Neurocognitive Disturbance in Obstructive Sleep Apnoea: Mechanisms of Harm

Michelle Olaithe

University of Western Australia

This thesis is submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy (Ph.D.) at the University of Western Australia

School of Psychology

The University of Western Australia

April 2014

Word count: 29,447

Declaration

“I declare that this report is an original piece of research, conducted by myself and does not contain data or materials which have been previously submitted by myself or anyone for academic credit. I further declare that this report does not contain any materials previously presented by myself or another person, except where due reference is made in the text. This study was conducted with the approval of the Human Research Ethics Committee of the University of Western Australia.”

Name: Michelle Olaithe

Date: 

Signature:
Abstract

The overarching aim addressed by this thesis was the investigation of the relationship between cognitive dysfunction and mechanisms of harm (sleep fragmentation and hypoxia) in Obstructive Sleep Apnoea (OSA). It begins with a general introduction to the nocturnal features and cognitive profile of OSA (Chapter 1), continues with three research studies (reported over Chapters 2, 4 and 5, Chapter 3 details methods and recruitment), and concludes with a general discussion (Chapter 6).

Study 1 (Chapter 2) considered the profile of executive dysfunction in individuals with OSA. OSA is a frequent and often under-diagnosed condition that is associated with upper airway collapse, oxygen desaturation, and sleep fragmentation leading to cognitive dysfunction. There is good meta-analytic evidence that sub-domains of attention and memory are affected by OSA. However, a thorough investigation of the impact of OSA on different sub-domains of executive function had yet to be conducted. Study 1 (Chapter 2) investigated the impact of untreated and treated OSA, in adult patients, on five, theorised, sub-domains of executive function. An extensive literature search was conducted of published and unpublished materials, returning 35 studies that matched selection criteria. Meta-analysis was used to synthesise the results from studies examining the impact of OSA on executive functioning compared to controls (21 studies) and before and after treatment (19 studies); 5 studies met inclusion in both categories. All domains of executive function (Shifting, Updating, Inhibition, Generativity and Fluid Reasoning) demonstrated medium to very large impairments in OSA independent of age, and disease severity. All domains improved (small to medium effects) with CPAP treatment, and this
improvement was not moderated by age or disease severity. Further studies are needed to explore the extent of primary (neural damage in a region, with corresponding behavioural dysfunction, e.g., damage to areas responsible for memory, and demonstrated memory difficulties) or secondary nature (neural damage resulting in secondary dysfunction, e.g., damage in areas responsible for attention control, impacting on memory capacity) of these deficits, and the impact of age and pre-morbid ability (cognitive reserve).

Chapter 3 reports the recruitment procedures, participants and methods used in this thesis.

Study 2 (Chapter 4) assessed the measurement of sleepiness using the Epworth Sleepiness Scale (ESS) prior to inclusion of this construct as a covariate in Study 3. The ESS is a widely used tool for measuring sleepiness. In addition to providing a single measure of sleepiness (a one factor structure), the ESS also has the capacity to provide additional information about specific factors that facilitate sleep-onset, including a person’s posture, activity and environment. These features of sleepiness are referred to as somnificity. Study 2 (Chapter 3) evaluated the fit of a 1-factor structure (sleepiness) and a 3-factor structure (reflecting low, medium and high levels of somnificity) for the ESS using Confirmatory Factor Analysis (CFA). Two samples (a community sample \( N = 356 \) and a clinical sample \( N = 679 \)) were administered the ESS. In both samples, a 3-factor structure (community sample adjusted \( \chi^2 = 2.95, \) RMSEA = .07, CFI = .95; clinical sample adjusted \( \chi^2 = 3.98, \) RMSEA = .07, CFI = .98) provided a level of model fit that was at least as good as the 1-factor structure (community sample adjusted \( \chi^2 = 5.01, \) RMSEA = .11, CFI = .87; clinical sample adjusted \( \chi^2 = 8.87, \) RMSEA = .11, CFI = .92). In addition to a single measure of
sleepiness, the ESS can provide subscale scores that relate to three underlying levels of somnificity. These findings suggest that the ESS can be used to measure an individual’s overall sleep propensity as well as more specific measures of sleep propensity in low, moderate and high levels of situational somnificity. Due to the findings of this chapter, the ESS was included as a covariate in Study 3, run in the model as a single factor, and as three factors to evaluate model fit.

Study 3 (Chapters 5) determined the influence of hypoxia and sleep fragmentation on cognition in OSA, while controlling for potentially confounding variables including subjective sleepiness (using results from Study 2), age and premorbid intelligence. Participants with and without OSA (N = 150) were recruited from the general community and a tertiary hospital sleep clinic. All underwent comprehensive, laboratory-based polysomnography and completed assessments of cognition including attention, short-term & long-term memory and executive function. Structural Equation Modelling (SEM) was used to construct a theoretically driven model to examine the relationships between hypoxia and sleep fragmentation, and cognitive function. Increased sleep disturbance was a significant predictor of decreased attention (p = .04) and decreased executive function (p = .05), after controlling for IQ. No significant predictors of memory function were found, and hypoxia was not related to any cognitive domain. Controlling for age removes the significant relationships between sleep fragmentation, attention and executive function.

The implications of the findings are discussed in detail in Chapter 6. This chapter elucidates the strengths, originality of this thesis, and a range of suggestions concerning how future research can move forward from the present research program.
Table of Contents

Abstract........................................................................................................................................ii

Table of Contents.........................................................................................................................vi

List of Tables ..............................................................................................................................ix

List of Figures ............................................................................................................................x

Acknowledgements....................................................................................................................xi

Statement of Contribution............................................................................................................xii

Publications....................................................................................................................................xv

List of Abbreviations ....................................................................................................................xx

Chapter 1 General Introduction .................................................................................................1

Adult Obstructive Sleep Apnea ...................................................................................................3

Cognition in OSA.........................................................................................................................4

Conceptual models of cognitive harm in OSA ..........................................................................5

Measuring nocturnal disturbance in OSA ..................................................................................8

Measurement of cognition in OSA .............................................................................................10

Use of problematic statistical methods ....................................................................................11

Inter-individual differences moderating the impact of OSA on cognition ... 12

Summary and contents of chapters .........................................................................................13

Chapter 2 Executive Function in OSA ....................................................................................16

Preface .........................................................................................................................................16

Abstract .......................................................................................................................................17
List of Tables

Table 1 Subdomains of executive function and tests ........................................... 28
Table 2 Effect sizes for OSA to control and pre to post groups ......................... 33
Table 3 Demographic data for the Clinical and Community participants ........ 47
Table 4 Measurements obtained during polysomnography ............................... 59
Table 5 Sleep stages in healthy adults .............................................................. 64
Table 6 ESS total and subscale scores .............................................................. 77
Table 7 Fit indices for the one and three factor ESS models ............................ 79
Table 8 Descriptive statistics for sleep and cognitive measures ....................... 97
Table 9 Descriptive statistics for the clinical and community samples ............ 99
Table 10 Fit indices for Model 1 ad Model 2 .................................................... 106
List of Figures

Figure 1 Flow chart of study inclusion and exclusion.........................23

Figure 2 Proposed 1 factor structure of the ESS...............................75

Figure 3 Proposed 3 factor structure of the ESS...............................76

Figure 4 Model 1 – Accounting for estimated IQ............................101

Figure 5 Model 2 – Accounting for estimated IQ and age ..................104
Acknowledgments

The work in this thesis was supported by an Australian Postgraduate Award, awarded by the Australian Government, the University of Western Australia ‘Top-up’ Scholarship, and funds from the School of Psychology, University of Western Australia. I gratefully acknowledge the assistance and support of so many people I am blessed to know, far too many to mention here. I could write a thesis of dedication to you all!

My wonderful supervisors, friends, and mentors Professor Romola Bucks, and Winthrop Professor Peter Eastwood, you have taught me so much more than research and writing, and I will carry these lessons and memories forever. Thank you for your patience and guidance.

My wonderful daughter, Madelynn, and the love of my life, Jesse, who are my life, and have taught me play is far more important than work.

My beautiful supportive Mother, Ann Leaver, and Sister, Leoni Leaver, who cooked many wonderful meals, dismissed my absence at family parties, suffered through much grumpiness and still rocked up at my door with coffee, flowers, and encouragement.

To the beautiful people in my life who grace me with their wonderful cooking, terrible stories and exceptional warmth: we have known each other since we were little, I hope we know each other ‘till we grow little again, Jennifer Free and Kimberley Pearman.

The amazing students, who anguished with me when I anguished, laughed when I laughed, and helped me hunt down the crickets plaguing our room because of Colin MacLeod’s anxiety studies; Laila Simpson, Alea Losch, Heather Liebregts, Melissa Burgess, Cindy Cabeleira, Alix Mellor, and Ines Pandzic.

Finally, huge thanks to Professor David Hillman, Christine McGuire, and the wonderful staff at Sir Charles Gairdner Hospital, West Australian Sleep Disorders Institute and Sleep Clinic, who gave me so many hours and so much guidance on the right buttons to push (people and machines).
Statement of Original Contribution

The work contained in this thesis has not previously been submitted for a degree or diploma at any other higher education institution. This thesis is entirely my own work and has been accomplished during enrolment in the degree of Doctor of Philosophy (Psychology). All external sources have been acknowledged.

