to a different EEG response to sleepiness, that the MSLT or MWT measured a different phenomenon to the EEG, or that the variation in sleepiness in the study sample was smaller than that seen during sleep deprivation, leading to a small signal (Sforza et al., 2002). Apart from hypoxia, the chronicity of the sleep apnea condition may also allow an individual to adapt to the effects of the OSA-mediated injury during sleep.

What are the implications if the findings were replicated by further research? It might suggest that the parameters from the EEG model might reflect the effects of sleep loss in terms of deteriorating prefrontal lobe function. The opposite direction of relationship between alpha power and the same outcome, executive function is unexpected.

Spurious results may have resulted from measurement error in the generation of the EEG spectra or the summary outcome measures, influential outlying values, or a type I error. The effect of

residual artefact was apparent in the examples of EEG spectra in Figure 4.2. More data from longer recordings of the resting EEG might have improved the reliability of the generated power spectra.

Alternatively, the eye movement artefact removal procedure may have removed the expected effect on the EEG by removing EEG power in the lower frequencies due to the slow eye movements accompanying drowsiness. The latter and eyelid closure and blinks might be an important contributor to the EEG power changes with increasing sleepiness reported in the literature (T. Akerstedt, 1995; Torsvall & Akerstedt, 1985).

The aggregated summary variables have yet to be validated, however the effect of measurement error might once again be mitigated by including a larger neurocognitive assessment battery. There were a small number of outlying values of some model parameters apparent in the scatterplots. The conclusions of the study were robust to their exclusion from the analysis. These outliers came from recordings with apparently good fits to the data: they had low values for the chi-squared statistic that were less than 5. The direction of the relationships between factor 5 and parameters X and Y were unchanged after only including the recordings with chi-squared values indicating excellent model fit (less than 5). It should be noted that the value of the chi-squared statistic in determining reliability of parameters derived from fitting the model has not been established. Furthermore, a good fit to an observed spectrum contaminated by artefact might also yield incorrect parameters. Lastly, type I error is possible given the number of comparisons made, and the relatively high p-value (mostly 0.01 to 0.03).

The strength of the associations between any of the EEG variables and functional outcomes identified were at most weak to moderate. A correlation coefficient of 0.30 still leaves more than 90% of the variance is unexplained, but studies relating functional outcomes to plausible

potential predictors have also found relationships of similar strength (Adams et al., 2001). The study did not demonstrate consistent relationships across a majority of the summary outcomes. While measurement error might be reduced by longer EEG recording times to improve measurement of the EEG spectrum, possible technical improvements in the model fitting algorithm, or more extensive neurocognitive testing to improve characterisation of functional outcomes, the lack of a strong relationship between the EEG model parameters and the outcomes evaluated here support the need for ongoing research to confirm the current findings, but do not as yet support the use of these measures as an indicator of sleepiness in clinical or occupational applications such as fitness to work or drive.

5 Evaluating the EEG as a measure of sleep depth

5.1 Chapter abstract

Aim: To evaluate the parameters from the EEG mathematical model as a measure of depth of nocturnal NREM sleep.

Design: Observational study, with N=8 subjects with regular sleep architecture and AHI < 10.

Methods: The first nocturnal NREM sleep cycle was divided into 20 time segments. EEG model parameters and delta and sigma band EEG power were calculated for each segment. Linear and quadratic trends were explored, using linear mixed-effects models.

Results: Net cortical excitation (Parameter X) showed a linear decreasing trend ($p = 0.03$). Parameter Z showed a negative linear trend ($p = 0.0005$), reflecting reduced inhibitory influence of the thalamic reticular nucleus on the corticothalamic relay nuclei with increasing depth of sleep. There was a significant decreasing linear trend for sigma power ($p=0.04$) and significant increasing linear and negative quadratic trends for delta power (both $p < 0.001$).

Conclusion: The present study has reproduced previous findings regarding the changes in EEG power with deepening NREM sleep. Despite this, EEG model predictions of thalamocortical interactions are not consistent with present theory: the model predicts reduced inhibitory thalamic reticular cell activity with increasing NREM sleep depth.

5.2 Introduction

Within a NREM sleep cycle, the progression from light to deep is marked by the transition through four arbitrarily-defined sleep stages. In examining EEG power, a more gradual progression can be appreciated. Delta power (0.75 – 4.5 Hz) gradually increases, while spindle frequency activity (12.5 – 15.25 Hz) rises to a peak within the first 10-15 minutes, followed by a slow decline (Aeschbach & Borbely, 1993). These changes are the most dramatic for the first NREM cycle, as slow-wave sleep is reduced as the night progresses.

The time course of slow wave activity and sleep spindles, and their reciprocal relationship (D. J. Dijk et al., 1993), is thought to arise from the progressive hyperpolarisation of thalamocortical relay neurons as sleep deepens (Steriade, 2003). GABA-ergic thalamic reticular neurons from the thalamic reticular nucleus interact with the thalamocortical neurons, and induce their hyperpolarisation. At lower levels of hyperpolarisation, oscillations at a spindle frequency occur, while with further hyperpolarisation, slow waves arise. With this process there is increased gating of sensory information from the periphery being transmitted to the cortex.

A model has been developed in reference to the theorised interaction between cortex and thalamic nuclei, and provides a means to link these interactions with the observed EEG. This EEG model has been fitted to waking EEG, as well as a small number of sleep EEG recordings (Robinson et al., 2005; Robinson et al., 1997).

This study will evaluate the parameters derived from the EEG mathematical model as a measure of sleep depth. In a group of subjects with regular sleep architecture, the study will show that previously-observed changes in slow frequency power with increasing NREM sleep are also reflected in changes in EEG model parameters describing interactions between cortex and

thalamus. The hypothesis was that Parameter Z, reflecting the inhibitory reciprocal interaction between the thalamic reticular nucleus and thalamocortical relay nuclei will rise (become more positive) as NREM sleep deepens.

5.3 Methods

Participants

Data from the overnight polysomnography of 8 subjects from the cross-sectional study were used for this analysis. Subjects were included if they had an apnea-hypopnea index (AHI) of <10, and a normal- or regular-appearing sleep architecture, in terms of the usual progression of sleep stages (see examples, Figure 5.1). Inclusion and exclusion criteria relevant to the cross-sectional study as a whole have been described previously (page 62).

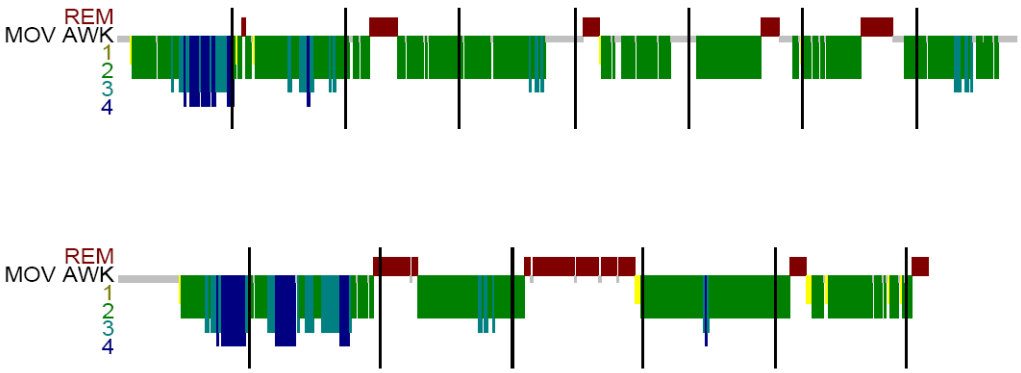


Figure 5.1 Illustration of sleep stages from 2 subjects

Data preparation and analysis

EEG and polysomnographic recordings were performed using standard sleep laboratory acquisition hardware (Compumedics E series, Compumedics, Melbourne, VIC). Sleep staging was performed manually, as described previously (page 85).

Start and end epoch of the first cycle of NREM sleep were identified, according to criteria described by Feinberg & Floyd (see Table 5.1), with the modification that the NREM cycle was considered terminated if there was a shift to a lighter sleep stage for greater than 5 minutes (Darchia et al., 2007; Feinberg & Floyd, 1979).

- 15-minute minimum duration for each NREM period.
- Last NREM period of the night has to be followed by 5 minutes or more REM before awakening.
- Last REM period has to be followed by 5 min or longer of NREM. (No sleep period complete if interrupted by final awakening.)
- During night, duration of each REM period is at least 5 min, except for the first REM period (which has no minimum length).
- After the first REM period, REM episodes of < 5 minute are added to the previous complete REM period.
- REM periods interrupted by less than 15 min of continuous NREM is treated as a single episode, with any intercurrent NREM sleep added to the previous complete NREM period.

Table 5.1 Feinberg criteria for defining NREM sleep cycles

EEG data were exported to European Data Format (EDF) files for further processing. The power spectra from lead Fz were generated in 1-second epochs by fast Fourier transform using a Welch window (C. Rennie, 2004a).

The first NREM cycle was divided into 20 equal time segments. Within each of the 20 segments, an average power spectrum was obtained by arithmetic mean of each frequency bin. Power within delta and sigma bands were derived by obtaining the sum of power over the 1-4 Hz and 12-14 Hz bins respectively.

The EEG model fitting computer program generated estimates for the parameters relating to each of the 20 NREM sleep segments (C. Rennie, 2004b). Frequency bins above 39 Hz were discarded after it was found that excessive EMG artefact from the higher frequencies led to problems with reliable model fitting. The starting value for the parameter lnorm (the power normalization parameter to adjust for the variation between recordings in total power) was set at 3.0 to optimise the fitting procedure.

Statistical analysis

The EEG parameter values and EEG power in the delta and sigma bands were analysed by linear mixed models. SEGMENT was entered as numerical variable with integer values 1-20. The extent to which the changes across the NREM cycle in the outcomes followed a linear or curved pattern was explored with linear and quadratic fixed effects terms of SEGMENT. Random effects of intercept and SEGMENT were considered, to allow for inter-individual random variation in overall (average) values of the outcome and in the linear slope of the outcome across time. The best fitting model was determined by minimising Akaike's Information Criterion (Sakamoto, Ishiguro, & Kitagawa, 1986).

The EEG model parameters of interest were the summary parameters X, a measure of net cortical excitation, Y, a measure of net cortico-thalamic excitation, and Z the internal feedback between the thalamic reticular and relay nuclei. As a secondary analysis, gain parameters G_{ee} and G_{srs} are also reported, to allow comparison with published literature reporting these parameters from the model.

5.4 Results

PSG data from 8 subjects from the cross-sectional study were included in this analysis. All were males, aged 44.9 years (range 26-60 years). The mean (SD) AHI was 5.3 (3.8) per hour, arousal index 11.7 (4.3) per hour. ESS was 10 (4.1). The mean length of the first NREM cycle was 43.7 (14.0) minutes, the range was 28 to 67.5 minutes.

Delta and sigma EEG power

As expected from previous reports (Aeschbach & Borbely, 1993), sigma power appeared to rise early in the NREM sleep cycle, and then gradually fell. (Figure 5.2) There was a significant negative linear effect of SEGMENT ($p = 0.04$, Output 5.1).

```
Fixed effects: eeg.sigma ~ segment

               Value Std.Error  DF   t-value p-value

(Intercept) 25.024000  5.493393 132  4.555290  0.0000

segment     -0.576085  0.281395 132 -2.047242  0.0426
```

Output 5.1

Delta power showed a gradual rise, with a slight fall at the end of the cycle prior to REM. (Figure 5.3) Positive linear and negative quadratic effects of SEGMENT were significant ($p < 0.001$, Output 5.2).

```
Fixed effects: eeg.delta ~ segment + I(segment^2)

                Value Std.Error  DF   t-value p-value

(Intercept)  6217.997 1847.3674 131  3.365869   0.001

segment       801.914  100.8946 131  7.948037   0.000

I(segment^2)  -21.956    3.1437 131 -6.984130   0.000
```

Output 5.2

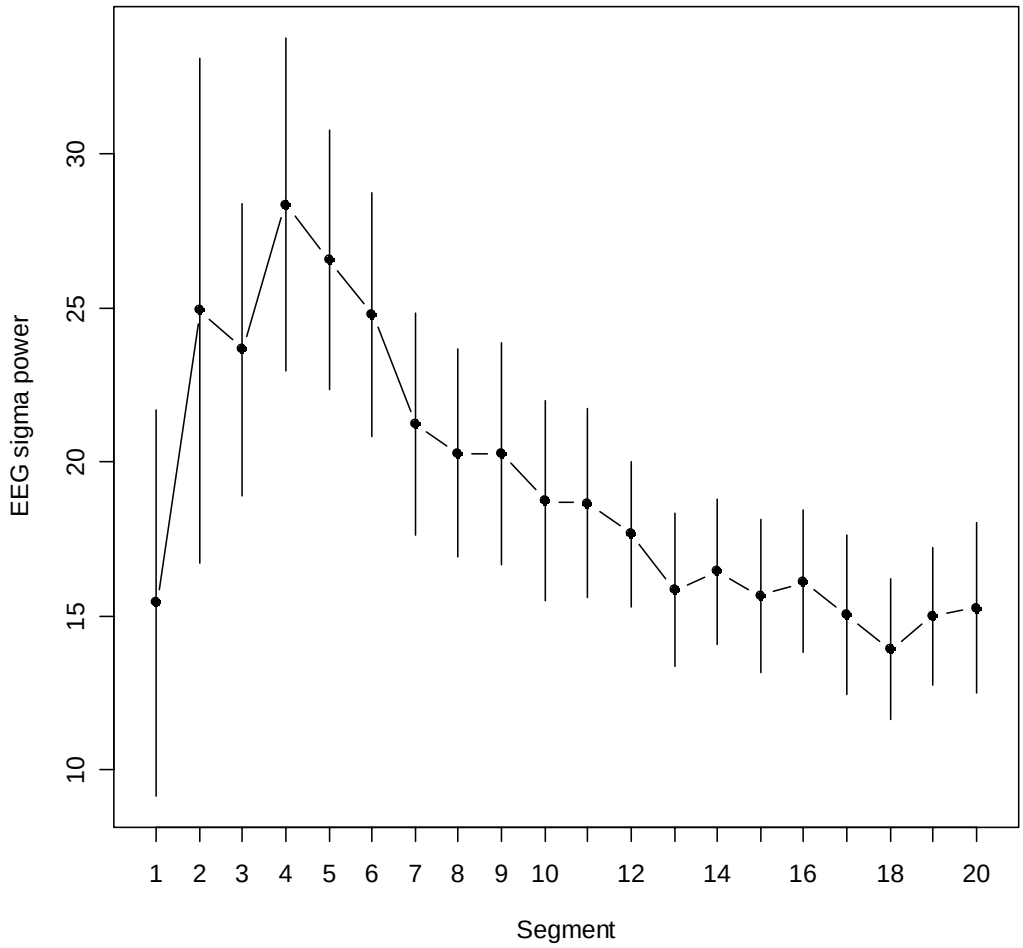


Figure 5.2 EEG sigma power across first NREM cycle

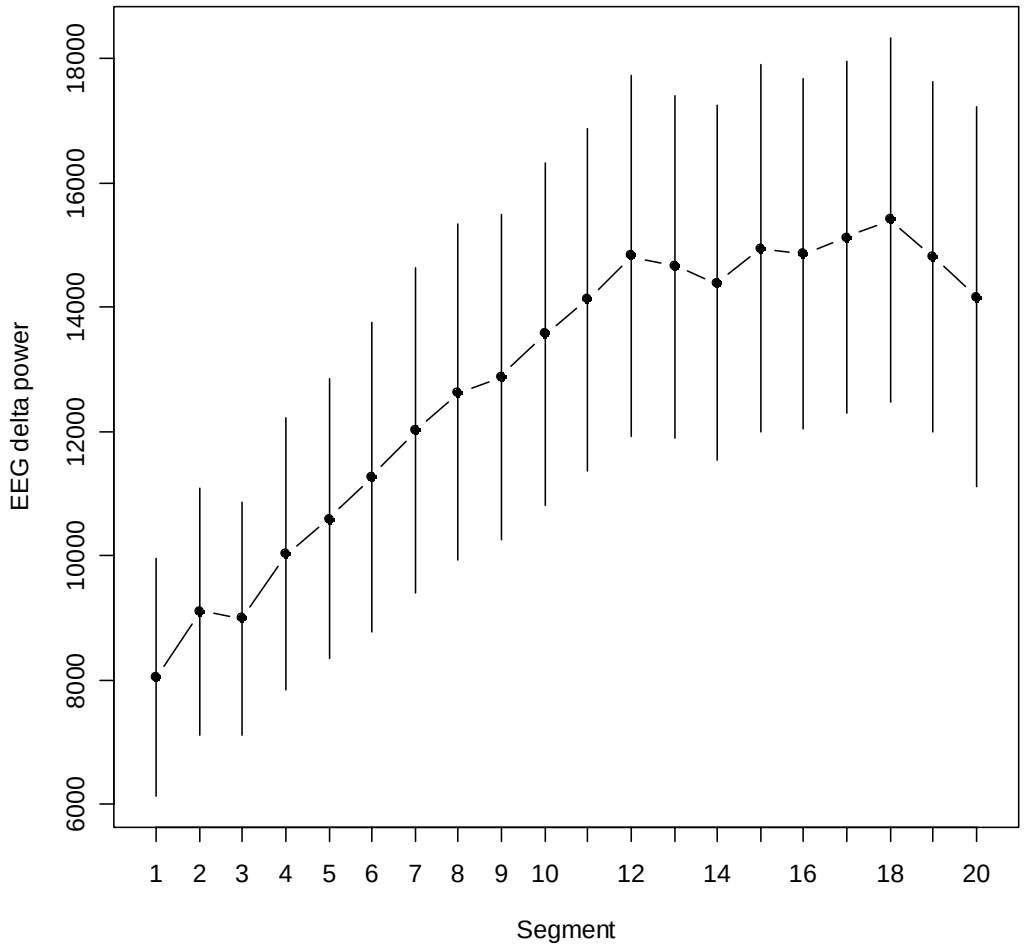


Figure 5.3 EEG delta power across first NREM cycle

Model parameters

Figure 5.4 shows an example of the original *observed* power spectrum during a segment within the first cycle of NREM sleep (solid line, segment 11 of subject 083). Parameters derived by fitting the EEG Model were then used to generate an expected (fitted) power spectrum (dotted line).