Components of the research were conducted in collaboration with other researchers or institutions, as part of a larger study titled “Predicting Usage of Continuous Positive Airway Pressure in Obstructive Sleep Apnoea” (PUCOSA), directed by Professor Romola Bucks and Professor Timothy Skinner, in collaboration with Professor David Hillman, West Australian Sleep Disorders Research Institute, Sir Charles Gardiner Hospital, Winthrop Professor Peter Eastwood, the Centre for Sleep Science, and the West Australian Participants Pool, as outlined in the Acknowledgements section. However, the design of the current project, data analysis, and the preparation of this thesis have been completed by the candidate alone, with guidance from her supervisors.

As part of “PUCOSA”, the student carried out the following tasks: trained three other researchers to administer assessments and to carry out study administration; assisted with drafting the ethics committee application/amendments; co-designed the protocol for contacting and testing participants; co-created a standardised neuropsychological assessment manual; set up and configured the computerised assessments used; created and administered an online survey; contacted by phone and post numerous study participants in bi-weekly recruitment sessions (for a total of 24 months); conducted the neuropsychological testing of 23 participants at baseline and at three months follow-up (total of 69 hours of face-to-face participant testing); conducted 24 over-night sleep studies at the centre for sleep
science (total of 288 hours of testing); conducted weekly visits to Sir Charles Gardiner Hospital to collect questionnaire data; gave one presentation to the hospital staff about study results/updates and to obtain feedback from the staff so as to review and refine study procedures; gave seven public presentations to Rotary, Lions, and adult education centres as part of community service and recruitment; entered neuropsychological assessment data into SPSS; assisted with the creation of a data entry handbook, defining variables and variable labels and their coding; approached participants at the hospital to complete PUCOSA questionnaires; liaised with the hospital about retrieval of descriptive and medical history data from hospital databases, including assisting with extracting data; and organised, attended and/or contributed to monthly PUCOSA management meetings for a total of 36 months.

The articles that were published as a result of the work undertaken for this thesis are included in chapters 2, 4 and 5. The student undertook a large portion of the data collection, completed all analyses, formulated and wrote the papers. The other authors on the papers provided intellectual input, supervised data collection procedures, advised on the analyses and interpretation, and assisted with formulation and editing of the final papers.

More specifically, in chapter 2, the author RSB helped to provided intellectual input into the conceptualisation of the concept, conducted quality assessment and categorisation of the papers included in the final analysis, and edited the manuscript.

In chapter 4, the authors RSB and TS provided intellectual input, advised on the analyses and interpretation, and assisted with editing the manuscript. The author JC assisted with data collection and edited the final manuscript. The author PE provided intellectual input and edited the final manuscript.
In chapter 5, the author RSB provided intellectual input, advised on the analyses and interpretation, and assisted with editing the manuscript. The authors TS, DH and PE provided intellectual input and edited the final manuscript.

In chapter 6, the author RSB provided intellectual input, and assisted with editing the manuscript. The author PE assisted with editing the manuscript.
Publications

Components of this thesis presented at conferences


**Olaithe, M.** & Bucks, R.S., A Meta-analysis of executive dysfunction in OSA: Before and after treatment, and correlates of nocturnal symptoms, Sleep up Top, Australasian Sleep Association (ASA), Darwin, 10-13 October 2012.


Other abstracts published during the course of this thesis


**Bucks, R.S., Olaithe, M., Eastwood, P. R., Skinner, T. C., & Hillman, D. (2013)** Impact of Sleep Disordered Breathing on self-reported memory function: It’s


Association (ASA), Sleep Down Under 2010 - Biodiversity of Sleep. Christchurch, New Zealand, 21-23 October 2010 *Journal of Sleep and Biological Rhythms, 8*(Suppl. 1), A54.


Peer reviewed publications

(Publications from this thesis are marked with an **)  

(Chapter 5) **Olaithe, M., Skinner, T., Hillman, D., Eastwood, P. & Bucks, R., (Submitted Dec 2013) Cognitive Dysfunction in Obstructive Sleep Apnoea; evaluating the contributions of sleep fragmentation and hypoxia. Sleep and Breathing


# List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AHI</td>
<td>Apnoea Hypopnoea Index</td>
</tr>
<tr>
<td>BMI</td>
<td>Body Mass Index</td>
</tr>
<tr>
<td>CDR</td>
<td>Cognitive Drug Research Battery</td>
</tr>
<tr>
<td>C-FIT</td>
<td>Culture Fair Intelligence Test</td>
</tr>
<tr>
<td>COWAT</td>
<td>Controlled Oral Word Association Test</td>
</tr>
<tr>
<td>CPAP</td>
<td>Continuous Positive Airway Pressure</td>
</tr>
<tr>
<td>DASS-21</td>
<td>Depression Anxiety and Stress Scale, 21 item version</td>
</tr>
<tr>
<td>EF</td>
<td>Executive Functioning</td>
</tr>
<tr>
<td>EF CLOX</td>
<td>Executive Functioning CLOX task</td>
</tr>
<tr>
<td>ISI</td>
<td>Insomnia Severity Index</td>
</tr>
<tr>
<td>LLT</td>
<td>Location Learning Test</td>
</tr>
<tr>
<td>NART</td>
<td>National Adult Reading Test</td>
</tr>
<tr>
<td>OSA</td>
<td>Obstructive Sleep Apnoea</td>
</tr>
<tr>
<td>PSG</td>
<td>Polysomnography</td>
</tr>
<tr>
<td>PUCOSA</td>
<td>Predicting Usage of Continuous Positive Airway Pressure in Obstructive Sleep Apnoea</td>
</tr>
<tr>
<td>REM</td>
<td>Rapid eye movement</td>
</tr>
<tr>
<td>TMT</td>
<td>Trail Making Test</td>
</tr>
</tbody>
</table>
CHAPTER 1: GENERAL INTRODUCTION

Neurocognitive Disturbance in Obstructive Sleep Apnoea: Mechanisms of Cognitive Harm

Sleep had traditionally been understood as a passive event that occurs in the absence of alertness (Data & MacLean, 2007). However, as the technology for measuring changes in physiology has evolved, so too has the understanding of sleep. Modern science reveals that sleep is an important, dynamic process (Data & MacLean, 2007). Scalp electrical recordings were first used in the 1930s visually to identify patterns of brain activity in sleep (AASM, 2007), and in 1967 the first standardized scoring manual was produced by Rechtschaffen & Kales (AASM, 2007).

The specific physiological and neurological features of each sleep stage can be measured and recorded using polysomnography (PSG). PSG is the continuous measurement of the ebb and flow of physiological processes during sleep (Butkov & Lee-Chiong, 2007). Overnight, laboratory PSG is the gold-standard method for diagnosing sleep disorders such as obstructive sleep apnoea (Kryger, 2010). During the overnight sleep study, measurements of oxygen saturation, brain activity, muscle tone, heart activity and rhythm, breathing rhythm, airflow, eye movements, sound and gross body movements are obtained. These physiological measurements are used to stage sleep, and disturbances of sleep.

Human sleep features two broad stages; Rapid Eye Movement sleep (REM), so named for the characteristic eye movements present during this stage, and Non-Rapid Eye Movement sleep (NREM) (Kryger, 2010). NREM sleep can be further broken into 3 sub-stages of sleep; NREM stage 1 (N1), NREM stage 2 (N2) and NREM stage 3 (N3) (AASM, 2007). Stages N1 and N2 are considered shallow,
transitory sleep stages, and N3 is considered deep sleep, characterised by slow, rhythmic brain activity (AASM, 2007; Kryger, 2010). Healthy adults cycle through each of the four sleep stages in 90-120 minute cycles, totaling approximately 4-6 cycles a night (Carskadon & Dement, 2011). Typically, sleep will start in shallow stages progressing from N1 to N2 then N3, to REM (Carskadon & Dement, 2011). However, this pattern is by no means a clean progression, and a person can experience periods of wake or shallow sleep during or before deeper sleep states or REM sleep. Sleep stages are classified, or ‘scored’, during PSG by the type of waveform present in the EEG signal. The waveform present in an epoch of sleep (30 seconds of recorded sleep), and the pattern of waveforms over the course of the night is known as sleep architecture. Disturbances to sleep are said to disturb sleep architecture, as they change the waveforms of, transitions between and length of stages seen in healthy sleep.

Staging adult sleep and scoring events is guided by specific criteria laid out in The American Academy of Sleep Medicine (AASM) Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications (AASM, 2007). This guide provides a comprehensive set of rules for equipment application and signal evaluation in a PSG. The manual builds on the original scoring manual written by Rechtschaffen and Kales (Rechtschaffen & Kales, 1968). These guidelines define not only sleep stages but also sleep disordered events (AASM, 2007).

A sleep-disordered event is the occurrence of an important disturbance to regular sleep patterns. Such an event may be a pause in breath for greater than 10 seconds (an apnoea), a reduction in breathing (a hypopnoea), regular and periodic
movement of the legs (restless legs), frequent wakenings (sleep fragmentation), or other disturbance (e.g. nocturia, nightmares, or movement during REM).

Some of these events can occur several times a night, even in healthy sleepers. For example, a healthy adult will often wake during the night, many times without awareness (Sforza et al, 2008). These events become problematic only when they reach a critical number of times per night, affecting sleep architecture or affecting daytime functioning (Kryger, 2010). A sleep disorder is a medical disorder disrupting sleep, and is diagnosed when the number of sleep disrupting events is above a critical point (Kryger, 2010). Sleep Disordered Breathing (SDB) is a group of disorders characterised by abnormalities in respiratory patterns during sleep, of which the most common is Obstructive Sleep Apnoea (OSA) (Al-Lawati, Patel, & Ayas, 2009).

**Adult Obstructive Sleep Apnoea (OSA)**

**Description, burden of health, risk factors and complications**

Adult OSA is a frequent and often under-diagnosed condition characterised by repeated upper airway (pharyngeal) collapse resulting in sleep fragmentation and oxygen desaturation. (Butkov & Lee-Chiong, 2007). The condition often manifests in snoring, nocturnal gasping and choking, excessive daytime sleepiness, un-refreshed sleep, poor concentration and memory, fatigue, and reduced quality of life (Al-Lawati et al., 2009; Butkov & Lee-Chiong, 2007; Kryger, 2010). The estimated prevalence of OSA in the general population is 9% of middle-aged women and 27% of middle-aged men (Young, Peppard, & Gottlieb, 2002). However, the clinical prevalence, that is those who experience daytime symptoms and seek assessment (OSA Syndrome), is between 1-5%, leaving a large proportion of people with OSA undiagnosed and untreated (Butkov & Lee-Chiong, 2007).
Untreated OSA is associated with increased healthcare utilization, occupational injuries, and motor-vehicle accidents (Al-Ghanim, Comondore, Fleetham, Marra, & Aya, 2008). Hillman, Murphy, Antic, and Pezzulo (2006a) completed a comprehensive study evaluating the financial cost of sleep disorders in Australia, and estimated that the total economic burden of sleep disorders in Australia was $US7.5 billion in the year 2004, representing 1.4% of the total burden of disease for Australia. Sassani et al. (2004) estimate that, for the United States in the year 2000, there were more than 800,000 motor-vehicle collisions, costing more than $15.9 billion, and claiming more than 14,000 lives, attributable to OSA. These same authors estimated that these figures could be more than halved if OSA were appropriately treated (Sassani et al., 2004).

OSA is effectively treated with adherent use of Continuous Positive Airway Pressure (CPAP) devices. Treatment with CPAP assists, to some extent, with deficits in cognition, improves quality of life and reduces car accidents (Butkov & Lee-Chiong, 2007). However, research demonstrates that between 46-83% of people with OSA are not adherent to the minimum 4 hours a night (Weaver & Grunstein, 2008).

Cognition in OSA

Adequate, undisturbed sleep is theorised to be critical for brain health and cognitive function. Poe, Walsh, and Bjorness (2010) have reported that increases in N3 and REM sleep are associated with new learning, and improvements in task performance. Sleep disturbances such as OSA, disrupt the ability to attain deep sleep states such as N3 and REM, and are associated with cognitive disturbance (Bucks, Olaithe, & Eastwood, 2012).

OSA is also associated with disturbances in attention (Findley et al., 1986), memory and new learning (Bedard, Montplaisir, Richer, Rouleau, & Malo, 1991),
and executive function (Fulda & Schulz, 2003; Saunamäki, Himanen, Polo, & Jehkonen, 2010; Saunamäki & Jehkonen, 2007). A recent meta-review (Bucks et al., 2012) concluded that individuals with OSA show deficits in broadly in attention; in memory, specifically episodic visual and verbal memory, visuospatial/constructional abilities; and, in executive function. Cognitive areas that appear to remain unaffected are language abilities, visual immediate recall, visuospatial learning and psychomotor functions. This review concluded that there was clear evidence for dysfunction within the domains of attention and memory, and probable harm in the domain of executive function, although this has yet to be clarified.

The mechanisms of harm through which OSA causes cognitive dysfunction are less well understood. The dominant model of neural harm and cognitive dysfunction in OSA proposes that harm is caused through long-term episodic hypoxia and sleep fragmentation.

**Conceptual models of cognitive harm in OSA**

Beebe and Gozal (2002) have proposed a conceptual framework based around critical roles for sleep fragmentation and nocturnal hypoxia in the development of cognitive dysfunction in individuals with OSA. In their model, sleep is viewed as a necessary restorative process for efficient executive functioning and reinforcing foundations for learning and memory. The nocturnal disturbances of sleep fragmentation and hypoxia interrupt and corrupt these processes.

**Sleep fragmentation.** OSA causes sleep fragmentation through the interruption of sleep by frequent, brief arousals following a respiratory disturbance, such as an apnoea or hypopnoea (Kryger, 2010). During an apnoea or hypopnea the individual exhibits partial or full airway obstruction, this restricted flow causes the person to arouse, which restarts breathing. This can happen many times a night.
Sleep fragmentation in OSA is proposed to contribute to neurocognitive dysfunction, specifically decrements in attention (Verstraeten, 2007), and memory (Daurat, Foret, Bret-Dibat, Fureix, & Tiberge, 2008). Sleep fragmentation contributes to these dysfunctions by slowing cognitive processing speed (Torun-Yazihan, Aydin, & Karakas, 2007; Verstraeten, 2007), and disruption of the restorative processes of sleep (Beebe & Gozal, 2002; Tartar et al., 2010). McKenna et al. (2007) have demonstrated that sleep fragmentation elevates behavioural, electrographic, and neuro-chemical measures of sleepiness in rats.

The gold standard treatment in OSA, is with Continuous Positive Airway Pressure (CPAP) devices. These devices apply gentle air pressure into the pharynx in order to maintain airway patency (Sullivan, Berthon-Jones, Issa, & Eves, 1981). Treatment with CPAP improves quality of life and reduces car accidents and, to some extent, assists with deficits in cognition (Weaver & Grunstein, 2008). However, between 46-83% of people with OSA use their device for less than 4 of use hours of use per night (Weaver & Grunstein, 2008). Such low usage may also result in inadequate treatment of OSA-related cognitive impairment.

In OSA, even with successful amelioration of sleep fragmentation with CPAP treatment, neurocognitive dysfunction often persists (Beebe, Groesz, Wells, Nichols, & McGee, 2003). There is evidence of cell death and structural abnormalities in regions of the brain associated with memory (hippocampus), attention and executive function (frontal cortices) in animal (Xu et al., 2004) and human models (Zhang, Lin, Shunwei, Yuping, & Luning, 2011). As such, it seems likely that these deficits are a primary and direct consequence of OSA (see Beebe and Gozal (2002) for a review; but see Durmer and Dinges (2005b) for evidence that similar frontal changes are also seen post sleep loss), rather than as a secondary result of cognitive slowing.
Neurocognitive Disturbance in OSA

and sleep disturbance. Indeed, one strong proponent of the secondary deficits argument (Verstraeten, Cluydts, Pevemagie, & Hoffman, 2004), conducted a study controlling for attention deficits while exploring executive dysfunction differences between OSA and controls. They concluded that while most of the executive difficulties were secondary to attention, deficits in the ‘Shifting’ facet of executive function remained.

**Hypoxia.** Other authors argue that the deficits remaining after successful CPAP treatment are caused by the long-term night-time hypoxia experienced by individuals with OSA (Beebe, 2006). There is a substantial body of literature that suggests that hypoxia damages the central nervous system through oxidative damage and contributes to chronic and untreatable cognitive dysfunction (Tsai, 2010).

An oxidant is a toxic substance that can cause oxidative damage to proteins and lipids within the human body (Butkov & Lee-Chiong, 2007). Peroxidation is the process by which free radicals or oxidants disrupt the structure of the cellular membrane of proteins and lipids. Free radicals are produced in normal cellular respiration, however, in healthy individuals systems have evolved to minimise oxidative damage and eliminate excess free radicals (Butkov & Lee-Chiong, 2007). In individuals with OSA these systems either do not function at their fullest capacity or are burdened by an over-production of free radicals (Butkov & Lee-Chiong, 2007). What results from a long chain of oxidative damage, is damage to cellular membranes and the production of mutagenic and carcinogenic end products (Xu et al., 2004). Chronic intermittent hypoxia (CIH) is the long-term cycle of deoxygenation and reoxygenation of bodily tissues, as is seen in OSA.

CIH is believed to contribute to memory and executive dysfunction seen in OSA (Beebe & Gozal, 2002). Neuroimaging shows that individuals with OSA
experience reduced cell neurogenesis and density of the hippocampus (Guzman-Marin, Bashir, Suntsova, Szymusiak, & McGinty, 2007), the frontal cortex (Hatipoglu & Rubinstein, 2004), and generalised grey matter reductions (Macey et al., 2003). Additionally, CIH has been associated with decline in the cognitive domains associated with these brain regions: memory, attention/speed of processing, and executive functioning (Beebe, 2006; Beebe et al., 2003).