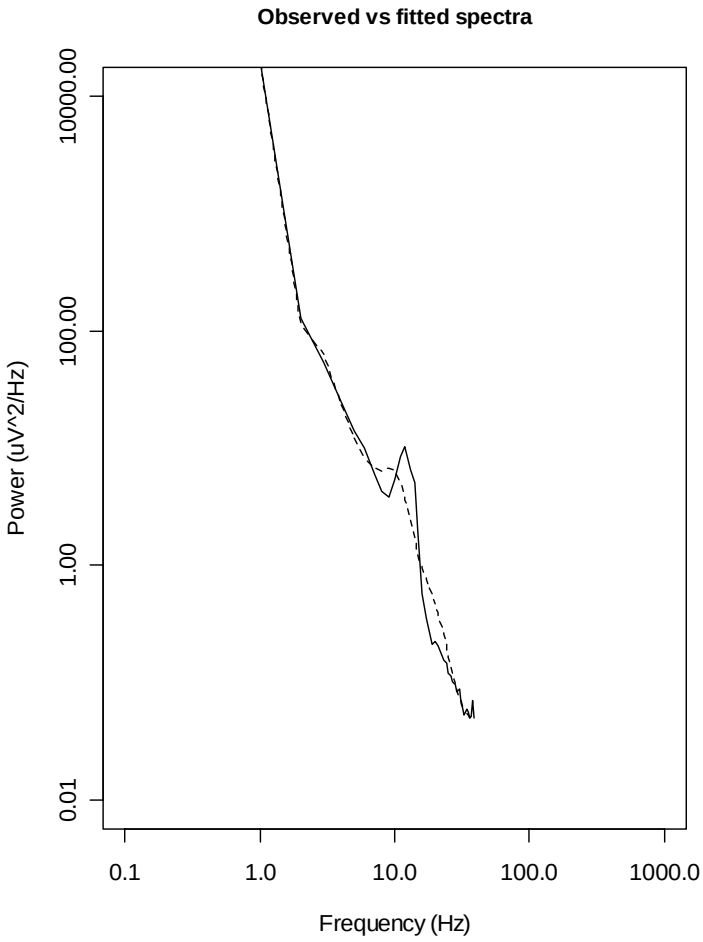


Figure 5.4 Example of observed vs. fitted spectra for a single segment

Parameter X

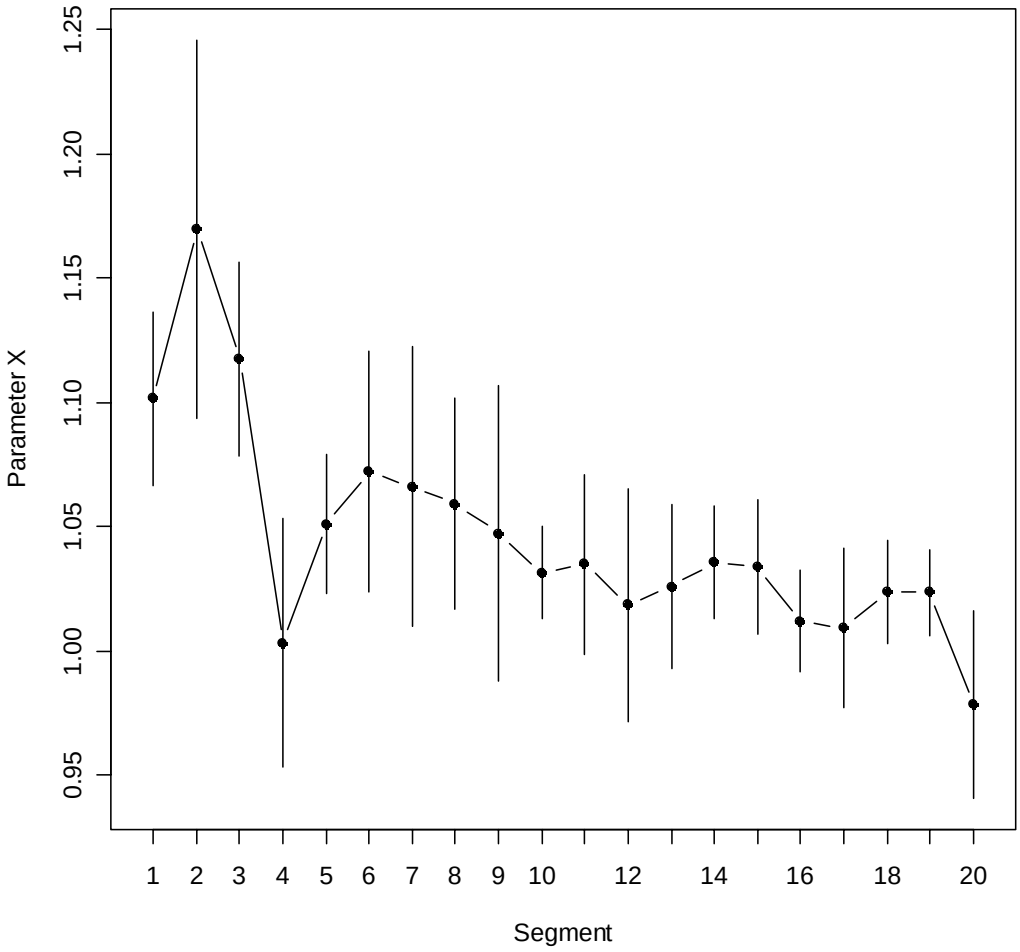


Figure 5.5 Dynamics of net cortical excitation (parameter X) across the first NREM cycle

Parameter X (Figure 5.5), which reflects net cortical excitation fell progressively across the NREM cycle (linear trend p = 0.03, Output 5.3), with no significant quadratic effects indicating no significant curvature.

```
Fixed effects: X ~ segment

                Value  Std.Error  DF  t-value p-value

(Intercept)  1.104802 0.05163218 132 21.39755   0.000

segment     -0.005739 0.00253123 132 -2.26727   0.025
```

Output 5.3

Figure 5.6 shows an augmented predictions plot. Observed values for Parameter X are plotted against segment for each subject. The lines show the individual-subject predicted linear effect of segment, augmented by the estimated random effects. It illustrates the capability of the mixed-effects model to allow for inter-subject variation in response: that is, there can be some random variation in the slope of the line. The plots show the downward linear trend to be consistent across subjects.

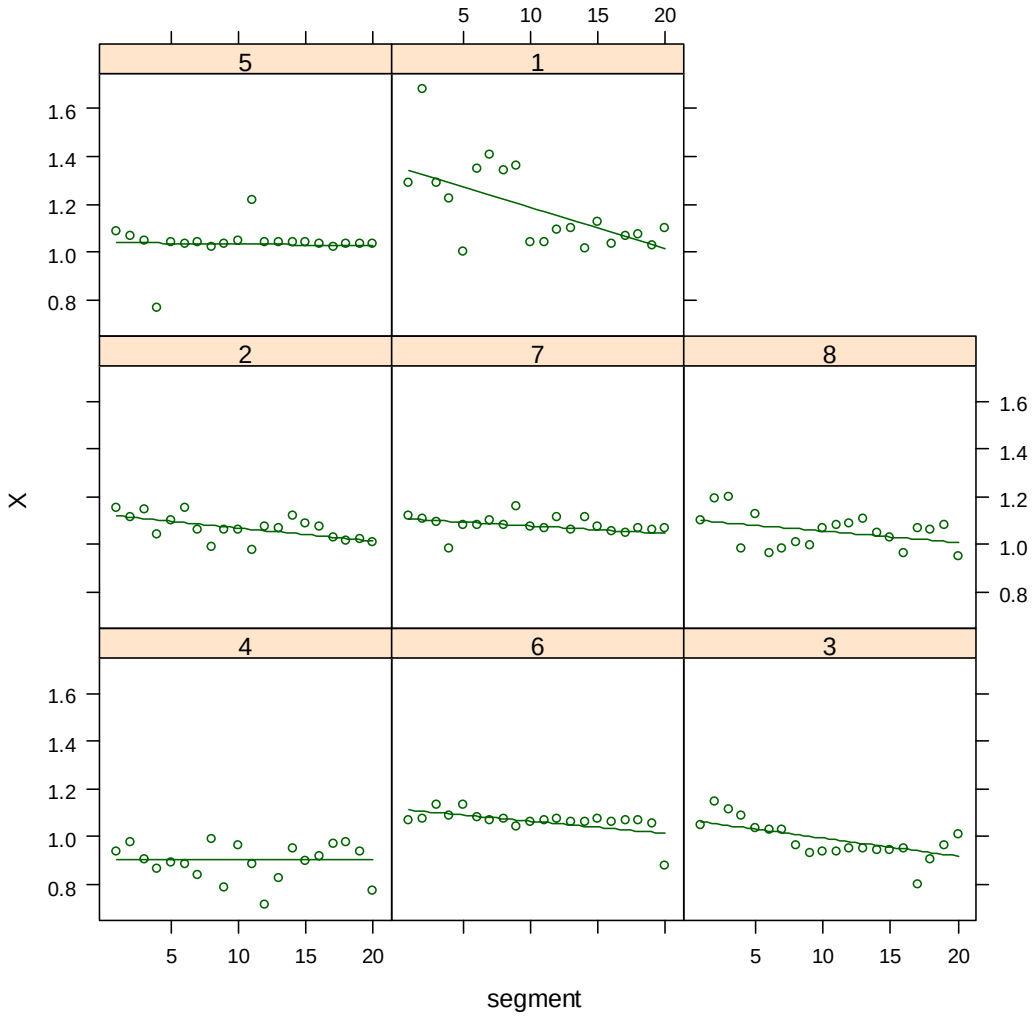


Figure 5.6 Individual subject data for Parameter X

Parameter Y

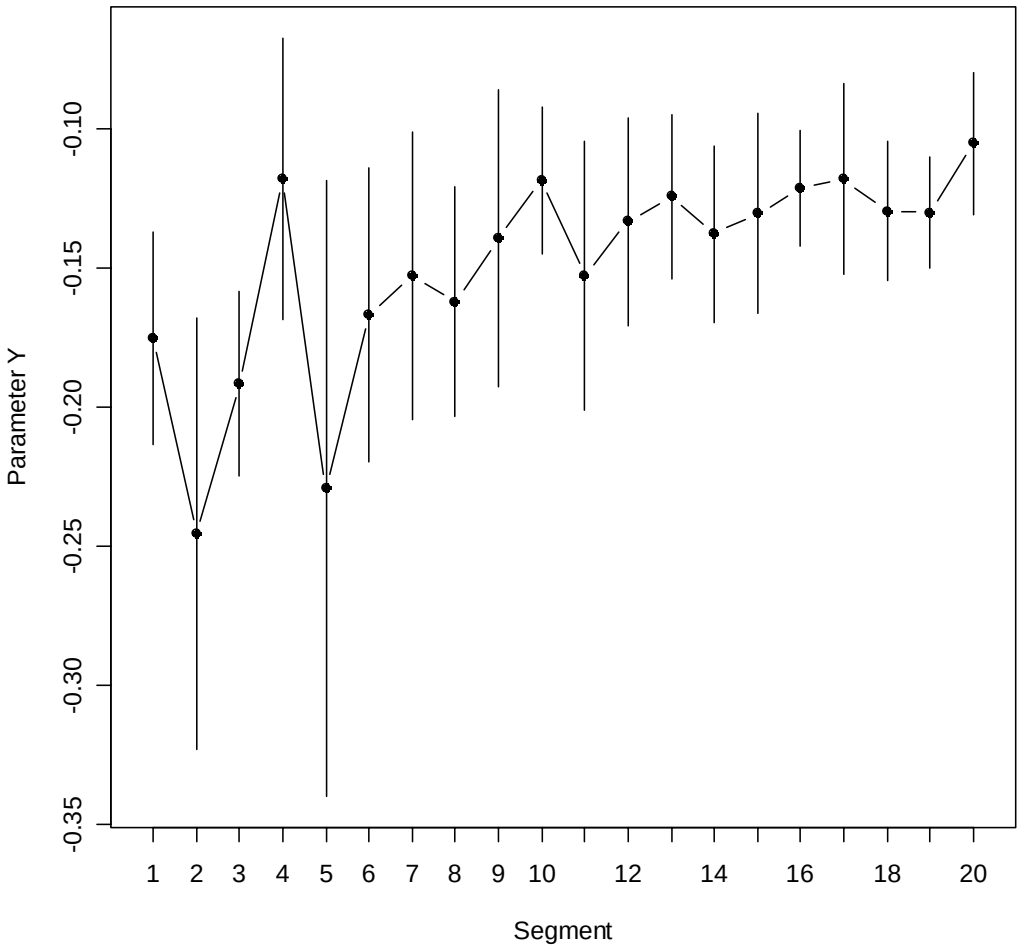


Figure 5.7 Dynamics of net cortico-thalamic excitation (parameter Y) across the first NREM cycle

Parameter Y reflects net excitation within the corticothalamic circuit. The plot for parameter Y (Figure 5.7) appeared to show a gradual increase across the first NREM sleep cycle, but the test for linear trend was not significant (p=0.18, Output 5.4).

```
Fixed effects: Y ~ segment

                  Value  Std.Error  DF   t-value p-value

(Intercept) -0.19678271 0.07086574 132 -2.776838  0.0063

segment      0.00518883 0.00385958 132  1.344404  0.1811
```

Output 5.4

Figure 5.8 shows an augmented predictions plot for Parameter Y, with a model including a random effect for segment. Open circles depict observed values for Parameter Y for each individual subject. The lines show the individual-subject predicted linear effect of segment, augmented by the estimated random effects. The plots show Parameter Y to be relatively unvarying across the NREM period, with the exception of subject 1, who displays an increasing trend to Parameter Y.