Summary of proposed mechanisms of harm. Despite strong evidence that sleep fragmentation and hypoxia have the potential to cause cognitive harm and dysfunction, and clear evidence that individuals with OSA exhibit cognitive dysfunction, there remains a dearth of knowledge about the nature of the relationship between cognitive dysfunction and nocturnal disturbance. Studies have consistently failed to find a ‘dose-response’ relationship between the severity of OSA (either hypoxia or fragmentation) and the severity of cognitive deficits found. Indeed, a search of the literature reveals no empirical test of Beebe and Gozal’s full conceptual framework, and conflicting results for studies investigating the relationship between OSA and cognition (Aloia, Arnedt, Davis, Riggs, & Byrd, 2004). This may be for one of several reasons; 1) use of measures of nocturnal indices of disturbance that do not separate hypoxia and sleep fragmentation, 2) a lack of theoretically-driven measurement of cognitive domains, 3) use of problematic statistical methods, or 4) a failure to account for potentially confounding, inter-individual factors such as premorbid IQ, sleepiness and age. Each is considered, in turn, below.

Measuring nocturnal disturbance in OSA

People with OSA exhibit both sleep fragmentation and hypoxia, however these indices are not captured by the primary measure of disease severity in OSA, the Apnoea Hypopnoea Index (AHI). The AHI is a summation of the total number of
times an individual experiences an apnoea and/or hypopnoea. However, individuals can differ on the magnitude and length of desaturation during these events (Kryger, 2010). The AHI makes the assumption that multiple, brief, shallow desaturations are as problematic for cognition, and other functions affected by OSA (e.g., the cardiovascular system, and mood), as fewer, deeper, full obstructions leading to profound oxygen desaturation. People can also exhibit more arousals than marked by apnoeas and hypopnoeas, since many respiratory disturbances do not meet the thresholds given by the AASM (AASM, 1999). It is possible that the relative imprecision of the AHI is the reason that research has consistently failed to find an association between OSA severity (indexed by the AHI) and cognitive dysfunction (Aloia et al., 2004).

The studies that have considered sleep fragmentation or hypoxia in isolation suggest differential harm. This idea is captured in both conceptual models (Beebe, 2006; Beebe et al., 2003), and research. There is evidence to suggest that attention may be more impaired by sleep fragmentation (Verstraeten, 2007), whilst hypoxia greatly affects memory and executive functioning (Beebe & Gozal, 2002). This proposal is tentative at best, and needs further examination. A review by (Aloia et al., 2004) examined the relationship between hypoxia and sleep fragmentation and domains of cognitive dysfunction, across 37 peer-reviewed papers, of which 11 had examined the relationship between some measure of nocturnal disturbance and a cognitive domain. The review revealed that approximately half the papers reported a relationship between attention and hypoxia (5 of 10 papers) and, attention and sleep fragmentation (6 of 9 papers), and less than half reported a relationship between executive function or memory and hypoxia (3 of 10 and 2 of 8 papers, respectively) and/or sleep fragmentation (2 of 9 and 2 of 7 papers, respectively). The authors
concluded that no relationship could be established between the degree of cognitive
dysfunction and the extent of hypoxia and sleep fragmentation. However, they
highlighted that there were many issues with this conclusion. For instance, the many
different papers utilized a variety of nocturnal disturbance measures which may not
measure the same aspect of disturbance. For example in calculating sleep
fragmentation Aloia et. al (2004) necessarily combined measures of total arousal
(ArI) with global measures of disturbance (AHI). The authors also aggregated many
different cognitive assessments across studies (more on this point below). The
authors concluded by saying their results were equivocal, and more studies needed to
be conducted to elucidate the link between cognition and mechanisms of harm in
OSA.

Despite no clear link between cognition and mechanisms of harm from meta-
analyses and laboratory based studies neuroimaging studies do show that sleep
disturbance, hypoxia in particular, visibly affects the brain (Canessa et al., 2011).
This harm occurs through similar mechanisms to those that cause stress and damage
in the cardiovascular system (Hamilton & Naughton, 2013), without the usual
protective mechanisms such as vasodilation, probably due to the cyclic nature of
hypoxia in OSA (Kato et al., 2000).

Despite neuroimaging and experimental evidence that sleep fragmentation and
hypoxia do cause harm to brain structures and cognitive function, a lack of clarity
remains in the cognitive profile of individuals with OSA, and the relationship
between cognitive dysfunction and mechanisms of harm. This may be due to several
reasons, of which 3 are commonplace in the literature; 1) a lack of theoretically
sound division and measurement of cognition domains, 2) use of problematic
statistical methods, and 3) no measurement of inter-individual differences such as age or Premorbid IQ that can moderate the impact of OSA on cognition.

**Measurement of cognition in OSA**

Further complicating an understanding of the dose-response relationship between nocturnal harm and cognition is the wide range of tests used in different research papers. Researchers have utilised an array of tasks to measure whole domains of cognition such as memory or executive function. However, these different tasks likely capture different facets of the domain. For example, executive function in OSA has been assessed by tasks as diverse as the Trail Making Test (a test of the ability to join numbers and letters in alternating sequence, as quickly as possible, 1 – A – 2- B – 3 –C and so on), and verbal fluency (a test of the ability to generate words beginning with a letter, e.g. F, A. and S) (Saunamäki & Jehkonen, 2007). Each of these tasks captures a different facet of executive function: set-shifting or cognitive flexibility and generativity, respectively. A review of the impact of OSA on episodic memory by Wallace and Bucks (2012) highlighted the need to separate cognitive domains into sub-domains. Wallace and Bucks (2012) showed that the discrepant findings in episodic memory testing may have been a function of deficits in some domains of memory (verbal episodic memory (immediate recall, delayed recall, learning and recognition) and visuo-spatial episodic memory (immediate and delayed recall)), but not visual immediate recall or visuo-spatial learning.

Likewise, an examination of the many tasks used to examine executive function, theoretically divided, may explain some of the discrepant findings within the literature. Chapter 2 reports the first meta-analysis of executive deficit in OSA that considers this issue in detail.
Use of problematic statistical methods

Many examinations of the relationship between cognition and nocturnal disturbance in OSA have used small sample sizes (Beebe et al., 2003) and traditional regression techniques. These regression techniques are unable to capture individual variation and measurement error (Byrne, 2010). This lack of precision may contribute to the differential findings in the literature.

In order to account for measurement error and individual variation the present thesis used Structural Equation Modelling (SEM). SEM is an extension of multiple regression designed to test a set of hypothesised relationships between variables, estimated simultaneously (Ullman & Bentler, 2012). It provides a mechanism through which to examine relationships between hypothesized constructs, whilst controlling for individual differences, such as premorbid intelligence and sleepiness, and accounting for measurement error.

SEM is particularly useful when examining hypothesized constructs such as memory, attention, and executive function, as it can account for measurement error that naturally exists between the ‘pure’ construct and its measurement, providing a stringent test of the latent structure (Byrne, 2010; Iaccobucci, 2010; MacCallum & Austin, 2010). In addition, it allow the simultaneous exploration of the impact of OSA (through hypoxia and sleep fragmentation) on multiple aspects of cognition (i.e. memory, attention and executive function), thus reducing the risk of a Type 1 error which would otherwise arise from testing the impact on each cognitive outcome variable individually (Byrne, 2010).
Inter-individual differences moderating the impact of OSA on cognition

**Premorbid cognitive functioning.** Pre-morbid cognitive functioning alters the neurocognitive expression of OSA (Tsai, 2010). Cognitive reserve is the concept that a high level of pre-morbid cognitive ability acts to 'buffer' the effect of neurocognitive trauma (LaRue, 2010). Alchanatis et al. (2005) reported that high-intelligence (an index of cognitive reserve) participants with OSA had the same attention and alertness patterns as high-intelligence participants without OSA. However, normal-intelligence participants with OSA experienced decline in attention and alertness compared to normal-intelligence controls. These authors theorised that high-intelligence protected participants from the negative impact of OSA on cognition.

**Subjective sleepiness.** Individual differences in subjective sleepiness in OSA may modify the way an individual performs on neurocognitive tests (Alchanatis, Zias, Deligiorgis, Liappas, et al., 2008). This may be because high sleepiness lessens the ability to direct cognitive resources, and attend to the task at hand. Consistent with this view, sleepy individuals perform less well on neurocognitive tasks, than non-sleepy individuals (Durmer & Dinges, 2005b; Naismith, Winter, Gotsopoulos, Hickie, & Cistulli, 2004).

**Age.** Participant age is an important variable in research exploring OSA and cognition. OSA is related to age in two ways. First, the risk of having OSA increases with increasing age (Alchanatis, Zias, Deligiorgis, Chroneou, et al., 2008). Indicating that the older the individual the more likely they are to have OSA. Secondly, OSA is commonly undiagnosed (Young, Evans, Finn, & Palta, 1997) as it can only be
diagnosed with an overnight sleep study. This means that older people may have had undiagnosed OSA for a longer period of time, leading to larger cumulative deficits. Our best estimate of disease duration comes from subjective reports of snoring or demographic risk factors, such as obesity (Marcus, Pothineni, Marcus, & Bisognano, 2014) of which the individual (or partner) may or may not be aware, and which may or may not indicate OSA (Cirignotta et al., 2009).

**Summary and contents of chapters**

OSA causes cognitive dysfunction. Reviews and meta-analyses have clarified the pattern of deficits in attention, and memory, however the pattern of executive dysfunction remains unclear. Hypoxia and sleep fragmentation are the hypothesised mechanisms of harm in OSA, as posited by published conceptual models, however these models and the relationship between nocturnal disturbance (hypoxia and sleep fragmentation) and attention, memory and executive dysfunction have yet to tested.

This thesis aimed to examine and clarify the relationship between cognition, and sleep disturbance in OSA. In particular it examined the relationship between the broad domains of cognition shown to be disordered in OSA: attention, memory, and executive function, and nocturnal disturbance: sleep fragmentation and hypoxia.

In order to do this, first it was necessary to clarify the pattern of executive function deficits. The results of a meta-analysis of executive dysfunction in OSA before and after treatment are presented in Chapter 2. This study was published in the journal *Sleep* (2013); 36 (9); 1297-305.

Detailed information about recruitment, participants, materials and methodology for Chapters 4 and 5 are presented in Chapter 3. As Chapters 4 and 5
were submitted for publication, and due to journal word restrictions, these chapters have little detail on recruitment, participants, materials and methodology.

Premorbid IQ, age and daytime sleepiness, independently affect cognitive performance, and there is a need to control for these inter-individual factors when examining the relationships between cognition and OSA (Alchanatis et al., 2005; Alchanatis, Zias, Deligiorgis, Chroneou, et al., 2008). Age can be quantified, and premorbid IQ can be estimated using psychometrically validated tools (Nelson & Willison, 1991), however, the factor structure of the most widely-used measure of daytime sleepiness, the Epworth Sleepiness Scale (ESS: used in this thesis) has been questioned (Smith et al., 2008). Factor analyses of the ESS have revealed 1, 2, or 3 possible underlying constructs. In order to understand how best to utilise the ESS, Chapter 4 examines the factor structure of the ESS in community and clinical samples using confirmatory factor analysis. This chapter was published in *Sleep and Breathing* (2012); 17(2); 763-9.

The final, investigative chapter used the findings from the meta-analysis and investigation of the ESS, in structural equation models examining the relationships between attention, memory and executive function, and hypoxia and sleep fragmentation, accounting for intelligence, age and sleepiness. This chapter has been submitted to *Sleep and Breathing* for review as of the 17.12.2013.

The final chapter of the thesis, Chapter 6, presents a general discussion of the findings, strengths, original contributions, and future directions.
CHAPTER 2: EXECUTIVE FUNCTION IN OSA

This chapter was published in the journal *Sleep*: Olaíthe, M. and Bucks, R.S. (2013). Executive Dysfunction in OSA Before and After Treatment: A Meta-Analysis. *SLEEP*, 36(9); 1297-305. It is presented below, as published, but formatted for consistency with the rest of the thesis.

Preface

Cognition is affected by OSA, however, until recently the literature has been divided as to the profile of cognitive dysfunction in OSA. Recent reviews have summarised which aspects of attention and memory are impacted by OSA. However, it remains unclear what specific aspects of executive function are impaired.

Chapter 2 presents a meta-analysis reviewing the literature regarding which specific aspects of executive function are impaired in OSA and improved by CPAP.
Abstract

Study Objectives: Obstructive Sleep Apnoea (OSA) is a frequent and often under-diagnosed condition that is associated with upper airway collapse, oxygen desaturation, and sleep fragmentation leading to cognitive dysfunction. There is good meta-analytic evidence that sub-domains of attention and memory are affected by OSA. However, a thorough investigation of the impact of OSA on different sub-domains of executive function has yet to be conducted. This report investigates the impact of OSA and its treatment, in adult patients on five, theorised, sub-domains of executive function.

Design: An extensive literature search was conducted of published and unpublished materials, returning 35 studies that matched selection criteria. Meta-analysis was used to synthesise the results from studies examining the impact of OSA on executive functioning compared to controls (21 studies) and before and after treatment (19 studies); 5 studies met inclusion in both categories.

Measurements: Research papers were selected which assessed five sub-domains of executive function: Shifting, Updating, Inhibition, Generativity and Fluid Reasoning.

Results: All 5 domains of executive function demonstrated medium to very large impairments in OSA independent of age, and disease severity. Furthermore, all sub-domains of executive function demonstrated small to medium improvements with CPAP treatment.

Discussion: Executive function is impaired across all five domains in OSA, these difficulties improve with CPAP treatment. Age and disease severity did not moderate
the effects found, however, further studies are needed exploring the extent of primary and secondary effects, and the impact of age and pre-morbid ability (cognitive reserve).
Executive Dysfunction in OSA Before and After Treatment: A Meta-Analysis

Obstructive sleep apnoea (OSA) is a frequent and often under-diagnosed condition that is associated with upper airway collapse, oxygen desaturation, and sleep fragmentation leading to sleepiness, hypertension, increased risk of cardiac disease, and neurocognitive disturbance. (Al-Lawati et al., 2009; Butkov & Lee-Chiong, 2007; Young, Palta, & Dempsey, 1993) Untreated OSA is associated with increased healthcare utilization, occupational injuries, motor-vehicle accidents (Al-Ghanim et al., 2008; Hillman, Murphy, Antic, & Pezzulo, 2006b; Sassani et al., 2004) and neurocognitive sequelae in memory, attention and executive function. (Butkov & Lee-Chiong, 2007; Tsai, 2010) The gold standard treatment for OSA is Continuous Positive Airway Pressure (CPAP). (Kryger, 2010; Weaver & Grunstein, 2008)

To date, most reviews of cognitive functioning in OSA have inspected cognition as a whole, collapsing research findings into ‘memory’, ‘executive function’ or ‘attention’ domains (for example see, (Aloia et al., 2004; Beebe et al., 2003)). As the evidence base grows, however, it is both possible and desirable to explore the cognitive burden of OSA within subcomponents. Based on current neurocognitive theory, there are functional and biological grounds for segregating cognitive domains or functional systems into such subcomponents. (Adrover-Roig, Sesé, Barceló, & Palmer, 2012; Lezak, Howieson, & Loring, 2004) These subcomponents work in concert to produce what we colloquially know as memory, attention and executive function. (Elliot, 2003; Larner, 2008)

Recently, Wallace and Bucks (Wallace & Bucks, 2012) divided episodic memory into theoretically-driven subcomponents, revealing deficits in individuals with OSA in verbal episodic memory (immediate recall, delayed recall, learning and
Neurocognitive Disturbance in OSA

Neurocognitive Disturbance in OSA (recognition) and visuo-spatial episodic memory (immediate and delayed recall), but not visual immediate recall or visuo-spatial learning. This theoretically-driven division of memory reveals that not all components of memory are dysfunctional in OSA and provides an explanation for the mixed findings in this field. A similar approach might prove fruitful when exploring executive dysfunction in OSA. In a review, Saunamaki et al. (Saunamäki & Jehkonen, 2007) demonstrated that aspects of executive function may also be impaired or preserved in OSA. They divided executive functioning by test, demonstrating deficits in Digit Span Forwards, Corsi Block Tapping task, Double encoding task, Wisconsin Card Sorting Test, Phonemic fluency, Rey-Osterreith Complex Figure test, and Maze tasks. However, by meta-analysing the data by test, this review did not aggregate executive functions using a theoretical framework. In addition, Saunamaki et al. included some tests that do not primarily measure executive function (i.e. Digit Span Forwards, Rey-Osterrieth Complex Figure test, the Trail Making Test Part A and the Corsi Block Tapping task), making it difficult to determine which subcomponents of EF, mapped by which tests, are impaired in OSA.

Executive function is an individually controlled and conscious effort to guide the operation of various cognitive processes and thereby regulate cognition. (Banich, 2009; Elliot, 2003; Funahashi, 2001; Lezak et al., 2004; Miyake, Friedman, Emerson, Witzki, & Howarter, 2000; Suchy, 2009) Like other cognitive domains, executive function is multidimensional. (Chan, Chen, Cheung, Chen, & Cheung, 2006; Lezak et al., 2004; Miyake et al., 2000; Suchy, 2009) Miyake et al., (Miyake et al., 2000) Fisk & Sharp (Fisk & Sharp, 2004) and Adrover-Roig et al. (Adrover-Roig et al., 2012) present an empirical basis for specifying how executive functions are organised, and what roles different subcomponents play. Miyake et al. (Miyake et
al., 2000) divide executive functioning into (a) *Shifting* between tasks or mental sets, (b) *Updating* and monitoring of working memory representations and (c) *Inhibition* of dominant or pre-potent responses. (Lehto, Juuja¨rvi, Kooistra, & Pulkkinen, 2003) Fisk and Sharp (Fisk & Sharp, 2004) and Adrover-Roig et al (Adrover-Roig et al., 2012) utilized this same 3-factor structure, but proposed a fourth component; efficiency of access to long term memory (called Generativity in the present report, for brevity). This four-component structure has been confirmed in multiple populations with factor analysis (exploratory and confirmatory) (Adrover-Roig et al., 2012; Fisk & Sharp, 2004; Lin, Chan, Zheng, Yang, & Wang, 2007; Montgomery, Fisk, Newcombe, & Murphy, 2005).