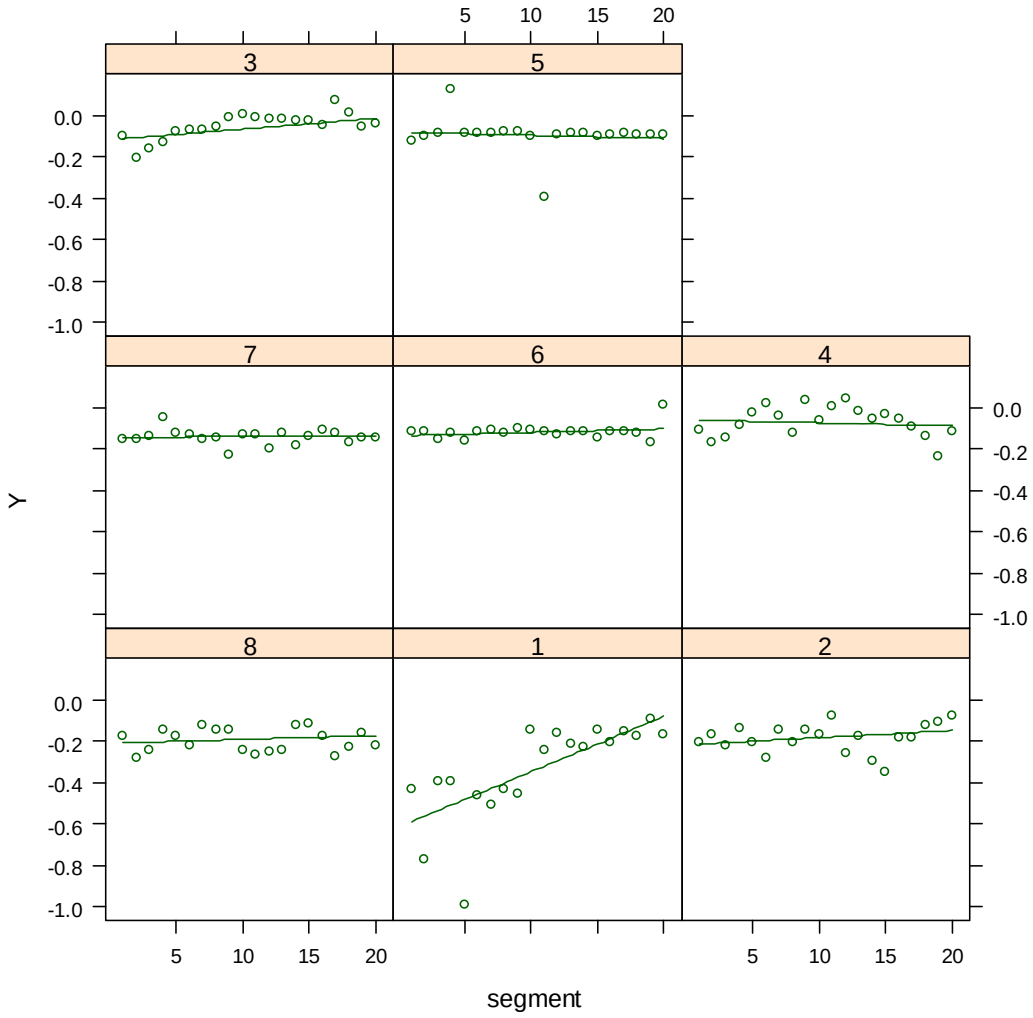


Figure 5.8 Individual subject data for Parameter Y

Parameter Z

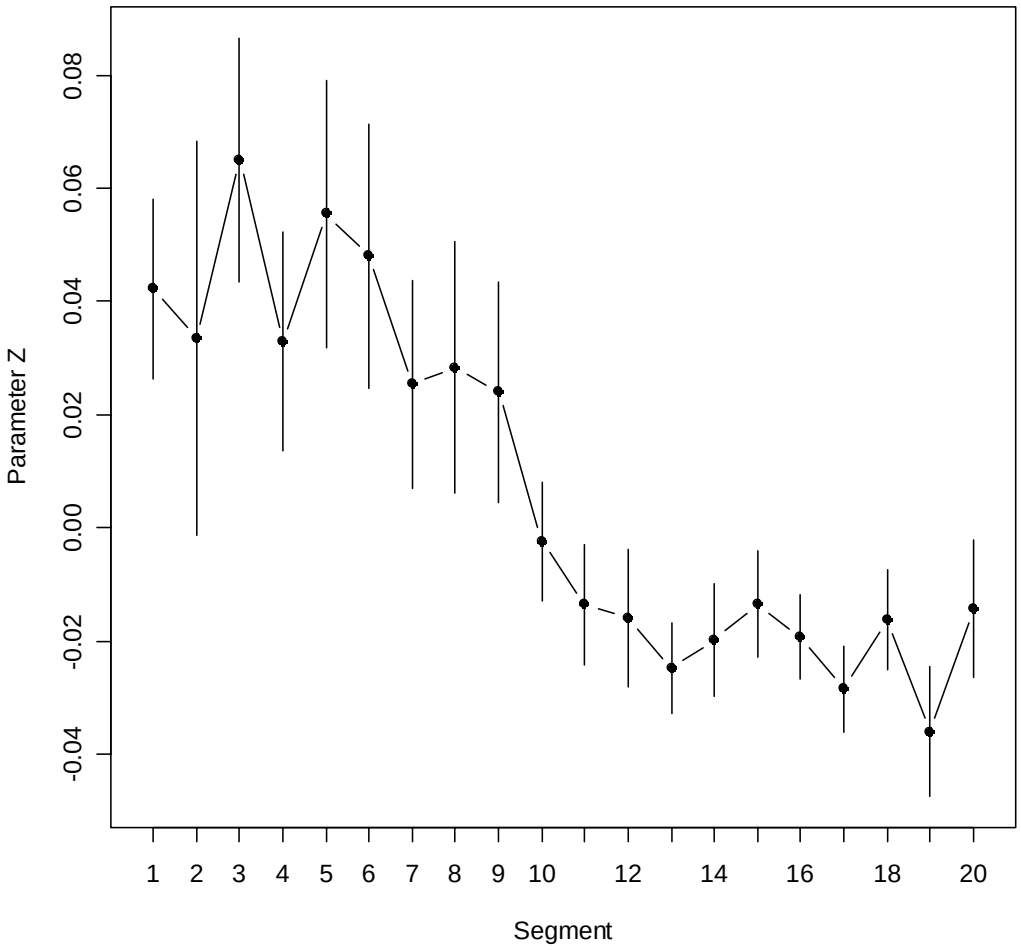


Figure 5.9 Dynamics of internal feedback between dorsal and ventral thalamus (parameter Z) across the first NREM sleep cycle

Parameter Z reflects the internal feedback between the thalamic reticular nuclei and the thalamic relay nuclei. There was a significant linear decrease in this parameter across the first NREM sleep cycle ($p = 0.0005$, Output 5.5). (Figure 5.9) It should be noted that Z is defined such that a

greater inhibitory influence upon this intrathalamic circuit should lead to a more positive value of the parameter Z.

```
Fixed effects: Z ~ segment

                  Value   Std.Error  DF   t-value p-value

(Intercept)  0.06310075 0.021687206 132  2.909584  0.0042

segment     -0.00477014 0.001332612 132 -3.579542  0.0005
```

Output 5.5

Figure 5.10 shows an augmented predictions plot. Open circles depict observed values for Parameter Z for each individual subject. The lines show the individual-subject predicted linear effect of segment, augmented by the estimated random effects. The plots show a downward trend to be consistent across subjects. There appeared to be an initial peak in Parameter Z in the first 2-5 segments, however quadratic effects of SEGMENT were not statistically significant.

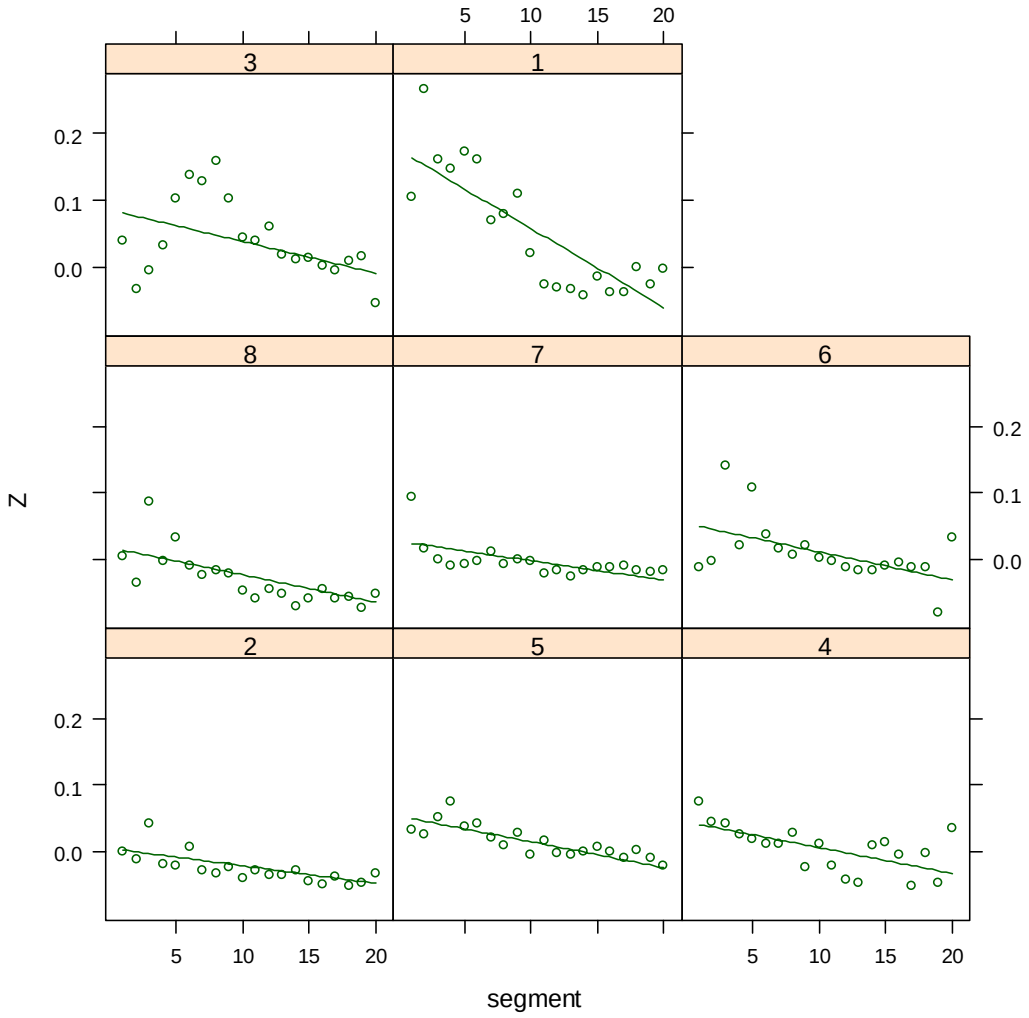


Figure 5.10 Individual subject data for Parameter Z

Gain parameter G_{ee}

Figure 5.11 shows the time course of the parameter expressing excitatory gain in the pyramidal cells (G_{ee}). There is a significant linear negative trend for SEGMENT, as well as a significant quadratic trend (both $p < 0.0001$).

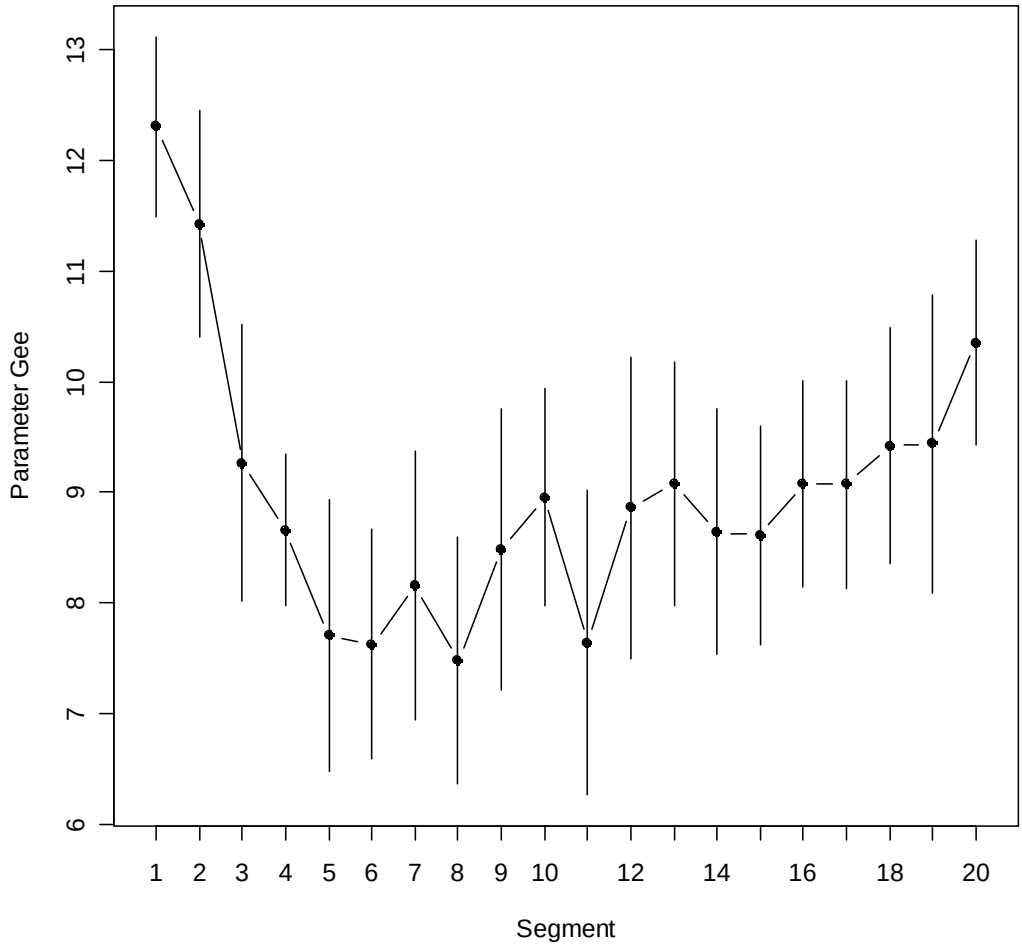


Figure 5.11 Dynamics of Parameter Gee with increasing depth of NREM sleep

```
Fixed effects: Gee ~ segment + I(segment^2)

                 Value Std.Error  DF   t-value p-value

(Intercept)  11.636912 0.9384275 150 12.400438       0

segment      -0.676717 0.1214879 150 -5.570239       0
```

```
I(segment^2)  0.031221 0.0055460 150  5.629380       0
```

Output 5.6

Gain parameter G_{srs}

Figure 5.12 shows the time course of parameter G_{srs} across the first NREM sleep cycle. It should be noted, as shown in cross-sectional data, that G_{srs} is a negative quantity. It is strongly negatively correlated with parameter *Z*, owing to the mathematical relationship that defines *Z*. There was a significant linear trend (p = 0.002) indicating reduced (towards zero) inhibitory influence of the thalamic reticular nucleus with increasing depth of NREM sleep (Output 5.7). Quadratic and cubic effects of SEGMENT were not significant.

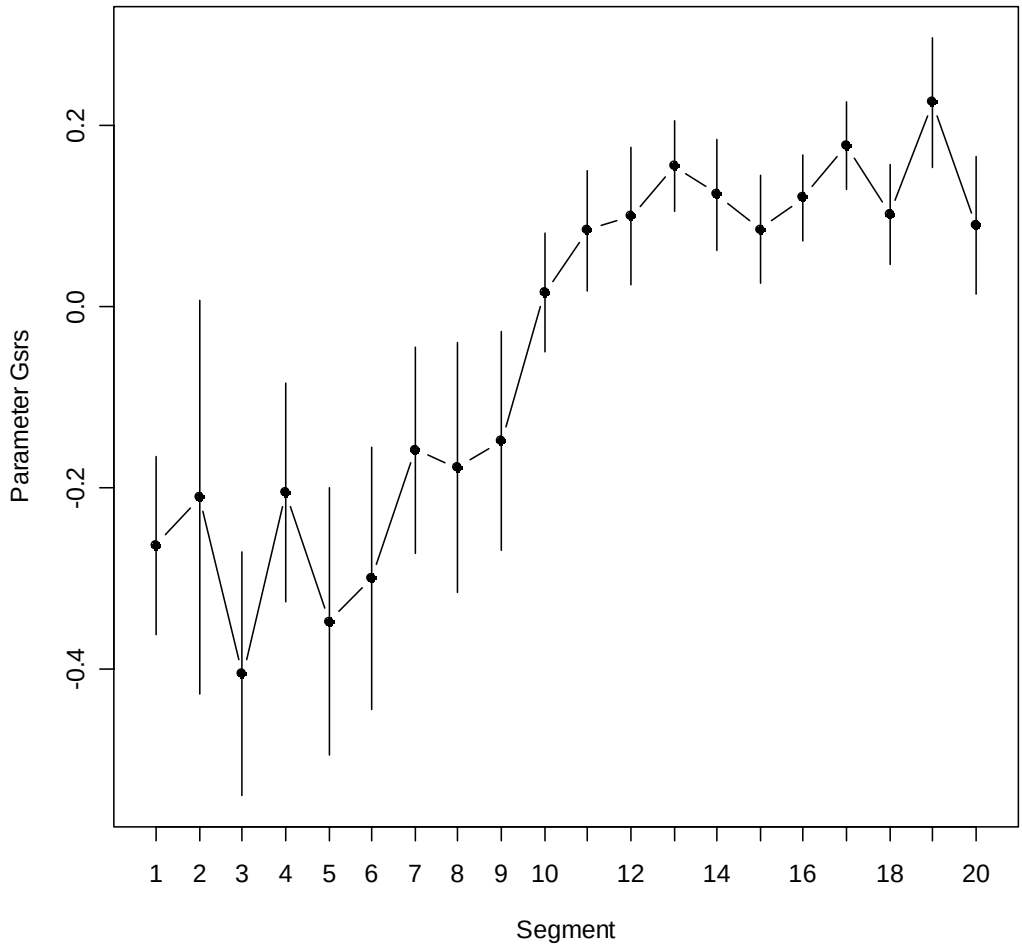


Figure 5.12 Dynamics of internal feedback between dorsal and ventral thalamus (parameter Gsrs) across the first NREM sleep cycle

```
Fixed effects: Gsrs ~ segment + I(segment^2)

                  Value  Std.Error  DF   t-value p-value

(Intercept)  -0.4201272 0.13088071 150 -3.210001  0.0016

segment       0.0460379 0.01426444 150  3.227456  0.0015

I(segment^2) -0.0007688 0.00058325 150 -1.318104  0.1895
```

Output 5.7

5.5 Conclusion

Consistent with previous studies, EEG delta power rises and sigma power falls across a NREM sleep cycle in subjects with typical sleep architecture. Over the same period, parameters derived from EEG modelling predict a gradual decrease in net cortical excitation (Parameter *X*), but a decrease in the inhibitory feedback loop within the thalamus (Parameter *Z*). The fall in cortical activity as sleep deepens during a NREM sleep cycle is consistent with expectations. However, the finding of a decrease in inhibitory feedback in the intra-thalamic loop appears to contradict the theorised role of the reticular nucleus in modulating the corticothalamic interaction in NREM sleep, and our interpretation of underlying processes that the model parameters represent.

This study also confirmed that spectra generated from averaging segments of a NREM sleep cycle can be fitted to a physiologically-based mathematical model. Parameters derived from fitting the observed spectra to the mathematical model show significant and biologically-plausible trends across time. The unexpected fall in parameter *Z* raises the possibility that the current interpretation for the process represented by parameter *Z* (and G_{srs}) is incorrect, or that this process reflects a secondary, antagonistic, process to inhibitory thalamic reticular nucleus activity.

It is clear that the functional interpretation of these EEG model parameters needs further study. Nevertheless, values of the parameters appear to vary with the expected fluctuation in sleep depth. Further investigation would be warranted to confirm these patterns, and perhaps confirm the neurophysiological significance of these parameter fluctuations.