Lezak, Howieson and Loring (Lezak et al., 2004) and Strauss, Sherman and Spreen (Strauss, Sherman, & Spreen, 2006) define a set of tasks that do not tap executive function per say but rather an overarching system of reasoning and problem solving. These tasks involve complex, higher order abstraction, problem solving and concept formation and include tasks such as Porteus Mazes or Clock drawing tasks. (Lezak et al., 2004; Strauss et al., 2006) The four-factor model defined above does not account for such tasks; however they abound in OSA literature on executive function and are considered a part of executive functioning in neuropsychological theory (Lezak et al., 2004; Strauss et al., 2006). Hence, in the present paper a class of executive function tasks, called Fluid Reasoning, was created to capture this concept.

The present paper builds on past reviews and meta-analyses examining executive functioning in adults with OSA within current neuropsychological theory of executive function. No previous meta-analysis in OSA has assessed EF dysfunction, or the effect of treatment, within these 5, theoretically-motivated
domains: Shifting, Updating, Inhibition, Generativity, and Fluid Reasoning. We addressed three questions: 1) which specific executive functions are affected by the presence of untreated OSA?; 2) if executive functions are impaired, does treatment help to remediate these deficits?; and 3) are any of these effects moderated by publication status, sample source, study design, age, disease severity, or control screening?

Method

Search Strategy

Data for this meta-analysis consisted of empirical articles published in peer-reviewed journals over the past 24 years (Jan 1987 – Nov 2011). Details of the search methodology employed are outlined in Figure 1.
Search terms: (Apnoea OR sleep-disordered breathing) AND (Cognition OR Cognitive ability OR Mental Status OR Neuropsychology OR Memory OR Attention OR Vigilance OR Executive OR Psychomotor)

Electronic Databases searched (Keyword and MeSH explode): Medline R (n = 463), Psych Info (n = 127), PubMed (n = 1757), EMBASE (n = 771), CINAHL (n = 118), CCTR (n = 31), NHS EED (n = 47)

Grey Literature: SIGLE (n = 15), NTIS (n = 1)

Conference proceedings: Conference proceedings citation index science (n = 212)

Dissertations and Theses: Proquest dissertations and theses (n = 71)

Internet: Google scholar (n = 8,070)

Handsearching: Index Medicus, Exerpta Medica, References of included articles (n = 60), Contact with authors of

11,302 duplicates and articles not relevant to topic removed

Titles and abstracts screened (n = 446)

Full text copies retrieved for evaluation (n = 269) using quality assessment criteria

Extracted descriptive data (n = 35 (34 + 1 personal communication)): author/s, publication status, year of publication, study design, sample size, participant details, co-morbidities screened for, source of OSA sample and executive function assessments employed.

Data Analysis:
Calculated effect sizes
Calculated statistical heterogeneity
Publication bias

Abstracts excluded (n = 177) Reasons:
1. Participants did not have primarily obstructive sleep apnoea
2. Paper did not examine executive function
3. Participants were <18 years
4. The article was not in English
5. Paper was not in an appropriate format (e.g. review article)

Studies excluded (n = 234) Reasons:
1. Paper was not in an appropriate format (e.g. review article)
2. The article was not in English
3. Participants were <18 years
4. Participants did not have primarily obstructive sleep apnoea
5. Did not assess EF
6. Test used was inadequately described
7. Test did not have acceptable validity and/or reliability
8. Data were presented in such a way that effect sizes could not be calculated
9. PSG not done on participants
10. PSG conducted >12months before/after cognitive assessment
11. Special population was used (e.g. OSA in TBI or insomnia)
12. Data were published elsewhere
13. Group matching was inappropriate (e.g. IQ higher in one group than the other)

Figure 1. Flow chart of study inclusion and exclusion
An extensive computer assisted literature search was conducted using electronic databases (Keyword and MeSH explode) for published articles (Medline R, PsychInfo, PubMed, EMBASE, CINAHL, CCTR, NHS EED), grey literature (SIGLE, NTIS), conference proceedings (Conference Proceedings Citation Index: Science), dissertations and theses (Proquest Dissertations and Theses), the Internet (Dogpile, Omni Medical search engine, Mednet) and via handsearching (Index Medicus, Exerpta Medica, references of included articles, contact with authors of unpublished studies). Unpublished studies were included in the search, to avoid publication bias.

The terms ‘apnoea OR sleep-disordered breathing’ were combined with ‘Cognition OR Cognitive ability OR Mental Status OR Neuropsychology OR Memory OR Attention OR Vigilance OR Executive OR Psychomotor’. The terms chosen covered a wide range of cognitive functions to capture tests that had been mislabelled or utilised to measure other cognitive domains.

Additional relevant articles were retrieved from the reference lists of studies included in the original search, conference proceedings and dissertations. Furthermore, key authors who have published articles on the relationship between OSA, cognition and CPAP treatment were contacted asking if they were aware of any other relevant published or unpublished studies.

**Study selection criteria**

This review included studies that assessed executive function in adults with OSA as defined by an Apnoea Hypopnoea Index (AHI) of > 5 per hour of sleep. (AASM, 1999) In all instances, except one, studies were excluded if OSA participants were not diagnosed using overnight polysomnography and/or if they did not include a control sample, if group matching was inappropriate (e.g. IQ
statistically different between control and OSA groups) or there were no baseline data (participants were assessed after treatment only). The exception to this rule was Antic et al., (N. Antic, September, 2012; N. A. Antic et al., 2011) where participants were administered overnight oximetry instead of PSG. This paper was included as the oximetry was validated against in-laboratory PSG in a random selection of 50% of the participants.

Additionally, papers were excluded if the PSG was conducted more than 12 months before/after neuropsychological profiling was completed. These studies were excluded as individuals may lose or gain weight, or change their lifestyle habits (e.g. drink, smoke or exercise more or less) which may alter the severity of their OSA. (Cowan & Livingston, 2012; Ong, O’Driscoll, Truby, Naughton, & Hamilton, 2012)

This review considered only studies with adult participants (≥18 years), not from special populations (e.g. people with Down’s syndrome, insomnia, or traumatic brain injury). Research demonstrates that there are etiological differences between adult and childhood OSA (Cheng, Dai, Wu, & Chen, 2012) thus the latter was not addressed here.

The present review aimed to delineate the pattern of executive deficits in OSA, hence studies that included a majority of central or mixed sleep apnoea patients were excluded. Research demonstrates that the pathophysiology, epidemiology, and clinical characteristics of central sleep apnoea and OSA are distinct. (T Young, P.E Peppard, & D. J. Gottlieb, 2002)

Papers were excluded if the tests used were inadequately described such that acceptable validity and/or reliability could not be confirmed, the paper was a review paper not a study, if it reported data already included in the present review (in this instance the most complete data set was selected) or was a cross-over trial. Cross-
over trials were excluded as research does not provide any definitive information regarding length of washout period required. (Phillips et al., 1990; Sullivan et al., 1981; Sullivan & Issa, 1985)

The present study was only able to examine CPAP treatment, as after evaluating studies with the exclusion criteria, there remained an insufficient number of other treatment studies (No oxygen therapy, positional therapy, drug trial or weight loss studies, 1 surgical study, 3 Mandibular Advancement Splint (MAS) studies, 3 studies with mixed treatments). Nor did the present review examine medication studies as these (e.g. modafinil and armodafinil) may alter alertness, cognitive function and judgement without treating underlying nocturnal symptoms. (Estrada, Kelley, Webb, Athy, & Crowley, 2012; Ray et al., 2012) Although research demonstrates that these medications can be helpful in conjunction with CPAP where there is residual sleepiness, the present study aimed to look at the effect of OSA on cognitive function, and such medications may have confounded these results.

Furthermore, studies were not considered if the data were presented in such a way that effect sizes could not be calculated even after contact with the author. We contacted 32 authors for further details on 33 research papers. Five authors or their representatives replied. Of these, 2 authors were deceased, 1 had no more detail to provide, and 2 emailed further data. Data received from N. Antic (N. Antic, September, 2012) was utilised in the present meta-review. We also received further data from M. Barnes, (Barnes et al., 2002) however this was later excluded as it was from a cross-over trial.

Given that it is difficult to keep participants and experimenters blinded to group (OSA or Control, Treatment or No treatment) in OSA studies when assessing
neuropsychological function, (Redline et al., 1997) this review did not exclude unblinded studies.

Finally the present paper included studies in which controls were screened using PSG or with questionnaires. Despite the risk of undetected OSA in the control sample, (Sharwood et al., 2012) evidence from Wallace and Bucks (Wallace & Bucks, 2012) suggested that comparing OSA participants with controls within memory domains, with and without PSG screening of controls, did not dramatically reduce the significance of the effects found, and would have reduced the number of studies available per subcomponent. Rather, the present meta-analysis considered control screening method as a moderator instead.