If validated, the model might have the potential to provide insight into the interactions between the cortex and thalamus with deepening NREM. There are many potential applications for this

technique as an indicator of sleep depth that would need to be evaluated. There are other ways in which a sleep EEG recording can be analysed. Event-related analyses from stimuli such as apneas, noise, or periodic limb movements might be employed to estimate of the depth or quality of sleep (and hence the arousal threshold) at the termination of the apnea, to greater understand the differences between events leading to the different types of arousals that occur as a consequence of these stimuli (for example cortical arousals observed on EEG, as opposed to other types of arousals detected with autonomic measures). The clinical relevance of events distinguished by the EEG-parameter-based scoring could be assessed in reference to the severity of daytime consequences arising, and in examining the changes with treatment of the sleep disorder or removal of the experimental disruption.

If validated as a marker of sleep depth, metrics derived from the EEG model parameters may provide an alternative method to stage or grade sleep that does not rely on a limited number of arbitrarily-defined sleep stages. Aggregate measures may be devised to provide an all-night measure of the quality of sleep that might predict daytime function or the effect of therapy for sleep apnea, or allow investigation of sleep quality in other sleep disorders such as insomnia and restless legs syndrome. These novel applications bear further investigation.

6 Overall conclusion

6.1 Summary of results

This study confirmed that tests within a comprehensive neurocognitive battery were sensitive to the effects of OSA, in subjects recruited from the community. It also confirmed that a questionnaire screening tool for sleep apnea could identify a subset of an otherwise-healthy normative population with neurocognitive impairment and differences on evoked response potential testing, relative to the rest of the normative population. When applied to a clinical population presenting with suspected OSA, outcomes from this neurocognitive battery, enriched with other questionnaires and tests of particular interest in sleep apnea, were able to be grouped by factor analysis five summary outcomes relevant to sleepiness and daytime function.

As expected from the literature, functional outcomes did not show strong relationships to polysomnographic measures of OSA severity. When related to parameters from fitting the EEG model, weak-to-moderate correlations were found with the summary factor outcomes. A greater inhibitory effect of the thalamic reticular nucleus on the thalamocortical relay nuclei (parameter *Z*) was associated with worse performance on tests of executive function. This finding was consistent with the hypothesis, based on a current theory of thalamocortical interactions. The direction of effect with regard to executive function and other measures was unexpected, showing worse performance on executive function with greater eyes-closed alpha power (generally representing greater alertness) or lower polysomnographically-determined severity of OSA.

EEG model parameters reflecting corticothalamic interactions showed significant changes with increasing depth of NREM sleep. A gradual fall in cortical excitation (parameter *X*) during this period was consistent with greater depth of sleep. In contradiction to current theory concerning the corticothalamic interaction during NREM sleep, with increasing sleep depth there was reduced inhibitory feedback between the thalamic reticular nucleus and the thalamocortical relay nuclei, represented by EEG model parameter Z.

6.2 Potential significance

The development of a readily accessible, theoretically-based, measure of sleepiness that related to important functional outcomes would have great significance. **Key domains of function** were assayed, including sleepiness, mood, executive function and driving ability. The self-reported sleepiness and executive function outcomes that correlated significantly with the EEG model parameters are known to be affected in sleep loss, as well as in OSA. Although OSA was used as a disease model of sleepiness, this EEG-model based measure may potentially be useful in the measurement of sleepiness in healthy individuals, such as in work situations requiring high levels of alertness, or to study the effects of shiftwork or monitor fitness for duty. It may also have utility in other sleep disorders such as narcolepsy and insomnia in which sleep loss or sleepiness may be occur.

The EEG model is **linked to a current theory** concerning the modulation of vigilance and sleep (Steriade, 2000), and thus predictions can be made regarding underlying neurophysiological interactions, based on values of the fitted parameters. This may provide a greater understanding of these underlying processes, and also provide a means for the model to be validated, if methods

of assaying the neuronal interactions or the levels of the putative 'sleep factor' become available. Recently-published data also suggest that the model parameters reflected response to stimulant therapy, and were associated with elements of impaired function (such as memory loss) of concern to the individual (Alexander et al., 2006; Rowe, Robinson, & Gordon, 2005).

The technology employed to obtain the recordings is **accessible** to the sleep clinician or researcher. Recordings were made using standard equipment found in a clinical sleep laboratory. Data acquisition was relatively simple, with the subject resting with their eyes closed for two minutes. Processing of the EEG signal and fitting of the EEG model were performed on software running within a personal computer.

Potential uses of such a technique are broad. As a **research tool**, the EEG model measure may present a relatively rapid means to accurately phenotype an individual's vulnerability to sleep loss, thus facilitating research into genetic factors mediating inter-individual variability. It may provide improved methods to measure sleep and sleep quality in sleep disorders such as insomnia and restless legs syndrome. It may assist in the development of new therapies, by providing insights into the neurobiological effects of new drugs, and allowing rapid measurement of desired and undesired central nervous system effects. It may also have application in analyses of events during sleep, such as the sleep-related breathing events and their termination.

As a **clinical tool**, the EEG model-based measure may help in distinguishing people with sleep disorders such as OSA who are at risk due to their sleepiness from those who are not abnormally sleepy. The measure might also be applied in determining that an adequate response to therapy has occurred, or assisting with dose titration of wakefulness promoting treatments. In occupational settings, the measurement has the potential to provide a relatively quick assessment

of the level of alertness that did not rely on self-report or the retrospective detection of errors, before commencing duty, or perhaps during a work shift.

6.3 Limitations and unresolved issues

> *"In applying mathematics to subjects such as physics or statistics we make tentative assumptions about the real world which we know are false but which we believe may be useful nonetheless* (Box, 1976)*."*

In comparing the results with expectations, one must acknowledge that the development of the EEG model, and its evaluation as a measure of sleepiness is at a preliminary stage. The EEG model is a simplification, in order to facilitate greater understanding of underlying complex brain processes. An appreciation of the properties of a model, and its limitations, is important to interpreting its predictions. Greater complexity in a model need not necessarily assist our understanding. Ultimately, the appropriateness of the degree of simplification can be judged by the degree to which the model, and predictions arising from it, are useful. In this work, the usefulness of the EEG model as a measure of sleepiness was explored, with the main comparison between the model and other measures of sleepiness conducted with cross-sectional data. Limitations of the study findings relate to the deployment of the test and the interpretation of the new measure.

Deployment of the test

In regard to the deployment of the test, there are **technical considerations** relating to the EEG model fitting, and practical considerations concerning the application of the EEG model measure

in the field. Further research is required to determine the extent to which the current measure of model fit (the chi-squared statistic) influences the predictive ability of the fitted parameters, and whether there are better criteria to determine model fit, perhaps by weighting key frequency ranges known to vary with sleepiness such as the theta and alpha bands. The current procedures to obtain the model parameters from an EEG spectrum require 10-40 minutes of computing time, not including procedures for artefact reduction and the need to prepare the spectral data in the appropriate format.

The application of the measure is in its infancy. There may be potential for improvements in model conceptualization, and in the EEG acquisition, for example with possibly better methods to represent the parameters (for example, the *X, Y* and *Z* parameters are themselves mathematical transformations of a set of gain parameters derived directly from the original model), better methods for artefact correction, simplification in the placement of EEG leads. Further research to determine the minimum number of recording leads necessary, or exploring alternative lead attachment methods such as adhesive patches or gel caps may simplify EEG acquisition. Devices using a forehead applied EEG electrode patch and a simple user interface are now in routine clinical use for anaesthetic monitoring (Sleigh, Andrzejowski, Steyn-Ross, & Steyn-Ross, 1999), and illustrate the potential for a practical EEG-based measure.

The suitability of any test of sleepiness depends on the purpose to which it is to be applied. There are some important **practical considerations**. If used for vigilance monitoring in the field, then a robust instrument capable of real-time analysis would be desired. A proof of concept study comparing EEG with other physiological measures in vigilance monitoring during a supervised real-world car drive found that the potential for artefact from muscle tone, movement and

external interference, and the need to attach electrodes left the EEG at a practical disadvantage, relative to eye blink measures (Papadelis et al., 2007).

The EEG model was fitted to eyes closed EEG spectra. For the purposes of this initial evaluation, eyes closed recordings offered the advantage of less contamination by eye blink artefact, and the availability of a greater amount of adequate data for generation of the power spectrum. The presence of a more prominent alpha peak in an eyes closed spectrum may also facilitate model fitting (Rowe, Robinson, & Gordon, 2005). If to be applied as a continuous vigilance monitor, eyes-open recordings would be less intrusive. EEG power changes between eyes open and eyes closed conditions have been previously described, and in the eyes closed condition overall power is higher (Leproult et al., 2003), but the potential for sleep onset during high levels of sleepiness may complicate analysis (T. Akerstedt & Gillberg, 1990). Differences in model parameters between the two conditions have also been described in the normative dataset (Rowe, Robinson, & Rennie, 2004). The parameter G_{srs} reflecting the intrathalamic interaction was significantly reduced in magnitude in the eyes open condition. It is not yet known how eyes closed model parameters would relate to sleepiness measures, however EEG power in various bands have been shown to reflect sleepiness in either condition.

If used only intermittently, such as in research settings to investigate drug effects or explore neurophysiology, the current use of sleep laboratory equipment is practical. Better techniques to reduce artefact and improve model fitting would still be advantageous.

Interpretation of the test.

> *"Far better an approximate answer to the* right *question, which is often vague, than an exact answer to the* wrong *question, which can always be made precise* (Tukey, 1962)*."*

It remains to be clarified **what the model parameters are measuring**. In interpreting the direction of effect from the studies conducted, it is unclear whether the changes in EEG model parameters represent a primary effect or whether they represent compensation, feedback inhibition or antagonism of a primary process. It may also be possible that in some situations a process takes on the primary role, while in other situations it is secondary: this may explain the mutually-contradictory findings in regard to parameter Z between the study of OSA patients during wake, and the second study involving NREM sleep.

More research is needed to describe **state and trait characteristics** of the model parameters, and the degree of inter-individual variation. The cross-sectional design of the study relating sleepiness to the EEG model, does not allow determination of the extent to which the parameters we have identified as potential measures of sleepiness reflect state or trait properties of the brain. We would expect, as with any other measure, a combination of both short-term and more long term influences (Johns, 1994, 1998). Trait-like characteristics of the EEG have been previously reported (Hoptman & Davidson, 1998), as has inter-individual variability in performance deterioration with sleep loss (H. P. Van Dongen et al., 2004), and inter-individual variability in the relationship between EEG and sleepiness-related measures (Torsvall & Akerstedt, 1988). Trait aspects of the EEG have been investigated in the context of predicting later cognitive decline (van der Hiele et al., 2007), or marking the presence of a psychiatric disorder (Rowe, Robinson, Rennie et al., 2004). Thus any trait characteristics in EEG model parameters may

reflect genetic variation in the performance decrement and EEG response to sleep loss (De Gennaro, Ferrara, Vecchio, Curcio, & Bertini, 2005; Franken, Chollet, & Tafti, 2001; Viola et al., 2007), the neuronal and synaptic structure within the brain of that individual determining their thalamocortical interactions and their performance ability, or perhaps irreversible injury to the brain from the intermittent hypoxia of OSA (Murray, 2007). The experiment showing changes in model parameters concomitant with increasing depth of sleep support the presence of a state-dependent fluctuation, and test-retest reliability of the parameters under stable resting conditions has been studied (van Albada et al., 2005). Further research is still needed to clarify the contribution of trait effects in parameters relevant to sleepiness.

If the acknowledged significant inter-subject variability in EEG power can be extrapolated to the EEG model parameters, the measure may be more sensitive in **detecting changes within an individual**, rather than between individuals. This does not preclude its use in vigilance monitoring, as long as sufficient training or baseline data is collected. A large proportion of studies examining EEG measures of sleepiness have employed within-subjects analyses from sleep deprivation experiments. These experiments are efficient given the greater statistical power with a repeated measures design when there is substantial between subjects variability, and also given the broad range of sleepiness levels that can be induced with sleep deprivation.

Correlations between EEG and functional outcomes were weak-to-moderate, with approximately 10% shared variance (r^2) at best. However the **lack of a strong relationship** should be put in context with literature showing at most moderate correlations between measures of sleepiness, and even between items within a single questionnaire such as the ESS (r^2 10.9-29.2) (Johns, 2002). Correlations between measures of OSA severity and functional outcomes have also been of a similar magnitude (Adams et al., 2001).

An important with the interpretation of the results is the **lack of consistency of results across the majority of outcomes** within the cross-sectional study. This might suggest the possibility of a type 1 error. It would have been more desirable to have demonstrated relationships between the EEG parameters and the majority of the five summary outcomes. The findings need to be replicated, with more profound fluctuations in sleepiness induced by sleep deprivation.

There is also a **lack of consistency of findings between the studies.** Just as one of our studies is in keeping with our hypothesis while the other is not, the published studies in the two clinical groups of ADHD and memory loss have also showed contradictory findings with G_{srs} being negatively-correlated with greater (drug-induced) alertness in one, and positively correlated with greater neurocognitive impairment in the other. These differences might be explained by differing modulation of thalamocortical interactions in different disease processes, or the different states of wake and sleep.

Another issue is in regard to the **differences between OSA and other experimental models of sleepiness**, including sleep deprivation and noise-induced sleep fragmentation. While sleep is fragmented in OSA, there remain uncertainties as to the effect of intermittent hypoxia, the presence of possible irreversible brain injury, or the equally uncertain degree to which sufferers of this chronic condition might compensate or adapt.

As a measure of fitness to work or drive, the EEG model should to be validated against relevant **real-world outcomes** such as actual accidents or work-related serious errors. Simulated work or simulated driving tasks possess ecological validity, however motivational factors are likely to be different in a subject who knows that they are not truly at risk. EEG monitoring has been successfully applied during actual work in train drivers (Torsvall & Akerstedt, 1987), process workers (T. Akerstedt et al., 1991), truck drivers (Kecklund & Akerstedt, 1993), commercial

pilots (Signal, Gale, & Gander, 2005), medical residents (T. Akerstedt, Arnetz, & Anderzen, 1990; Lockley et al., 2004), and as well as during non-work-related driving (Brookhuis & De Waard, 1993; Papadelis et al., 2007). In many of these studies, data regarding work or driving-related errors have been collected. The determination that an individual is sleepy need not necessarily imply that they are a hazard to safety (Kecklund & Akerstedt, 1993). Apart from the reported large interindividual differences in the relationship between EEG and sleepiness-related measures (Torsvall & Akerstedt, 1988), there may need to be other adverse factors coinciding before an accident occurs. Furthermore, results based on group data may not necessarily be applicable to an individual.

6.4 Research agenda

> *"Whatever device and analysis technique is used, before there is wide-scale implementation there will need to be much greater numbers of subjects tested, both sleep deprived and not sleep deprived. It is important to gain a better understanding of the false positive and negative value in various populations, especially given the great person to person variation in the pattern of EEG changes in sub-vigil states* (Oken & Salinsky, 2007)*."*

Further information needs to be gathered in regard to the psychometric properties of this EEG model-based measure, as part of its **validation** (page 50). This validation process will crucial to **understanding the properties and limitations of the model**, and the how the parameters should be interpreted. To separate state from trait characteristics in the EEG model parameters,

test-retest reliability should be ascertained within a day, and also between longer durations of weeks or months. The effect of interventions known to alter the level of sleepiness, such as sleep deprivation, or hypnotic or stimulant drugs, should also be conducted to assess the degree to which the model parameters respond to these interventions. Sleep deprivation paradigms have been widely used in the literature, and provide a method to obtain measurements at normal alertness and very high levels of sleepiness. In sleep apnea, the effect of treatment could also be studied, with longer term follow-up studies to determine the presence changes on the EEG that are not reversed with treatments such as continuous positive airway pressure. These comparisons between EEG and sleepiness-related outcomes may conveniently be nested within other studies, including randomised trials of pharmacotherapy. A sleep deprivation study including both healthy and OSA subjects may help explore whether or not the processes represented by the EEG model are functionally similar between OSA and healthy individuals.

The need for further research and development to advance the **technical aspects of the measure** has been discussed in the previous section. Work may be directed towards improving portability and robustness of the testing equipment, improving data acquisition, improving analysis and model fitting, and providing simple methods to interpret results of testing.

The **role of the EEG model** as a measure of sleepiness needs to be determined, whether as a continuous vigilance monitor, a clinical measure to identify a sleepy subgroup, or as a research tool for neurophysiology. Its potential applications will determine the process to be taken with validation. If the measure is to be used as a vigilance monitor, it will need to be **compared against other tests** such as those based on eyelid movement, with consideration of the accuracy of the measurement as well as the convenience of application (D F Dinges et al., 1998).

From a wider perspective, measuring sleepiness is one part of an **overall strategy** to reduce the adverse effects of sleepiness and improve health status. Other elements in this strategy would include better community awareness of the causes of sleepiness, improved knowledge of countermeasures and mitigating techniques such as naps, better shift patterns, possibly pharmacotherapy, and better use of administrative, legislative, and technological systems to reduce the risk of errors, to detect errors early, and if possible correct them. Alternatively, there is a role for technological solutions to reduce the need for humans to work in situations requiring sustained vigilance. Analyses of the causes of major catastrophes which have been attributed to sleepiness, reveal that a single incident requires the failure of more than a single factor (National Transportation Safety Board, 1990).

7 Abbreviation and term glossary

Alpha: The alpha frequency band of the electroencephalogram. (range 8-12.5 Hz, or 8-12 Hz as used in BRID battery).

AHI: apnea hypopnea index: in polysomnography, the number of apneas and hypopneas per hour of sleep. Also referred to as the RDI, respiratory disturbance index.

Beta: 13.0-39.5 Hz frequency band of the electroencephalogram. (12-35 Hz used in BRID battery).

BRID: Brain Resource International Database, BRID battery

CPP, Cognitive Performance Profile: The neurocognitive testing component of the Brain Resource International Database battery.

DASS: Depression Anxiety Stress Scale.

Delta: 0.5-3.5 Hz frequency band of the electroencephalogram (including the 3.5 to less than 4.0 Hz range).

EC: eyes-closed (EEG recording)

EEG: electroencephalogram

EO: eyes-open (EEG recording)

EOG: electrooculogram

ERP: evoked response potential

ESS: Epworth Sleepiness Scale

FOSA: Functional Outcomes of Sleep Scale

MAPI: Multivariable Apnea Prediction Index.

NREM: non-rapid eye-movement (sleep).

OSA: Obstructive Sleep Apnea.

PSG: polysomnogram, polysomnography.

PSA: power spectral analysis, here relating to analyses to separate an EEG signal into its component frequencies.

RDI: respiratory disturbance index, used synonymously with the apnea hypopnea index (AHI)

REM: rapid eye-movement (sleep).

STAI: (Spielberger) State-Trait Anxiety Inventory.

SWA: slow-wave activity, used synonymously with delta power during sleep.

Theta: 4.0-7.5 Hz frequency band of the electroencephalogram inclusive.

UARS: upper airway resistance syndrome (Guilleminault et al., 1991; Philip, Stoohs, & Guilleminault, 1994).

8 Index

9 Acknowledgements

I have benefited greatly from the support and encouragement of others. I would like to thank my supervisor, Ron Grunstein, for his tireless encouragement and feedback. I would like to thank my co-supervisor Peter Robinson for the crash course on the EEG mathematical model, and Chris Rennie who answered my numerous perplexed queries with humour and who amended his EEG analysis programs to the vagaries of our hardware. For the cross-sectional study I would like to thank Delwyn Bartlett, who prepared the cross-sectional study and doubled as a counsellor, and Qiao Yang, who taught me the technical aspects of the data collection and helped with so many things with her boundless enthusiasm. Thank you, Naomi Rogers, for advising on future directions of study, and commenting on this text. Thanks to all the colleagues, friends, and family who never ceased to enquire as to the progress of my candidature.

Most importantly, I would like to thank my family: my parents, my wife Yin Wan, and relative latecomers to the endeavour Weilin and Weiming, who were nevertheless an inspiration.

10 References

Achermann, P., & Borbely, A. (1994). Simulation of daytime vigilance by the additive interaction of a homeostatic and a circadian process. *Biological Cybernetics, 71*(2), 115-121.

Adams, N., Strauss, M., Schluchter, M., & Redline, S. (2001). Relation of Measures of Sleep-Disordered Breathing to Neuropsychological Functioning. *American Journal of Respiratory and Critical Care Medicine, 163*(7), 1626-1631.

Aeschbach, D., & Borbely, A. A. (1993). All-night dynamics of the human sleep EEG. *Journal of Sleep Research, 2*(2), 70-81.

Aeschbach, D., Matthews, J. R., Postolache, T. T., Jackson, M. A., Giesen, H. A., & Wehr, T. A. (1997). Dynamics of the human EEG during prolonged wakefulness: evidence for frequency-specific circadian and homeostatic influences. *Neuroscience Letters, 239*(2-3), 121-124.

Aeschbach, D., Matthews, J. R., Postolache, T. T., Jackson, M. A., Giesen, H. A., & Wehr, T. A. (1999). Two circadian rhythms in the human electroencephalogram during wakefulness. *American Journal of Physiology, 277*(6 Pt 2), R1771-1779.

Aguirre, M., & Broughton, R. J. (1987). Complex event-related potentials (P300 and CNV) and MSLT in the assessment of excessive daytime sleepiness in narcolepsy-cataplexy. *Electroencephalography & Clinical Neurophysiology, 67*(4), 298-316.

Akerstedt, T. (1995). Work hours, sleepiness and the underlying mechanisms. *Journal of Sleep Research, 4*(S2), 15-22.

Akerstedt, T. (1998). Sleepiness. *Sleep Medicine Reviews, 2*(1), 1-2.

Akerstedt, T., Arnetz, B. B., & Anderzen, I. (1990). Physicians during and following night call duty--41 hour ambulatory recording of sleep. *Electroencephalography & Clinical Neurophysiology., 76*(2), 193-196.

Akerstedt, T., & Folkard, S. (1995). Validation of the S and C components of the three-process model of alertness regulation. *Sleep., 18*(1), 1-6.

Akerstedt, T., & Gillberg, M. (1990). Subjective and objective sleepiness in the active individual. *International Journal of Neuroscience., 52*(1-2), 29-37.

Akerstedt, T., Kecklund, G., & Knutsson, A. (1991). Manifest sleepiness and the spectral content of the EEG during shift work. *Sleep., 14*(3), 221-225.

Akerstedt, T., Peters, B., Anund, A., & Kecklund, G. (2005). Impaired alertness and performance driving home from the night shift: a driving simulator study. *Journal of Sleep Research, 14*(1), 17-20.

Aldrich, M. S. (1994). Parkinsonism. In W. C. Dement, T. Roth & M. H. Kryger (Eds.), *Principles and Practice of Sleep Medicine* (2nd ed.). Philadelphia: Saunders.

Alexander, D. M., Arns, M. W., Paul, R. H., Rowe, D. L., Cooper, N., Esser, A. H., et al. (2006). EEG markers for cognitive decline in elderly subjects with subjective memory complaints. *Journal of Integrative Neuroscience, 5*(1), 49-74.

Aloia, M. S., Arnedt, J. T., Davis, J. D., Riggs, R. L., & Byrd, D. (2004). Neuropsychological sequelae of obstructive sleep apnea-hypopnea syndrome: a critical review. *Journal of the International Neuropsychological Society, 10*, 772-785.

American Academy of Sleep Medicine Task Force. (1999). Sleep-related breathing disorders in adults: recommendations for syndrome definition and measurement techniques in clinical research. *Sleep, 22*(5), 667-689.

American Electroencephalographic Society. (1991). American Electroencephalographic Society guidelines for standard electrode position nomenclature. *Journal of Clinical Neurophysiology., 8*(2), 200-202.

Amzica, F., & Steriade, M. (1997). The K-complex: its slow (<1-Hz) rhythmicity and relation to delta waves.[see comment]. *Neurology, 49*(4), 952-959.

Anderson, C., & Horne, J. A. (2003). Prefrontal cortex: Links between low frequency delta EEG in sleep and neuropsychological performance in healthy, older people. *Psychophysiology, 40*(3), 349-357.

Aserinsky, E., & Kleitman, N. (1955). Two Types of Ocular Motility Occurring in Sleep. *Journal of Applied Physiology, 8*(1), 1-10.

Aston-Jones, G., Chen, S., Zhu, Y., & Oshinsky, M. L. (2001). A neural circuit for circadian regulation of arousal. *Nature Neuroscience, 4*(7), 732-738.

Austroads, & National Road Transport Commission. (2003). *Assessing Fitness to Drive for Commercial and Private Vehicle Drivers: Medical Standards for Licensing and Clinical Management Guidelines* (3rd ed.). Sydney: Austroads.

Ayas, N. T., White, D. P., Manson, J. E., Stampfer, M. J., Speizer, F. E., Malhotra, A., et al. (2003). A prospective study of sleep duration and coronary heart disease in women. *Archives of Internal Medicine, 163*(2), 205-209.

Baddeley, A., Emslie, H., & Nimmo-Smith, I. (1993). The Spot-the-Word test: a robust estimate of verbal intelligence based on lexical decision. *British Journal of Clinical Psychology, 32*(Pt 1), 55-65.

Banks, S., Barnes, M., Tarquinio, N., Pierce, R. J., Lack, L. C., & McEvoy, R. D. (2004). The maintenance of wakefulness test in normal healthy subjects. *Sleep, 27*(4), 799-802.

Banks, S., Catcheside, P., Lack, L., Grunstein, R. R., & McEvoy, R. D. (2004). Low levels of alcohol impair driving simulator performance and reduce perception of crash risk in partially sleep deprived subjects. *Sleep, 27*(6), 1063-1067.

Banks, S., Catcheside, P., Lack, L. C., Grunstein, R. R., & McEvoy, R. D. (2005). The Maintenance of Wakefulness Test and Driving Simulator Performance. *Sleep, 28*(11), 1381-1385.

Bastuji, H., & Garcia-Larrea, L. (1999). Evoked potentials as a tool for the investigation of human sleep. *Sleep Medicine Reviews, 3*(1), 23-45.

Bearpark, H., Grunstein, R., Touyz, S., Channon, L., & Sullivan, C. (1988). Cognitive and Psychological Dysfunction in Sleep-Apnea before and after Treatment with CPAP. *Australian and New Zealand Journal of Medicine, 18*(3), 529-529.

Beebe, D. W., & Gozal, D. (2002). Obstructive sleep apnea and the prefrontal cortex: towards a comprehensive model linking nocturnal upper airway obstruction to daytime cognitive and behavioral deficits. *Journal of Sleep Research., 11*(1), 1-16.

Beebe, D. W., Groesz, L., Wells, C., Nichols, A., & McGee, K. (2003). The neuropsychological effects of obstructive sleep apnea: a meta-analysis of norm-referenced and case-controlled data. *Sleep, 26*(3), 298-307.

Bennett, L. S., Stradling, J. R., & Davies, R. J. (1997). A behavioural test to assess daytime sleepiness in obstructive sleep apnoea. *Journal of Sleep Research, 6*(2), 142-145.

Benton, A., & Hamsher, K. (1989). Multilingual aphasia examination. Iowa City, IA: AJA Associates.

Besset, A., Villemin, E., Tafti, M., & Billiard, M. (1998). Homeostatic process and sleep spindles in patients with sleep-maintenance insomnia: effect of partial (21 h) sleep deprivation. *Electroencephalography & Clinical Neurophysiology., 107*(2), 122-132.

Bjerner, B. (1949). Alpha depression and lowered pulse rate during delayed reactions in a serial reaction task. *Acta Physiologica Scandinavica, 19*(1), 1-93.

Black, J. (2003). Sleepiness and residual sleepiness in adults with obstructive sleep apnea. *Respiratory Physiology & Neurobiology, 136*(2-3), 211-220.

Blake, H., & Gerard, R. W. (1937). Brain potentials during sleep. *American Journal of Physiology., 119*, 692-703.

Bliwise, D. L. (2001). Is the measurement of sleepiness the Holy Grail of sleep medicine? *American Journal of Respiratory & Critical Care Medicine., 163*(7), 1517-1519.

Bonnet, M. H., & Arand, D. L. (1998). Sleepiness as measured by modified multiple sleep latency testing varies as a function of preceding activity. *Sleep., 21*(5), 477-483.

Bonnet, M. H., & Arand, D. L. (2005). Impact of motivation on multiple sleep latency test and maintenance of wakefulness test measurements. *Journal of Clinical Sleep Medicine, 1*(4), 386-390.

Borbely, A. A. (1982). A two process model of sleep regulation. *Human Neurobiology., 1*(3), 195-204.

Borbely, A. A., & Achermann, P. (1999). Sleep homeostasis and models of sleep regulation. *Journal of Biological Rhythms., 14*(6), 557-568.

Borbely, A. A., Baumann, F., Brandeis, D., Strauch, I., & Lehmann, D. (1981). Sleep deprivation: effect on sleep stages and EEG power density in man. *Electroencephalography & Clinical Neurophysiology., 51*(5), 483-495.

Borbely, A. A., & Tobler, I. (1989). Endogenous sleep-promoting substances and sleep regulation. *Physiological Reviews, 69*(2), 605-670.

Borg, G. (1990). Psychophysical scaling with applications in physical work and the perception of exertion. *Scandinavian Journal of Work, Environment & Health, 16 Suppl 1*, 55-58.

Box, G. E. P. (1976). Science and Statistics. *Journal of the American Statistical Association, 71*(356), 791-799.

Brain Resource Company. (2005). *Brain Resource International Database Methodology Document* (Version 1, November 2005 ed.). Sydney: Brain Resource Company.

Brookhuis, K. A., & De Waard, D. (1993). The use of psychophysiology to assess driver status. *Ergonomics, 36*(9), 1099 - 1110.

Broughton, R. (1982). Performance and evoked potential measures of various states of daytime sleepiness. *Sleep, 5 Suppl 2*, S135-146.

Brown, W. D. P. D. (2005). The Psychosocial Aspects of Obstructive Sleep Apnea. *Seminars in Respiratory & Critical Care Medicine Sleep and Respiration, 26*(1), 33-43.

Cajochen, C., Brunner, D. P., Krauchi, K., Graw, P., & Wirz-Justice, A. (1995). Power density in theta/alpha frequencies of the waking EEG progressively increases during sustained wakefulness. *Sleep., 18*(10), 890-894.

Cajochen, C., Foy, R., & Dijk, D. J. (1999). Frontal predominance of a relative increase in sleep delta and theta EEG activity after sleep loss in humans. *Sleep Research Online., 2*(3), 65-69.

Cajochen, C., Khalsa, S. B., Wyatt, J. K., Czeisler, C. A., & Dijk, D. J. (1999). EEG and ocular correlates of circadian melatonin phase and human performance decrements during sleep loss. *American Journal of Physiology., 277*(3 Pt 2), R640-649.

Cajochen, C., Knoblauch, V., Krauchi, K., Renz, C., & Wirz-Justice, A. (2001). Dynamics of frontal EEG activity, sleepiness and body temperature under high and low sleep pressure. *NeuroReport., 12*(10), 2277-2281.

Cajochen, C., Krauchi, K., von Arx, M. A., Mori, D., Graw, P., & Wirz-Justice, A. (1996). Daytime melatonin administration enhances sleepiness and theta/alpha activity in the waking EEG. *Neuroscience Letters., 207*(3), 209-213.

Campagne, A., Pebayle, T., & Muzet, A. (2004). Correlation between driving errors and vigilance level: influence of the driver's age. *Physiology & Behavior, 80*(4), 515-524.

Campbell, I. G., & Feinberg, I. (2005). Homeostatic sleep response to naps is similar in normal elderly and young adults. *Neurobiology of Aging, 26*(1), 135-144.

Carskadon, M. A., & Dement, W. C. (1979). Effects of total sleep loss on sleep tendency. *Perceptual & Motor Skills, 48*(2), 495-506.

Carskadon, M. A., & Dement, W. C. (1982). The multiple sleep latency test: what does it measure? *Sleep, 5 Suppl 2*, S67-72.

Carskadon, M. A., & Dement, W. C. (1987). Daytime sleepiness: Quantification of a behavioral state. *Neuroscience & Biobehavioral Reviews, 11*(3), 307-317.

Carskadon, M. A., Dement, W. C., Mitler, M. M., Roth, T., Westbrook, P. R., & Keenan, S. (1986). Guidelines for the multiple sleep latency test (MSLT): a standard measure of sleepiness. *Sleep., 9*(4), 519-524.

Casagrande, M., De Gennaro, L., Violani, C., Braibanti, P., & Bertini, M. (1997). A finger-tapping task and a reaction time task as behavioral measures of the transition from wakefulness to sleep: which task interferes less with the sleep onset process. *Sleep, 20*(4), 301-312.

Casagrande, M., Violani, C., Curcio, G., & Bertini, M. (1997). Assessing vigilance through a brief pencil and paper letter cancellation task (LCT): effects of one night of sleep deprivation and of the time of day. *Ergonomics, 40*(6), 613-630.

Chapotot, F., Pigeau, R., Canini, F., Bourdon, L., & Buguet, A. (2003). Distinctive effects of modafinil and d-amphetamine on the homeostatic and circadian modulation of the human waking EEG. *Psychopharmacology., 166*(2), 127-138.

Chervin, R. D. (2003). Epworth Sleepiness Scale? *Sleep Medicine, 4*(3), 175-176.

Chervin, R. D., & Aldrich, M. S. (1999). The Epworth Sleepiness Scale may not reflect objective measures of sleepiness or sleep apnea. *Neurology, 52*(1), 125-131.

Chervin, R. D., Burns, J. W., Subotic, N. S., Roussi, C., Thelen, B., & Ruzicka, D. (2004). Method for detection of respiratory cycle-related EEG changes in sleep-disordered breathing. *Sleep, 27*(1), 110-115.

Cheshire, K., Engleman, H., Deary, I., Shapiro, C., & Douglas, N. J. (1992). Factors impairing daytime performance in patients with sleep apnea/hypopnea syndrome. *Archives of Internal Medicine, 152*(3), 538-541.

Cirelli, C., Bushey, D., Hill, S., Huber, R., Kreber, R., Ganetzky, B., et al. (2005). Reduced sleep in Drosophila Shaker mutants. *Nature, 434*(7037), 1087-1092.

Clark, C. R., Paul, R. H., Williams, L. M., Arns, M., Fallahpour, K., Handmer, C., et al. (2006). Standardized assessment of cognitive functioning during development and aging using an automated touchscreen battery. *Archives of Clinical Neuropsychology, 21*(5), 449-467.

Cluydts, R., De Valck, E., Verstraeten, E., & Theys, P. (2002). Daytime sleepiness and its evaluation. *Sleep Med Rev, 6*(2), 83-96.

Corsi-Cabrera, M., Ramos, J., Arce, C., Guevara, M. A., Ponce-de Leon, M., & Lorenzo, I. (1992). Changes in the waking EEG as a consequence of sleep and sleep deprivation.[comment]. *Sleep., 15*(6), 550-555.

Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation, 10*(7), 1-9.

Cote, K. A., Milner, C. E., Osip, S. L., Ray, L. B., & Baxter, K. D. (2003). Waking quantitative electroencephalogram and auditory event-related potentials following experimentally induced sleep fragmentation. *Sleep, 26*(6), 687-694.

Crenshaw, M. C., & Edinger, J. D. (1999). Slow-wave sleep and waking cognitive performance among older adults with and without insomnia complaints. *Physiology & Behavior., 66*(3), 485-492.

Croft, R. J., Chandler, J. S., Barry, R. J., Cooper, N. R., & Clarke, A. R. (2005). EOG correction: A comparison of four methods. *null, 42*(1), 16-24.

Curcio, G., Casagrande, M., & Bertini, M. (2001). Sleepiness: evaluating and quantifying methods. *International Journal of Psychophysiology, 41*(3), 251-263.

Daniel, R. S. (1967). Alpha and theta EEG in vigilance. *Perceptual & Motor Skills., 25*(3), 697-703.

Darchia, N., Campbell, I. G., Tan, X., & Feinberg, I. (2007). Kinetics of NREM delta EEG power density across NREM periods depend on age and on delta-band designation. *Sleep, 30*(1), 71-79.

Davis, H., Davis, P., Loomis, A., Harvey, E., & Hobart, G. (1938). Human brain potentials during the onset of sleep. *Journal of Neurophysiology., 1*(1), 24-38.

De Gennaro, L., Devoto, A., Lucidi, F., & Violani, C. (2005). Oculomotor changes are associated to daytime sleepiness in the multiple sleep latency test. *Journal of Sleep Research, 14*(2), 107-112.

De Gennaro, L., Ferrara, M., Vecchio, F., Curcio, G., & Bertini, M. (2005). An electroencephalographic fingerprint of human sleep. *Neuroimage, 26*(1), 114-122.

De Valck, E., & Cluydts, R. (2003). Sleepiness as a state-trait phenomenon, comprising both a sleep drive and a wake drive. *Medical Hypotheses, 60*(4), 509-512.

Decary, A., Rouleau, I., & Montplaisir, J. (2000). Cognitive deficits associated with sleep apnea syndrome: a proposed neuropsychological test battery. *Sleep., 23*(3), 369-381.

Delis, D. C., Jacobson, M., Bondi, M. W., Hamilton, J. M., & Salmon, D. P. (2003). The myth of testing construct validity using factor analysis or correlations with normal or mixed clinical populations: lessons from memory assessment.[see comment]. *Journal of the International Neuropsychological Society, 9*(6), 936-946.

Dement, W., & Kleitman, N. (1957). Cyclic variations in EEG during sleep and their relation to eye movements, body motility, and dreaming. *Electroencephalography and Clinical Neurophysiology, 9*(4), 673-690.

Dement, W. C., & Carskadon, M. A. (1982). Current perspectives on daytime sleepiness: the issues. *Sleep., 5*(Suppl 2), S56-66.

Dement, W. C., & Mitler, M. M. (1993). It's time to wake up to the importance of sleep disorders. *JAMA, 269*(12), 1548-1550.

Desai, A. V., Marks, G. B., Jankelson, D., & Grunstein, R. R. (2006). Do Sleep Deprivation and Time of Day Interact with Mild Obstructive Sleep Apnea to Worsen Performance and Neurobehavioral Function? *Journal of Clinical Sleep Medicine, 2*(1), 63-70.

Desai, A. V., Wilsmore, B., Bartlett, D. J., Unger, G., Constable, B., Joffe, D., et al. (2007). The utility of the AusEd driving simulator in the clinical assessment of driver fatigue. *Behavior Research Methods, 39*(3), 673-681.

Dijk, D.-J., Shanahan, T. L., Duffy, J. F., Ronda, J. M., & Czeisler, C. A. (1997). Variation of electroencephalographic activity during non-rapid eye movement and rapid eye movement sleep with phase of circadian melatonin rhythm in humans. *The Journal of Physiology, 505*(3), 851-858.

Dijk, D. J., & Beersma, D. G. (1989). Effects of SWS deprivation on subsequent EEG power density and spontaneous sleep duration. *Electroencephalography & Clinical Neurophysiology., 72*(4), 312-320.

Dijk, D. J., Beersma, D. G., & Daan, S. (1987). EEG power density during nap sleep: reflection of an hourglass measuring the duration of prior wakefulness. *Journal of Biological Rhythms., 2*(3), 207-219.

Dijk, D. J., Brunner, D. P., & Borbely, A. A. (1990). Time course of EEG power density during long sleep in humans. *American Journal of Physiology., 258*(3 Pt 2), R650-661.

Dijk, D. J., Hayes, B., & Czeisler, C. A. (1993). Dynamics of electroencephalographic sleep spindles and slow wave activity in men: effect of sleep deprivation. *Brain Research., 626*(1-2), 190-199.

Dimpfel, W., Schober, F., & Spuler, M. (1993). The influence of caffeine on human EEG under resting conditions and during mental loads. *Clinical Investigator., 71*(3), 197-207.

Dinges, D. F. (1989). Napping patterns and effects in human adults. In D. F. Dinges & R. J. Broughton (Eds.), *Sleep and Alertness: Chronobiological, Behavioral and Medical Aspects of Napping. New York: Raven* (pp. 171-204). New York: Raven Press.

Dinges, D. F. (1995). An overview of sleepiness and accidents. *Journal of Sleep Research, 4*(Supplement 2), 4-14.

Dinges, D. F., & Kribbs, N. B. (1991). Performing while sleepy: effects of experimentally-induced sleepiness. In T. H. Monk (Ed.), *Sleep, sleepiness and performance*. Chichester: John Wiley & sons.

Dinges, D. F., Mallis, M., Maislin, G., & Powell, J. W. (1998). *Evaluation of techniques for ocular measurement as an index of fatigue and the basis for alertness management* (No. Final Report # DOT HS 808 762): US Department of Transportation, National Highway Traffic Safety Administration.

Dinges, D. F., Pack, F., Williams, K., Gillen, K. A., Powell, J. W., Ott, G. E., et al. (1997). Cumulative sleepiness, mood disturbance, and psychomotor vigilance performance decrements during a week of sleep restricted to 4-5 hours per night. *Sleep., 20*(4), 267-267.

Doghramji, K., Mitler, M. M., Sangal, R. B., Shapiro, C., Taylor, S., Walsleben, J., et al. (1997). A normative study of the maintenance of wakefulness test (MWT). *Electroencephalography & Clinical Neurophysiology., 103*(5), 554-562.

Doran, S. M., Van Dongen, H. P., & Dinges, D. F. (2001). Sustained attention performance during sleep deprivation: evidence of state instability. *Archives Italiennes de Biologie, 139*(3), 253-267.

Dorrian, J., Rogers, N. L., & Dinges, D. F. (2005). Psychomotor vigilance performance: neurocognitive assay sensitive to sleep loss. In C. Kushida (Ed.), *Sleep deprivation: clinical issues, pharmacology and sleep loss effects* (Vol. 193, pp. 39-68). New York: Marcel Dekker Inc.

Drummond, S. P., Brown, G. G., Stricker, J. L., Buxton, R. B., Wong, E. C., & Gillin, J. C. (1999). Sleep deprivation-induced reduction in cortical functional response to serial subtraction. *Neuroreport, 10*(18), 3745-3748.

Edgar, D. M., Dement, W. C., & Fuller, C. A. (1993). Effect of SCN lesions on sleep in squirrel monkeys: evidence for opponent processes in sleep-wake regulation. *Journal of Neuroscience, 13*(3), 1065-1079.

Engleman, H. M., & Joffe, D. (1999). Neuropsychological function in obstructive sleep apnoea. *Sleep Medicine Reviews, 3*(1), 59-78.

Feinberg, I. (1974). Changes in sleep cycle patterns with age. *Journal of Psychiatric Research, 10*(3-4), 283-306.

Feinberg, I., & Floyd, T. C. (1979). Systematic trends across the night in human sleep cycles. *Psychophysiology., 16*(3), 283-291.

Feinberg, I., Maloney, T., & Campbell, I. G. (2000). Effects of hypnotics on the sleep EEG of healthy young adults: new data and psychopharmacologic implications. *Journal of Psychiatric Research., 34*(6), 423-438.

Finelli, L. A., Baumann, H., Borbely, A. A., & Achermann, P. (2000). Dual electroencephalogram markers of human sleep homeostasis: correlation between theta activity in waking and slow-wave activity in sleep. *Neuroscience, 101*(3), 523-529.

Folstein, M. F., & Luria, R. (1973). Reliability, validity, and clinical application of the Visual Analogue Mood Scale. *Psychological Medicine, 3*(4), 479-486.

Foulkes, D., & Vogel, G. (1965). Mental activity at sleep onset. *Journal of Abnormal Psychology, 70*, 231-243.

Franken, P., Chollet, D., & Tafti, M. (2001). The homeostatic regulation of sleep need is under genetic control. *Journal of Neuroscience, 21*(8), 2610-2621.

Freedman, R. R. (1986). EEG power spectra in sleep-onset insomnia. *Electroencephalography & Clinical Neurophysiology, 63*(5), 408-413.

Frey, D. J., Badia, P., & Wright, K. P. (2004). Inter- and intra-individual variability in performance near the circadian nadir during sleep deprivation. *J Sleep Res, 13*(4), 305-315.

Fulda, S., & Schulz, H. (2003). Cognitive dysfunction in sleep-related breathing disorders: a meta-analysis. *Sleep Research Online, 5*(1), 19-51.

Gillberg, M., Kecklund, G., & Akerstedt, T. (1994). Relations between performance and subjective ratings of sleepiness during a night awake. *Sleep., 17*(3), 236-241.

Gordon, E. (2003). Integrative neuroscience in psychiatry: the role of a standardized database. *Australas Psychiat, 11*(2), 156-163.

Gordon, E., Cooper, N., Rennie, C., Hermens, D., & Williams, L. M. (2005). Integrative neuroscience: the role of a standardized database. *Clinical EEG & Neuroscience: Official Journal of the EEG & Clinical Neuroscience Society (ENCS), 36*(2), 64-75.

Gottlieb, D. J., Whitney, C. W., Bonekat, W. H., Iber, C., James, G. D., Lebowitz, M., et al. (1999). Relation of sleepiness to respiratory disturbance index: the Sleep Heart Health Study. *American Journal of Respiratory & Critical Care Medicine, 159*(2), 502-507.

Gratton, G., Coles, M. G., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography & Clinical Neurophysiology., 55*(4), 468-484.

Grunstein, R. R., Ho, K. Y., & Sullivan, C. E. (1991). Sleep apnea in acromegaly. *Annals of Internal Medicine., 115*(7), 527-532.

Guilleminault, C., Do Kim, Y., Chowdhuri, S., Horita, M., Ohayon, M., & Kushida, C. (2001). Sleep and daytime sleepiness in upper airway resistance syndrome compared to obstructive sleep apnoea syndrome. *European Respiratory Journal., 17*(5), 838-847.

Guilleminault, C., Stoohs, R., & Duncan, S. (1991). Snoring (I). Daytime sleepiness in regular heavy snorers. *Chest, 99*(1), 40-48.

Gundel, A., & Witthoft, H. (1983). Circadian rhythm in the EEG of man. *International Journal of Neuroscience, 19*(1-4), 287-292.

Gunstad, J., Cohen, R. A., Paul, R. H., Luyster, F. S., & Gordon, E. (2006). Age effects in time estimation: relationship to frontal brain morphometry. *Journal of Integrative Neuroscience, 5*(1), 75-87.

Halstead, W. C. (1947). *Brain and Intelligence: A Quantitative Study of the Frontal Lobes*. Chicago: University of Chicago Press.

Harrison, Y., & Horne, J. A. (2000). The Impact of Sleep Deprivation on Decision Making: A Review. *Journal of Experimental Psychology: Applied, 6*(3), 236-249.

Heinzer, R., Gaudreau, H., Decary, A., Sforza, E., Petit, D., Morisson, F., et al. (2001). Slow-wave activity in sleep apnea patients before and after continuous positive airway pressure treatment: contribution to daytime sleepiness. *Chest, 119*(6), 1807-1813.

Henry, J. D., & Crawford, J. R. (2005). The short-form version of the Depression Anxiety Stress Scales (DASS-21): construct validity and normative data in a large non-clinical sample. *British Journal of Clinical Psychology, 44*(Pt 2), 227-239.

Himanen, S. L., Joutsen, A., & Virkkala, J. (2004). Visual assessment of selected high amplitude frontopolar slow waves of sleep: differences between healthy subjects and apnea patients. *Clinical EEG & Neuroscience, 35*(3), 125-131.

Hoddes, E., Zarcone, V., Smythe, H., Phillips, R., & Dement, W. C. (1973). Quantification of sleepiness: a new approach. *Psychophysiology., 10*(4), 431-436.

Hoptman, M. J., & Davidson, R. J. (1998). Baseline eeg asymmetries and performance on neuropsychological tasks. *Neuropsychologia, 36*(12), 1343-1353.

Horne, J. A. (1991). Dimensions to sleepiness. In T. H. Monk (Ed.), *Sleep, Sleepiness and Performance* (pp. 169-196). Chichester: John Wiley and Sons.

Horne, J. A. (1993). Human sleep, sleep loss and behaviour. Implications for the prefrontal cortex and psychiatric disorder. *British Journal of Psychiatry, 162*, 413-419.

Horne, J. A., & Pettitt, A. N. (1985). High incentive effects on vigilance performance during 72 hours of total sleep deprivation. *Acta Psychologica, 58*(2), 123-139.

Huber, R., Felice Ghilardi, M., Massimini, M., & Tononi, G. (2004). Local sleep and learning. *Nature, advanced online publication*.

Huber, R., Tononi, G., & Cirelli, C. (2007). Exploratory behavior, cortical BDNF expression, and sleep homeostasis. *Sleep, 30*(2), 129-139.

Hyoki, K., Shigeta, M., Tsuno, N., Kawamuro, Y., & Kinoshita, T. (1998). Quantitative electro-oculography and electroencephalography as indices of alertness. *Electroencephalography and Clinical Neurophysiology, 106*(3), 213-219.

Inoue, Y., Nanba, K., Kojima, K., Mitani, H., & Arai, A. H. (2001). P300 abnormalities in patients with severe sleep apnea syndrome. *Psychiatry and Clinical Neurosciences, 55*(3), 247-248.

Jenkinson, C., Davies, R. J., Mullins, R., & Stradling, J. R. (1999). Comparison of therapeutic and subtherapeutic nasal continuous positive airway pressure for obstructive sleep apnoea: a randomised prospective parallel trial.[see comment]. *Lancet, 353*(9170), 2100-2105.

Johns, M. W. (1991). A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep., 14*(6), 540-545.

Johns, M. W. (1992). Reliability and factor analysis of the Epworth Sleepiness Scale. *Sleep, 15*(4), 376-381.

Johns, M. W. (1993). Daytime sleepiness, snoring, and obstructive sleep apnea. The Epworth Sleepiness Scale. *Chest, 103*(1), 30-36.

Johns, M. W. (1994). Sleepiness in different situations measured by the Epworth Sleepiness Scale. *Sleep, 17*(8), 703-710.

Johns, M. W. (1998). Rethinking the assessment of sleepiness. *Sleep Medicine Reviews, 2*(1), 3-15.

Johns, M. W. (2000). Sensitivity and specificity of the multiple sleep latency test (MSLT), the maintenance of wakefulness test and the Epworth sleepiness scale: failure of the MSLT as a gold standard.[see comment]. *Journal of Sleep Research, 9*(1), 5-11.

Johns, M. W. (2002). Sleep propensity varies with behaviour and the situation in which it is measured: the concept of somnificity. *Journal of Sleep Research, 11*(1), 61-67.

Johns, M. W. (2003). The Amplitude-Velocity Ratio of Blinks: A new Method for Monitoring Drowsiness. *Sleep, 26*(Supplement), A51-52.

Johns, M. W., & Hocking, B. (1997). Daytime sleepiness and sleep habits of Australian workers. *Sleep, 20*(10), 844-849.

Jurado, J. L., Luna-Villegas, G., & Buela-Casal, G. (1989). Normal human subjects with slow reaction times and larger time estimations after waking have diminished delta sleep. *Electroencephalography & Clinical Neurophysiology, 73*(2), 124-128.

Kaida, K., Takahashi, M., Akerstedt, T., Nakata, A., Otsuka, Y., Haratani, T., et al. (2006). Validation of the Karolinska sleepiness scale against performance and EEG variables. *Clinical Neurophysiology, 117*(7), 1574-1581.

Kecklund, G., & Akerstedt, T. (1993). Sleepiness in long distance truck driving: an ambulatory EEG study of night driving. *Ergonomics., 36*(9), 1007-1017.

Kemp, A. H., Stephan, B. C., Hopkinson, P., Sumich, A. L., Paul, R. H., Clark, C. R., et al. (2005). Toward an integrated profile of depression: evidence from the brain resource international database. *J Integr Neurosci, 4*(1), 95-106.

Kim, H., Dinges, D., & Young, T. (2007 in press). Sleep-disordered breathing and psychomotor vigilance in a community-based sample. *Sleep*.

Kingshott, R. N., Cosway, R. J., Deary, I. J., & Douglas, N. J. (2000). The effect of sleep fragmentation on cognitive processing using computerized topographic brain mapping. *Journal of Sleep Research, 9*(4), 353-357.