Quality Assessment

The authors (MO, RSB) independently reviewed articles according to the selection criteria. Where there were disagreements about whether or not to include an article, the authors discussed and came to an agreement.

The data for all included studies were extracted and coded by the first author. The second author extracted and coded the data for 10 randomly selected studies. The intra-class correlation coefficient between the data extracted by the first and second author was $r = 0.99$ (CI: 0.99-1.00).

Study Categorization

Included studies ($N = 35$; 34 studies + 1 personal communication) were divided into two non-exclusive categories, (1) comparisons of pre-treatment OSA groups to controls were used to identify the specific pattern of executive dysfunction present in untreated OSA ($n = 21$), and (2) CPAP treatment efficacy studies were used to establish whether executive impairments were permanent in OSA ($n = 19$): 5 studies met criteria for inclusion in both groups.
**Categorisation of Executive Function**

Table 1 presents the sub-domains of executive function and the tests ascribed to these sub-domains.

*Table 1.*

Sub domains of executive function and the tests accessing each sub-domain.

<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
<th>Tests that tap this cognitive skill</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Shifting</strong></td>
<td>Shifting back and forth between multiple tasks, operations or mental sets. Requires the disengagement of an irrelevant task set and subsequent engagement of a relevant task set when a new operation must be performed on a set of stimuli, necessary to overcome proactive interference or negative priming due to having recently performed a different operation.</td>
<td>- Wisconsin Card Sorting Test¹ (18)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Trails B (3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Switching task (26)</td>
</tr>
<tr>
<td><strong>Updating</strong></td>
<td>Updating and monitoring of working memory representations. Requires the monitoring and coding of incoming information for relevance to the task, and then appropriately revising items held in working memory by replacing old, no longer relevant information with new more relevant information. Dynamically manipulate the contents of working memory.</td>
<td>- N-back tasks (24)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Digit span backwards¹ (9)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- WAIS-R Arithmetic (13)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- ANAM Mathematical processing (14)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- ANAM running memory (17)</td>
</tr>
<tr>
<td><strong>Inhibition</strong></td>
<td>Inhibition of prepotent, dominant or automatic responses when necessary. An internally generated act of control.</td>
<td>- Towers¹ (19)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Stroop task (6)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Go No-go task (4)</td>
</tr>
<tr>
<td><strong>Generativity</strong></td>
<td>Speed and efficiency of access to long-term memory. An independent ability to create, generate or produce content without any input from what or whom?</td>
<td>- Verbal fluency tasks¹ (2)</td>
</tr>
<tr>
<td><strong>Fluid reasoning</strong></td>
<td>Concept formation/abstraction &amp; problem solving tasks. An intentional cognitive process that does not occur automatically, but rather involves the use of deliberate and controlled mental actions to solve novel problems.</td>
<td>- Mazes² (12)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Ravens progressive matrices² (1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Picture completion² (22)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- WAIS-R Picture arrangement² (11)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- WAIS-R Block design (8)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Stockings of Cambridge (23)</td>
</tr>
</tbody>
</table>
Data Extraction and Analysis

Data extracted and coded from the final articles included author/s, whether published or not, journal and year of publication (if applicable), study design, sample size and participant details when available (gender, years of formal education, body mass index (BMI), age, diagnostic criteria, AHI or RDI, oxygen desaturation indices [time spent below 90%: CT90] and sleep fragmentation indices [Arousal Index: ArI]), source of OSA sample (clinical vs. community) and neuropsychological assessments employed (See Appendix 1 for further details). Means, standard deviations and sample size were extracted to examine the relationships between the variables of interest.

In the instance that participants had been assessed at multiple time points after CPAP treatment we chose the most distant time point, as we wanted to examine the effect of continuous treatment on executive dysfunction. Furthermore, if participants were divided into compliant (> 4hrs for 80% of nights) and non-compliant users, only the compliant user information was included. These two decisions were made so as best to evaluate the benefit to individuals who utilise their CPAP devices as
recommended over the long term. These choices led us to lose only 1 subsample from a paper, but no whole papers.

**Data Processing**

The program Comprehensive Meta-Analysis version 2.2.064 (Borenstein, Hedges, Higgins, & Rothstein, 2005) was used to synthesize data, calculate effect sizes, and create Forest plots.

**Results**

**Description of Studies**

From the articles identified (N = 35), 21 studies compared people with untreated OSA to a control sample, 19 compared people with OSA before and after treatment; 5 studies had both a comparison to controls participants and neurocognitive testing performance before and after treatment. These studies represent 40 samples. Only 1 study was recruited from a community setting, (Quan et al., 2006) thereby making 98% of studies from clinical settings.

In total, there were 551 healthy controls (74% male; Mean age 49.46±8.96 years; Mean ESS 5.52±2.41; Mean BMI 25.42±2.50) and 1010 participants with OSA (81% male; Mean age 50.40±7.43 years; Mean AHI/RDI 47.58±15.98; Mean Arousal Index (ArI) 36.15±17.66; Mean Cumulative Time below 90% oxygen saturation (CT90) 40.07±28.55; Mean education 13.73±1.48years; Mean months of CPAP treatment 2.89±2.22 months; Mean Hours CPAP use per night 5.34±1.01; Mean Epworth Sleepiness Score (ESS) 12.02±2.38; Mean Body Mass Index (BMI) 33.12±2.76). Individual study details for sample size, publication source, age, indices of disease severity (AHI or RDI), oxygen desaturation (CT90) and sleep fragmentation (ArI), years of education, and length of CPAP treatment are given in Appendix 1.
In the present meta-analysis only studies with matched control and OSA participant variables were chosen, except that as expected, the control and OSA groups differed significantly in BMI, $t(1228) = -56.03, p < 0.001$, and ESS, $t(1118) = -51.15, p < 0.001$ scores.

**Calculation of Effect Sizes**

Random effect sizes were calculated. The random effects model assumes that each study has a different underlying ‘true’ effect size due to differing sample demographic variables. (Rosenthal, 1995) In the present meta-analysis, samples differed on such variables as disease severity, age, gender, screening measures, oxygen saturation and sleep fragmentation (See Appendix 1). A random effects model accounts for these between-studies differences, as well as within-study participant differences.

An effect was calculated for each sample across each of the five domains. For comparisons between the OSA group and healthy controls, Cohen’s $d$ was calculated according to the formula $d = \frac{\text{mean controls} - \text{mean OSA}}{\text{pooled SD}}$, where effect sizes of $d \leq 0.20$ are considered small, $d = 0.50$ medium, $d \geq 0.80$ large and $d \geq 1.00$ very large. (Cohen, 1977) Larger, positive effects indicate poorer performance for the OSA group.

In OSA group pre-treatment compared to post-treatment, Cohen’s $d$ effect sizes were calculated with the formula $d = \frac{\text{mean post-treatment} - \text{mean pre-treatment}}{\text{pooled SD}}$. Higher scores indicate greater improvement post CPAP treatment.

The random effect size estimates between OSA and control groups for each domain are displayed in Table 2. All five sub-domains of executive function were impaired, compared to controls. A very large effect was noted for Inhibition, a large
effect was present for Updating and Fluid Reasoning, and medium effect sizes were present for Shifting and Generativity. These results indicate medium to very large deficits in executive function performance across all five domains in individuals with OSA, when compared to control.

The random effect size estimates between OSA pre-treatment and post-treatment are displayed in Table 2. All five sub-domains of executive function demonstrated improvement after CPAP treatment. Medium effect sizes were found in Shifting and Inhibition, and small effect sizes were noted in Updating, Fluid Reasoning and Generativity. These results indicate small to medium size improvements across all five domains of executive function with CPAP treatment.
### Table 2.