Kirshner, B., & Guyatt, G. (1985). A methodological framework for assessing health indices. *Journal of Chronic Diseases, 38*(1), 27-36.

Kleitman, N. (1963). *Sleep and Wakefulness*: University of Chicago Press.

Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Research - Brain Research Reviews, 29*(2-3), 169-195.

Kotterba, S., Rasche, K., Widdig, W., Duscha, C., Blombach, S., Schultze-Werninghaus, G., et al. (1998). Neuropsychological investigations and event-related potentials in obstructive sleep apnea syndrome before and during CPAP-therapy. *J Neurol Sci, 159*(1), 45-50.

Kraemer, S., Danker-Hopfe, H., Dorn, H., Schmidt, A., Ehlert, I., & Herrmann, W. M. (2000). Time-of-day variations of indicators of attention: performance, physiologic parameters, and self-assessment of sleepiness. *Biological Psychiatry, 48*(11), 1069-1080.

Krieger, A. C., Ayappa, I., Norman, R. G., Rapoport, D. M., & Walsleben, J. (2004). Comparison of the maintenance of wakefulness test (MWT) to a modified behavioral test (OSLER) in the evaluation of daytime sleepiness. *J Sleep Res, 13*(4), 407-411.

Krikorian, R., Bartok, J., & Gay, N. (1994). Tower of London procedure: a standard method and developmental data. *Journal of Clinical & Experimental Neuropsychology, 16*(6), 840-850.

Lafrance, C., Paquet, J., & Dumont, M. (2002). Diurnal time courses in psychomotor performance and waking EEG frequencies. *Brain & Cognition., 48*(2-3), 625-631.

Lal, S. K., & Craig, A. (2002). Driver fatigue: electroencephalography and psychological assessment. *Psychophysiology., 39*(3), 313-321.

Lavie, P. (1979). Ultradian rhythms in alertness - a pupillometric study. *Biol Psychol, 9*(1), 49-62.

Leproult, R., Colecchia, E. F., Berardi, A. M., Stickgold, R., Kosslyn, S. M., & Van Cauter, E. (2003). Individual differences in subjective and objective alertness during sleep deprivation are stable and unrelated. *American Journal of Physiology - Regulatory, Integrative and Comparative Physiology, 284*(2), R280-290.

Lockley, S. W., Cronin, J. W., Evans, E. E., Cade, B. E., Lee, C. J., Landrigan, C. P., et al. (2004). Effect of reducing interns' weekly work hours on sleep and attentional failures. *New England Journal of Medicine, 351*(18), 1829-1837.

Loomis, A. L., Harvey, E. N., & Hobart, G. (1936). Electrical potentials of the human brain. *Journal of Experimental Psychology: General, 19*(3), 249-279.

Loomis, A. L., Harvey, E. N., & Hobart, G. A. (1937). Cerebral states during sleep, as studied by human brain potentials. *Journal of Experimental Psychology: General, 21*(2), 127-144:.

Lovibond, P. (2007, accessed 21 June). *Overview of the DASS and its uses*, from http://www2.psy.unsw.edu.au/groups/dass/

Lovibond, S. H., & Lovibond, P. F. (1995). *Manual for the Depression Anxiety Stress Scales*. Sydney: The Psychology Foundation of Australia.

Lubin, A., Hord, D. J., Tracy, M. L., & Johnson, L. C. (1976). Effects of exercise, bedrest and napping on performance decrement during 40 hours. *Psychophysiology, 13*(4), 334-339.

Macchi, M. M., Boulos, Z., Ranney, T., Simmons, L., & Campbell, S. S. (2002). Effects of an afternoon nap on nighttime alertness and performance in long-haul drivers. *Accident Analysis & Prevention, 34*(6), 825-834.

MacLean, A. W., Criollo, M., Fekken, G. C., & Knowles, J. B. (1990). The Stanford Sleepiness Scale in a clinic sample: the need for revision. *Sleep Research, 19*, 249.

MacLean, A. W., Saskin, P., Fekken, G. C., & Knowles, J. B. (1989). Should the Stanford sleepiness scale be revised? *Sleep Research, 18*, 370.

Magill, B. (2001). *Cronbach alpha function*, from https://stat.ethz.ch/pipermail/r-help/2001-March/011982.html

Maislin, G., Pack, A. I., Kribbs, N. B., Smith, P. L., Schwartz, A. R., Kline, L. R., et al. (1995). A survey screen for prediction of apnea. *Sleep., 18*(3), 158-166.

Marshall, N. S., Barnes, M., Travier, N., Campbell, A. J., Pierce, R. J., McEvoy, R. D., et al. (2006). Continuous positive airway pressure reduces daytime sleepiness in mild to moderate obstructive sleep apnoea: a meta-analysis. *Thorax, 61*(5), 430-434.

McCarthy, G., & Donchin, E. (1981). A metric for thought: a comparison of P300 latency and reaction time. *Science, 211*(4477), 77-80.

McCormick, D. A., & Bal, T. (1997). Sleep and Arousal: Thalamocortical Mechanisms. *Annual Review of Neuroscience, 20*(1), 185-215.

McLaren, J. W., Hauri, P. J., Lin, S.-C., & Harris, C. D. (2002). Pupillometry in clinically sleepy patients. *Sleep Medicine, 3*(4), 347-352.

McNair, D. M., Lorr, M., & Droppleman, L. F. (1971). *EITS manual for the profile of mood states*. San Diego: Educational and Industrial Test Services.

Merica, H., & Fortune, R. D. (1997). A Neuronal Transition Probability Model for the Evolution of Power in the Sigma and Delta Frequency Bands of Sleep EEG. *Physiology & Behavior, 62*(3), 585-589.

Miletin, M. S., & Hanly, P. J. (2003). Measurement properties of the Epworth sleepiness scale. *Sleep Medicine, 4*(3), 195-199.

Miller, J. D., Morin, L. P., Schwartz, W. J., & Moore, R. Y. (1996). New insights into the mammalian circadian clock. *Sleep, 19*(8), 641-667.

Mitler, M. M., Doghramji, K., & Shapiro, C. (2000). The maintenance of wakefulness test: normative data by age. *Journal of Psychosomatic Research., 49*(5), 363-365.

Mitler, M. M., Gujavarty, K. S., & Browman, C. P. (1982). Maintenance of wakefulness test: a polysomnographic technique for evaluation treatment efficacy in patients with excessive somnolence. *Electroencephalography & Clinical Neurophysiology., 53*(6), 658-661.

Monk, T. H. (1989). A Visual Analogue Scale technique to measure global vigor and affect. *Psychiatry Research., 27*(1), 89-99.

Moore, R. Y. (1983). Organization and function of a central nervous system circadian oscillator: the suprachiasmatic hypothalamic nucleus. *Federation Proceedings, 42*(11), 2783-2789.

Morisson, F., Decary, A., Petit, D., Lavigne, G., Malo, J., & Montplaisir, J. (2001). Daytime sleepiness and EEG spectral analysis in apneic patients before and after treatment with continuous positive airway pressure. *Chest, 119*(1), 45-52.

Morisson, F., Lavigne, G., Petit, D., Nielsen, T., Malo, J., & Montplaisir, J. (1998). Spectral analysis of wakefulness and REM sleep EEG in patients with sleep apnoea syndrome. *European Respiratory Journal, 11*(5), 1135-1140.

Morris, G. O., Williams, H. L., & Lubin, A. (1960). Misperception and disorientation during sleep deprivation. *Archives of General Psychiatry, 2*, 247-254.

Morris, R. G., Ahmed, S., Syed, G. M., & Toone, B. K. (1993). Neural correlates of planning ability: frontal lobe activation during the Tower of London test. *Neuropsychologia, 31*(12), 1367-1378.

Munch, M., Knoblauch, V., Blatter, K., Schroder, C., Schnitzler, C., Krauchi, K., et al. (2004). The frontal predominance in human EEG delta activity after sleep loss decreases with age. *European Journal of Neuroscience, 20*(5), 1402-1410.

Murray, B. J. (2007). Brain Death by a Thousand Hypoxic Cuts in Sleep. *Am. J. Respir. Crit. Care Med., 175*(6), 528-529.

National Transportation Safety Board. (1990). *Marine accident report: Grounding of the U. S. Tankship Exxon Valdez on Bligh Reef, Prince William Sound, near Valdez, Alaska, March 24, 1989*. Washington, D. C.: National Transportation Safety Board.

Newton, T. F., Kalechstein, A. D., Hardy, D. J., Cook, I. A., Nestor, L., Ling, W., et al. (2004). Association between quantitative EEG and neurocognition in methamphetamine-dependent volunteers. *Clinical Neurophysiology, 115*(1), 194-198.

Niggemyer, K. A., Begley, A., Monk, T., & Buysse, D. J. (2004). Circadian and homeostatic modulation of sleep in older adults during a 90-minute day study. *Sleep, 27*(8), 1535-1541.

O'Connor, S. C., & Robinson, P. A. (2004). Spatially uniform and nonuniform analyses of electroencephalographic dynamics, with application to the topography of the alpha rhythm. *Physical Review E (Statistical, Nonlinear, and Soft Matter Physics), 70*(1), 011911-011919.

Oken, B. S., & Salinsky, M. C. (2007). Sleeping and driving: Not a safe dual-task. *Clinical Neurophysiology, 118*(9), 1899-1900.

Olson, L., Cole, M., & Ambrogetti, A. (1998). Correlations among Epworth Sleepiness Scale scores, multiple sleep latency tests and psychological symptoms. *Journal of Sleep Research, 7*(4), 248-253.

Ondze, B., Espa, F., Dauvilliers, Y., Billiard, M., & Besset, A. (2003). Sleep architecture, slow wave activity and sleep spindles in mild sleep disordered breathing. *Clinical Neurophysiology, 114*(5), 867-874.

Pallant, J. (2004). *SPSS Survival Manual: A step by step guide to data analysis using SPSS*: Allen & Unwin.

Papadelis, C., Chen, Z., Kourtidou-Papadeli, C., Bamidis, P. D., Chouvarda, I., Bekiaris, E., et al. (2007). Monitoring sleepiness with on-board electrophysiological recordings for preventing sleep-deprived traffic accidents. *Clinical Neurophysiology, 118*(9), 1906-1922.

Patel, S. R., White, D. P., Malhotra, A., Stanchina, M. L., & Ayas, N. T. (2003). Continuous positive airway pressure therapy for treating sleepiness in a diverse population with obstructive sleep apnea: results of a meta-analysis. *Archives of Internal Medicine, 163*(5), 565-571.

Paul, R. H., Lawrence, J., Williams, L. M., Richard, C. C., Cooper, N., & Gordon, E. (2005). Preliminary Validity Of "IntegneuroTM": A New Computerized Battery Of Neurocognitive Tests. *International Journal of Neuroscience, 115*(11), 1549-1567.

Paus, T., Zatorre, R. J., Hofle, N., Caramanos, Z., Gotman, J., Petrides, M., et al. (1997). Time-related changes in neural systems underlying attention and arousal in humans. *J. Cogn. Neurosci, 9*, 392-408.

Pavlova, M., Berg, O., Gleason, R., Walker, F., Roberts, S., & Regestein, Q. (2001). Self-reported hyperarousal traits among insomnia patients. *Journal of Psychosomatic Research, 51*(2), 435-441.

Philip, P., Stoohs, R., & Guilleminault, C. (1994). Sleep fragmentation in normals: a model for sleepiness associated with upper airway resistance syndrome. *Sleep., 17*(3), 242-247.

Pivik, R. T. (1991). The several qualities of sleepiness: psychophysiological considerations. In T. H. Monk (Ed.), *Sleep, Sleepiness and Performance* (pp. 3-37). Chichester, England: John Wiley & Sons.

Polotsky, V. Y., Rubin, A. E., Balbir, A., Dean, T., Smith, P. L., Schwartz, A. R., et al. (2006). Intermittent hypoxia causes REM sleep deficits and decreases EEG delta power in NREM sleep in the C57BL/6J mouse. *Sleep Medicine, 7*(1), 7-16.

Porkka-Heiskanen, T., Alanko, L., Kalinchuk, A., & Stenberg, D. (2002). Adenosine and sleep. *Sleep Med Rev, 6*(4), 321-332.

Porkka-Heiskanen, T., Strecker, R. E., Thakkar, M., Bjorkum, A. A., Greene, R. W., & McCarley, R. W. (1997). Adenosine: a mediator of the sleep-inducing effects of prolonged wakefulness. *Science, 276*(5316), 1265-1268.

Press, W. H. (1992). *Numerical recipes in C: the art of scientific computing*. Cambridge: Cambridge University Press.

Priest, B., Brichard, C., Aubert, G., Liistro, G., & Rodenstein, D. O. (2001). Microsleep during a simplified maintenance of wakefulness test. A validation study of the OSLER test.[comment]. *American Journal of Respiratory & Critical Care Medicine., 163*(7), 1619-1625.

R Development Core Team. (2007). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.

Rechtschaffen, A., & Kales, A. (Eds.). (1968). *A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects.* Los Angeles: University of California.

Redline, S., Kirchner, H. L., Quan, S. F., Gottlieb, D. J., Kapur, V., & Newman, A. (2004). The effects of age, sex, ethnicity, and sleep-disordered breathing on sleep architecture. *Archives of Internal Medicine, 164*(4), 406-418.

Regestein, Q. R., Dambrosia, J., Hallett, M., Murawski, B., & Paine, M. (1993). Daytime alertness in patients with primary insomnia. *American Journal of Psychiatry, 150*(10), 1529-1534.

Reitan, R. M. (1955). The relation of the trail making test to organic brain damage. *Journal of Consulting Psychology, 19*(5), 393-394.

Rennie, C. (2000). edfcor program.

Rennie, C. (2004a). edfview program (version 20 July 2004).

Rennie, C. (2004b). eegfit program (version 31 August 2004). In.

Rennie, C. J., Robinson, P. A., & Wright, J. J. (2002). Unified neurophysical model of EEG spectra and evoked potentials. *Biological Cybernetics., 86*(6), 457-471.

Robinson, P. A., Rennie, C. J., & Rowe, D. L. (2002). Dynamics of large-scale brain activity in normal arousal states and epileptic seizures. *Physical Review E. Statistical, Nonlinear, & Soft Matter Physics., 65*(4 Pt 1), 041924.

Robinson, P. A., Rennie, C. J., Rowe, D. L., & O'Connor, S. C. (2004). Estimation of multiscale neurophysiologic parameters by electroencephalographic means. *Human Brain Mapping, 23*(1), 53-72.

Robinson, P. A., Rennie, C. J., Rowe, D. L., O'Connor, S. C., & Gordon, E. (2005). Multiscale brain modelling. *Philosophical Transactions of the Royal Society of London - Series B: Biological Sciences, 360*(1457), 1043-1050.

Robinson, P. A., Rennie, C. J., & Wright, J. J. (1997). Propagation and stability of waves of electrical activity in the cerebral cortex. *Physical Review E, 56*(1), 826-840.

Robinson, P. A., Whitehouse, R. W., & Rennie, C. J. (2003). Nonuniform corticothalamic continuum model of electroencephalographic spectra with application to split-alpha peaks. *Physical Review E Statistical, Nonlinear, & Soft Matter Physics, 68*(2 Pt 1), 021922.

Rogers, N. L., Dorrian, J., & Dinges, D. F. (2003). Sleep, waking and neurobehavioural performance. *Frontiers in Bioscience, 8*, s1056-1067.

Rosenthal, L., Roehrs, T. A., & Roth, T. (1993). The Sleep-Wake Activity Inventory: a self-report measure of daytime sleepiness. *Biological Psychiatry, 34*(11), 810-820.

Roth, D. L., & Crosson, B. (1985). Memory span and long-term memory deficits in brain-impaired patients. *Journal of Clinical Psychology, 41*(4), 521-527.

Rowe, D. L., Robinson, P. A., & Gordon, E. (2005). Stimulant drug action in attention deficit hyperactivity disorder (ADHD): inference of neurophysiological mechanisms via quantitative modelling. *Clinical Neurophysiology, 116*(2), 324-335.

Rowe, D. L., Robinson, P. A., Lazzaro, I. L., Powles, R. C., Gordon, E., & Williams, L. M. (2005). Biophysical modeling of tonic cortical electrical activity in attention deficit hyperactivity disorder. *International Journal of Neuroscience, 115*(9), 1273-1305.

Rowe, D. L., Robinson, P. A., & Rennie, C. J. (2004). Estimation of neurophysiological parameters from the waking EEG using a biophysical model of brain dynamics. *Journal of Theoretical Biology, 231*(3), 413-433.

Rowe, D. L., Robinson, P. A., Rennie, C. J., Harris, A. W., Felmingham, K. L., Lazzaro, I. L., et al. (2004). Neurophysiologically-based mean-field modelling of tonic cortical activity in post-traumatic stress disorder (PTSD), schizophrenia, first episode schizophrenia and attention deficit hyperactivity disorder (ADHD). *Journal of Integrative Neuroscience, 3*(4), 453-487.

Rumbach, L., Krieger, J., & Kurtz, D. (1991). Auditory event-related potentials in obstructive sleep apnea: effects of treatment with nasal continuous positive airway pressure. *Electroencephalogr Clin Neurophysiol, 80*(5), 454-457.

Sagaspe, P., Taillard, J., Chaumet, G., Guilleminault, C., Coste, O., Moore, N., et al. (2007). Maintenance of Wakefulness Test as a Predictor of Driving Performance in Patients With Untreated Obstructive Sleep Apnea. *Sleep, 30*(3), 327-330.

Sakamoto, Y., Ishiguro, M., & Kitagawa, G. (1986). *Akaike Information Criterion Statistics*. Dordrecht, Holland: D. Reidel Publishing Company.

Sallinen, M., Harma, M., Akila, R., Holm, A., Luukkonen, R., Mikola, H., et al. (2004). The effects of sleep debt and monotonous work on sleepiness and performance during a 12-h dayshift. *Journal of Sleep Research, 13*(4), 285-294.

Sangal, R. B., Mitler, M. M., & Sangal, J. M. (1999). Subjective sleepiness ratings (Epworth sleepiness scale) do not reflect the same parameter of sleepiness as objective sleepiness (maintenance of wakefulness test) in patients with narcolepsy. *Clinical Neurophysiology, 110*(12), 2131-2135.

Sangal, R. B., & Sangal, J. M. (1997a). Abnormal visual P300 latency in obstructive sleep apnea does not change acutely upon treatment with CPAP. *Sleep, 20*(9), 702-704.

Sangal, R. B., & Sangal, J. M. (1997b). Obstructive sleep apnea and abnormal P300 latency topography. *Clinical Electroencephalography, 28*(1), 16-25.

Sangal, R. B., Thomas, L., & Mitler, M. M. (1992). Maintenance of wakefulness test and multiple sleep latency test. Measurement of different abilities in patients with sleep disorders. *Chest., 101*(4), 898-902.

Santamaria, J., & Chiappa, K. H. (1987). The EEG of drowsiness in normal adults. *Journal of Clinical Neurophysiology., 4*(4), 327-382.

Saper, C. B., Chou, T. C., & Scammell, T. E. (2001). The sleep switch: hypothalamic control of sleep and wakefulness. *Trends in Neurosciences, 24*(12 SU -), 726-731.

Schmidt, M. (1996). *Rey Auditory Verbal Learning Test: RAVLT: a Handbook*. Los Angeles: Western Psychological Services.

Seugnet, L., Boero, J., Gottschalk, L., Duntley, S. P., & Shaw, P. J. (2006). Identification of a biomarker for sleep drive in flies and humans. *PNAS*, 0609463104.

Sforza, E., Grandin, S., Jouny, C., Rochat, T., & Ibanez, V. (2002). Is waking electroencephalographic activity a predictor of daytime sleepiness in sleep-related breathing disorders? *European Respiratory Journal., 19*(4), 645-652.

Sforza, E., & Krieger, J. (1992). Daytime sleepiness after long-term continuous positive airway pressure (CPAP) treatment in obstructive sleep apnea syndrome. *Journal of the Neurological Sciences, 110*(1-2), 21-26.

Shallice, T. (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society of London - Series B: Biological Sciences, 298*(1089), 199-209.

Shen, J., Barbera, J., & Shapiro, C. M. (2006). Distinguishing sleepiness and fatigue: focus on definition and measurement. *Sleep Medicine Reviews, 10*(1), 63-76.

Shwedyk, E., Balasubramanian, R., & Scott, R. N. (1977). A nonstationary model for the electromyogram. *IEEE Transactions on Biomedical Engineering, 24*(5), 417-424.

Siegel, J. M. (2005). Clues to the functions of mammalian sleep. *Nature, 437*(7063), 1264-1271.

Signal, T. L., Gale, J., & Gander, P. H. (2005). Sleep measurement in flight crew: comparing actigraphic and subjective estimates to polysomnography. *Aviation Space & Environmental Medicine, 76*(11), 1058-1063.

Silber, M. H., Ancoli-Israel, S., Bonnet, M. H., Chokroverty, S., Grigg-Damberger, M. M., Hirshkowitz, M., et al. (2007). The Visual Scoring of Sleep in Adults. *Journal of Clinical Sleep Medicine, 3*(2), 121-131.

Simpson, J. A., & Weiner, E. S. C. (1989). *The Oxford English Dictionary* (2nd ed.). Oxford: Clarendon Press.

Sleep Disorders Atlas Task Force of the American Sleep Disorders Association. (1992). EEG arousals: scoring rules and examples: a preliminary report from the Sleep Disorders Atlas Task Force of the American Sleep Disorders Association. *Sleep, 15*(2), 173-184.

Sleigh, J. W., Andrzejowski, J., Steyn-Ross, A., & Steyn-Ross, M. (1999). The bispectral index: a measure of depth of sleep? *Anesthesia & Analgesia, 88*(3), 659-661.

Smit, A. S., Eling, P. A., & Coenen, A. M. (2004). Mental effort affects vigilance enduringly: after-effects in EEG and behavior. *International Journal of Psychophysiology, 53*(3), 239-243.

Smit, A. S., Eling, P. A. T. M., & Coenen, A. M. L. (2004). Mental effort causes vigilance decrease due to resource depletion. *Acta Psychologica, 115*(1), 35-42.

Smulders, F. T. Y., Kenemans, J. L., Jonkman, L. M., & Kok, A. (1997). The effects of sleep loss on task performance and the electroencephalogram in young and elderly subjects. *Biological Psychology, 45*(1-3), 217-239.

Spielberger, C. D., Gorusch, R., & Lushene, R. (1970). *The State Trait Anxiety Inventory (STAI) Manual*. Palo Alto: Consulting Psychologists Press.

Stampi, C., Stone, P., & Michimori, A. (1995). A new quantitative method for assessing sleepiness: The alpha attenuation test. *Work and Stress, 9*(2-3), 368-376.

Stenberg, D., Litonius, E., Halldner, L., Johansson, B., Fredholm, B. B., & Porkka-Heiskanen, T. (2003). Sleep and its homeostatic regulation in mice lacking the adenosine A1 receptor. *Journal of Sleep Research, 12*(4), 283-290.

Steriade, M. (2000). Corticothalamic resonance, states of vigilance and mentation. *Neuroscience, 101*(2), 243-276.

Steriade, M. (2003). The corticothalamic system in sleep. *Frontiers in Bioscience, 8*(99), d878-899.

Steriade, M., & Amzica, F. (1998). Coalescence of sleep rhythms and their chronology in corticothalamic networks. *Sleep Research Online, 1*(1), 1-10.

Steriade, M., Nunez, A., & Amzica, F. (1993). A novel slow (< 1 Hz) oscillation of neocortical neurons in vivo: depolarizing and hyperpolarizing components. *Journal of Neuroscience, 13*(8), 3252-3265.

Streiner, D. L., & Norman, G. R. (1989). *Health Measurement Scales. A Practical Guide to their Development and Use*. Oxford: Oxford University Press.

Strijkstra, A. M., Beersma, D. G. M., Drayer, B., Halbesma, N., & Daan, S. (2003). Subjective sleepiness correlates negatively with global alpha (8-12 Hz) and positively with central frontal theta (4-8 Hz) frequencies in the human resting awake electroencephalogram. *Neuroscience Letters, 340*(1), 17-20.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *J. Exp. Psychol, 18*(6), 643-661.

Sugerman, J. L., & Walsh, J. K. (1989). Physiological sleep tendency and ability to maintain alertness at night. *Sleep, 12*(2), 106-112.

Szymusiak, R., Gvilia, I., & McGinty, D. (2007). Hypothalamic control of sleep. *Sleep Med, 8*(4), 291-301.

Taillard, J., Moore, N., Claustrat, B., Coste, O., Bioulac, B., & Philip, P. (2006). Nocturnal sustained attention during sleep deprivation can be predicted by specific periods of subjective daytime alertness in normal young humans. *Journal of Sleep Research, 15*(1), 41-45.

Tanaka, H., Hayashi, M., & Hori, T. (1996). Statistical features of hypnagogic EEG measured by a new scoring system. *Sleep, 19*(9), 731-738.

Thayer, R. E. (1967). Measurement of activation through self-report. *Psychological Reports, 20*(2), 663-678.

Thompson, C., & Harding, G. F. (1968). Circadian rhythm in the EEG and other physiological and psychological variables. *Electroencephalography & Clinical Neurophysiology, 25*(5), 509.

Thorpy, M. J. (1992). The clinical use of the Multiple Sleep Latency Test. The Standards of Practice Committee of the American Sleep Disorders Association.[erratum appears in Sleep 1992 Aug;15(4):381]. *Sleep, 15*(3), 268-276.

Tononi, G., & Cirelli, C. (2003). Sleep and synaptic homeostasis: a hypothesis. *Brain Research Bulletin, 62*(2), 143-150.

Tononi, G., & Cirelli, C. (2006). Sleep function and synaptic homeostasis. *Sleep Medicine Reviews, 10*(1), 49-62.

Torsvall, L., & Akerstedt, T. (1985). Eye closure, sleepiness and EEG spectra. In E. Koella, H. Ruther & H. Schulz (Eds.), *Sleep: Proceedings of the seventh European congress on sleep research, Munich, September 3-7, 1984.* (pp. 300-301). Stuttgart: Gustav Fischer Verlag.

Torsvall, L., & Akerstedt, T. (1987). Sleepiness on the job: continuously measured EEG changes in train drivers. *Electroencephalography & Clinical Neurophysiology., 66*(6), 502-511.

Torsvall, L., & Akerstedt, T. (1988). Extreme sleepiness: quantification of EOG and spectral EEG parameters. *International Journal of Neuroscience., 38*(3-4), 435-441.

Tukey, J. (1962). The Future of Data Analysis. *The Annals of Mathematical Statistics, 33*(1), 1-67.

Uchida, S., Feinberg, I., March, J. D., Atsumi, Y., & Maloney, T. (1999). A comparison of period amplitude analysis and FFT power spectral analysis of all-night human sleep EEG. *Physiology & Behavior., 67*(1), 121-131.

Valley, V., & Broughton, R. (1983). The physiological (EEG) nature of drowsiness and its relation to performance deficits in narcoleptics. *Electroencephalography & Clinical Neurophysiology, 55*(3), 243-251.

van Albada, S., Rennie, C., & Robinson, P. (2005). Traditional and model-based approaches to assessing the reproducibility of the EEG. *Presented in: 15th Australasian Society for Psychophysiology, Wollongong, December*.

van Boxtel, A. (2001). Optimal signal bandwidth for the recording of surface EMG activity of facial, jaw, oral, and neck muscles. *Psychophysiology, 38*(1), 22-34.

van den Berg, J., Neely, G., Nilsson, L., Knutsson, A., & Landstrom, U. (2005). Electroencephalography and subjective ratings of sleep deprivation. *Sleep Medicine, 6*(3), 231-240.

van der Hiele, K., Vein, A. A., Reijntjes, R. H. A. M., Westendorp, R. G. J., Bollen, E. L. E. M., van Buchem, M. A., et al. (2007). EEG correlates in the spectrum of cognitive decline. *Clinical Neurophysiology, 118*(9), 1931-1939.

Van Dongen, H. P., Baynard, M. D., Maislin, G., & Dinges, D. F. (2004). Systematic interindividual differences in neurobehavioral impairment from sleep loss: evidence of trait-like differential vulnerability. *Sleep, 27*(3), 423-433.

Van Dongen, H. P. A., & Dinges, D. F. (2000). Circadian rhythms in fatigue, alertness, and performance. In M. H. Kryger, T. Roth & W. C. Dement (Eds.), *Principles and practice of sleep medicine* (3rd ed., pp. 391-399). Philadelphia: W. B. Saunders.

Van Horn, J. D., Gold, J. M., Esposito, G., Ostrem, J. L., Mattay, V., Weinberger, D. R., et al. (1998). Changing patterns of brain activation during maze learning. *Brain Research, 793*(1-2), 29-38.

Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S* (4th ed.). New York: Springer.

Verstraeten, E., & Cluydts, R. (2004). Executive control of attention in sleep apnea patients: theoretical concepts and methodological considerations. *Sleep Medicine Reviews, 8*(4), 257-267.

Verstraeten, E., Cluydts, R., Pevernagie, D., & Hoffmann, G. (2004). Executive function in sleep apnea: controlling for attentional capacity in assessing executive attention. *Sleep, 27*(4), 685-693.

Vgontzas, A. N., Bixler, E. O., Tan, T. L., Kantner, D., Martin, L. F., & Kales, A. (1998). Obesity without sleep apnea is associated with daytime sleepiness. *Arch Intern Med, 158*(12), 1333-1337.

Viola, A. U., Archer, S. N., James, L. M., Groeger, J. A., Lo, J. C., Skene, D. J., et al. (2007). PER3 polymorphism predicts sleep structure and waking performance. *Current Biology, 17*(7), 613-618.

Violani, C., Lucidi, F., Robusto, E., Devoto, A., Zucconi, M., & Ferini Strambi, L. (2003). The assessment of daytime sleep propensity: a comparison between the Epworth Sleepiness Scale and a newly developed Resistance to Sleepiness Scale. *Clinical Neurophysiology, 114*(6), 1027-1033.

Vitaterna, M. H., Pinto, L. H., & Turek, F. W. (2005). Molecular Genetic Basis for Mammalian Circadian Rhythms. In W. C. Dement, T. Roth & M. H. Kryger (Eds.), *Principles and Practice of Sleep Medicine* (4th ed., pp. 363-374). Philadelphia: Elsevier/Saunders.

Walsh, K. W. (1991). *Understanding Brain Damage: A Primer of Neuropsychological Evaluation*. Edinburgh: Churchill Livingstone.

Walsleben, J. A., Squires, N. K., & Rothenberger, V. L. (1989). Auditory event-related potentials and brain dysfunction in sleep apnea. *Electroencephalogr Clin Neurophysiol, 74*(4), 297-311.

Wang, G., Chen, M., Bian, J., & He, B. (2002). Electroencephalogram spectral power analysis of obstructive sleep apnea syndrome patients before and during continuous positive airway pressure therapy. *Chung-Hua Chieh Ho Ho Hu Hsi Tsa Chih Chinese Journal of Tuberculosis & Respiratory Diseases, 25*(4), 199-203.

Ware, J. E., Jr., & Sherbourne, C. D. (1992). The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Medical Care, 30*(6), 473-483.

Weaver, T. E. (2001). Outcome measurement in sleep medicine practice and research. Part 1: assessment of symptoms, subjective and objective daytime sleepiness, health-related quality of life and functional status. *Sleep Medicine Reviews, 5*(2), 103-128.

Weaver, T. E. (2001). Outcome measurement in sleep medicine practice and research. Part 2: assessment of neurobehavioral performance and mood. *Sleep Medicine Reviews, 5*(3), 223-236.

Weaver, T. E., Laizner, A. M., Evans, L. K., Maislin, G., Chugh, D. K., Lyon, K., et al. (1997). An instrument to measure functional status outcomes for disorders of excessive sleepiness. *Sleep, 20*(10), 835-843.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale - Revised*. New York: Psychological Corporation.

Wichniak, A., Geisler, P., Brunner, H., Tracik, F., Cronlein, T., Friess, E., et al. (2003). Spectral composition of NREM sleep in healthy subjects with moderately increased daytime sleepiness. *Clinical Neurophysiology, 114*(8), 1549-1555.

Wierwille, W. (1999). Historical perspective on slow eyelid closure: Whence PERCLOS. *Technical Proceedings of Ocular Measures of Driver Alertness Conference, Herndon, VA.(FHWA Technical Report No. MC-99-136). Washington, DC: Federal Highway Administration, Office of Motor Carrier and Highway Safety*, 31-53.

Wilkinson, R., & Houghton, D. (1975). Portable four-choice reaction time test with magnetic tape memory. *Behav Res Meth Instrumentation, 7*, 441-446.

Wilkinson, R. T. (1961). Interaction of lack of sleep with knowledge of results, repeated testing, and individual differences. *Journal of Experimental Psychology, 62*, 263-271.

Williams, H. L., Gieseking, C. F., & Lubin, A. (1966). Some effects of sleep loss on memory. *Perceptual & Motor Skills, 23*(3), 1287-1293.

Williams, H. L., Lubin, A., & Goodnow, J. J. (1959). Impaired performance with acute sleep loss. *Psychological Monographs, 73*(14), 1-26.

Williams, L. M., Simms, E., Clark, C. R., Paul, R. H., Rowe, D. L., & Gordon, E. (2005). The Test-Retest Reliability Of A Standardized Neurocognitive And Neurophysiological Test Battery: "Neuromarker". *International Journal of Neuroscience, 115*(12), 1605-1630.

Williamson, A. M., & Feyer, A. M. (2000). Moderate sleep deprivation produces impairments in cognitive and motor performance equivalent to legally prescribed levels of alcohol intoxication. *Occupational & Environmental Medicine, 57*(10), 649-655.

Wong, K. K., Grunstein, R. R., Bartlett, D. J., & Gordon, E. (2006). Brain function in obstructive sleep apnea: results from the brain resource international database. *Journal of Integrative Neuroscience, 5*(1), 111-121.

Yallop, C. (2005). *Macquarie Dictionary* (4th ed.). North Ryde, NSW, Australia: Macquarie Library.

Young, T., Palta, M., Dempsey, J., Skatrud, J., Weber, S., & Badr, S. (1993). The occurrence of sleep-disordered breathing among middle-aged adults. *New England Journal of Medicine, 328*(17), 1230-1235.

Zee, P. C., & Manthena, P. (2007). The brain's master circadian clock: Implications and opportunities for therapy of sleep disorders. *Sleep Medicine Reviews, 11*(1), 59-70.

Zwyghuizen-Doorenbos, A., Roehrs, T., Schaefer, M., & Roth, T. (1988). Test-retest reliability of the MSLT. *Sleep, 11*(6), 562-565.