Mean effect sizes for the differences between OSA to control groups, and pre-treatment to post-treatment groups

<table>
<thead>
<tr>
<th>Domain</th>
<th>N</th>
<th>d</th>
<th>95% CI</th>
<th>Z</th>
<th>p</th>
<th>Q (df)</th>
<th>p</th>
<th>Tau</th>
<th>I²</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>OSA to controls</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shifting</td>
<td>15</td>
<td>0.53</td>
<td>0.38</td>
<td>0.92</td>
<td>4.78</td>
<td>&lt; 0.001</td>
<td>41.68(14)</td>
<td>&lt; 0.001</td>
<td>0.42</td>
</tr>
<tr>
<td>Updating</td>
<td>14</td>
<td>0.91</td>
<td>0.49</td>
<td>1.32</td>
<td>4.30</td>
<td>&lt; 0.001</td>
<td>71.12(13)</td>
<td>&lt; 0.001</td>
<td>0.71</td>
</tr>
<tr>
<td>Inhibition</td>
<td>9</td>
<td>1.12</td>
<td>0.55</td>
<td>1.69</td>
<td>3.83</td>
<td>&lt; 0.001</td>
<td>57.17(8)</td>
<td>&lt; 0.001</td>
<td>0.79</td>
</tr>
<tr>
<td>Generativity</td>
<td>8</td>
<td>0.59</td>
<td>0.34</td>
<td>0.85</td>
<td>4.58</td>
<td>&lt; 0.001</td>
<td>8.25(13)</td>
<td>0.311</td>
<td>0.14</td>
</tr>
<tr>
<td>Fluid Reasoning</td>
<td>11</td>
<td>0.80</td>
<td>0.42</td>
<td>1.19</td>
<td>4.11</td>
<td>&lt; 0.001</td>
<td>44.12(10)</td>
<td>&lt; 0.001</td>
<td>0.55</td>
</tr>
<tr>
<td><strong>Pre to post CPAP treatment</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shifting</td>
<td>14</td>
<td>0.66</td>
<td>0.31</td>
<td>1.00</td>
<td>3.75</td>
<td>&lt; 0.001</td>
<td>47.12(13)</td>
<td>&lt; 0.001</td>
<td>0.58</td>
</tr>
<tr>
<td>Updating</td>
<td>10</td>
<td>0.46</td>
<td>0.15</td>
<td>0.77</td>
<td>2.90</td>
<td>0.021</td>
<td>15.94(9)</td>
<td>0.068</td>
<td>0.33</td>
</tr>
<tr>
<td>Inhibition</td>
<td>9</td>
<td>0.57</td>
<td>0.31</td>
<td>0.86</td>
<td>4.17</td>
<td>&lt; 0.001</td>
<td>10.75(8)</td>
<td>0.216</td>
<td>0.21</td>
</tr>
<tr>
<td>Generativity</td>
<td>8</td>
<td>0.33</td>
<td>0.13</td>
<td>0.53</td>
<td>3.27</td>
<td>&lt; 0.001</td>
<td>5.21(7)</td>
<td>0.635</td>
<td>0.00</td>
</tr>
<tr>
<td>Fluid Reasoning</td>
<td>10</td>
<td>0.37</td>
<td>0.18</td>
<td>0.56</td>
<td>3.85</td>
<td>&lt; 0.001</td>
<td>10.02(9)</td>
<td>0.349</td>
<td>0.10</td>
</tr>
</tbody>
</table>

Forest plots for individual studies are available in Appendix 1.
Heterogeneity

Heterogeneity of effect sizes is a measure of difference between a study’s true effect size and the observed effect size. The true effect size is the actual effect size in the underlying population, whilst the observed effect size is the effect measured in the sample. (Borenstein, Hedges, Higgins, & Rothstein, 2009) Heterogeneity was investigated visually with Forest plots, and statistically using Cochrane’s Q statistic, the $T^2$ and $I^2$ statistics (See Tables 2 and 3, and Appendix 1, for individual Forest plots). When the Q statistic is significant, this suggests there is a significant difference between the observed and true effect. However the Q statistic is vulnerable to small sample size, hence $T^2$ and $I^2$ can provide an estimate of the proportion of real variance caused by extraneous study variables such as age or test used. (Borenstein et al., 2009) In any instance that Q was significant, $I^2$ was examined to quantify the degree of heterogeneity.

For the comparisons between controls and individuals with OSA, significant heterogeneity was present in all domains. However, the $I^2$ ranged between 77.33 and 86.01 suggesting at least 77% of the variance was generated from real, between-group differences.

For the comparisons of pre- and post-treatment, significant heterogeneity was present in the domain, Shifting. However, the $I^2$ was 72.41 suggesting at least 72% of the variance was generated from real, between-time differences.

Moderator analysis, using a random effects model, was conducted to explore potential, between-study differences that may explain this heterogeneity.

Moderator Analysis

Pre-treatment OSA to controls. Moderators investigated were age (Group 1 < 50, Group 2 ≥ 50 years), stringent inclusion/exclusion criteria (Group 1 = no,
Group 2 = yes), control group selection criteria (Group 1 = PSG, Group 2 = Questionnaire), and publication status (Group 1 = Published, Group 2 = Not published). Papers were considered to have stringent inclusion/exclusion criteria if they excluded participants with factors that could potentially affect cognition, such as a history of traumatic brain injury, certain medications or diseases. None of these moderators changed the effect significantly.

We were unable to examine the impact of disease severity as measured by AHI as there were insufficient studies reporting effects of moderate OSA (AHI 15-29) in each of the five sub-domains. A reviewer suggested dividing only those with AHI ≥30 into two groups and rerunning analyses. Accordingly, we divided the samples into severe OSA (AHI 30-50, N = 15) and very severe OSA (AHI 51+, N = 17). Where there were sufficient samples (i.e. 2 or more) for each executive domain, all effects remained significant and severity did not moderate the findings.

**Pre-treatment to post-treatment differences.** Between-study differences were also investigated using moderator analysis. Moderators were age, stringent inclusion/exclusion criteria, control selection criteria (PSG or questionnaire), length of CPAP use, and publication status. No variables significantly moderated the effect of CPAP treatment on the executive burden of OSA.

Likewise, disease severity, as measured by AHI, could not be examined as there were insufficient studies reporting effects of moderate OSA in each of the five sub-domains. As above, as recommended by a reviewer, we explored the impact of severity within individuals with AHI ≥ 30. As before, all effects remained significant, and severity did not moderate the findings.

Furthermore, we were unable to examine the impact of CPAP compliance on effect sizes as only one study divided the participants into compliant and non-
compliant users, all other studies excluded non-compliant individuals or reported only the group mean compliance which was always above 4hrs.

However, and as recommended by a reviewer, we explored months on CPAP by dividing samples into short (0-5 months, $N = 13$) and long ($\geq 5.5$ months, $N = 5$) term CPAP use. Where there were sufficient samples (i.e. 2 or more) for each executive domain, all effects remained significant and months on CPAP did not moderate the findings.

**Risk of Publication Bias**

There is evidence to suggest that studies with a significant result are more likely to be published, and that published studies are more likely to be available for meta-analysis. (Borenstein et al., 2009) Publication bias was inspected visually using funnel plots. These were asymmetrical, indicating the presence of bias. Egger’s test for asymmetry (Egger, Smith, Schneider, & Minder, 1997) was used to investigate this further.

**Pre-treatment OSA to controls.** For the OSA to control samples, Egger’s test was non-significant for Inhibition (intercept 4.79; 95% CI: 0.04 to 9.54; $p = 0.05$), and Updating (intercept 5.84; 95% CI: -0.63 to 12.32; $p = 0.07$), but was significant for Fluid Reasoning (intercept 4.49; 95% CI: 0.84 to 8.13; $p = 0.02$), Generativity (intercept 4.55; 95% CI: 0.65 to 8.45; $p = 0.03$) and Shifting (intercept 4.48; 95% CI: 1.72 to 7.25; $p = 0.002$). However, Rosenthal’s Fail-safe N, which represents the number of studies needed to create an overall non-significant effect, (Rosenthal, 1979) was 172 non-significant studies for Fluid Reasoning, 49 studies for Generativity and 232 studies for Shifting.

Duval and Tweedie’s *Trim and Fill* procedure (Duval & Tweedie, 2000) was used to determine the best estimate of an unbiased, overall effect size for the OSA of
control samples for the domains of Fluid Reasoning, Generativity and Shifting. For Fluid Reasoning, the overall effect size was reduced from a large effect of 0.80 to a small effect of 0.26. For Generativity the overall effect size shifted from a medium effect of 0.59 to a small effect of 0.45. For Shifting, the overall effect size was reduced from a medium effect of 0.53 to a small effect of 0.37. This suggests publication bias may be inflating the estimates in the domains of Fluid Reasoning, Generativity and Shifting, but that there were still significant differences between those with OSA and controls in these domains.

**Pre-treatment to post-treatment differences.** For the pre- to post-treatment effects, Egger’s test was non-significant for all five domains; Shifting (intercept 1.62; 95% CI: -1.39 to 4.63; \( p = 0.26 \)), Generativity (intercept 0.70; 95% CI: -1.16 to 2.56; \( p = 0.39 \)), Fluid Reasoning (intercept 0.17; 95% CI: -1.43 to 1.76; \( p = 0.82 \)), Inhibition (intercept 8.61; 95% CI: -8.29 to 25.52; \( p = 0.27 \)) and Updating (intercept 0.86; 95% CI: 12.73 to 14.45; \( p = 0.89 \)). This indicates no publication bias in these domains.

**Discussion**

The current paper builds on previous reviews by focusing on executive function within five, theoretically-driven subcomponents. Three questions were posed: 1) which specific executive functions are affected by the presence of untreated OSA?; 2) if executive functions are impaired, does treatment help to remediate these deficits?; and 3) are any of these effects moderated by publication status, sample source, study design, age, disease severity, treatment length or control screening?
The findings of the present review

The results from the present analysis indicate that executive function is impaired in OSA compared to control participants across all five subcomponents. People with OSA have difficulty Shifting between tasks or mental sets, Updating and monitoring working memory representations, Inhibiting dominant or pre-potent responses, they struggle with Generating new information without external input or efficiently accessing long term memory, and they have significant problems with Fluid Reasoning or problem solving. Further to this, the present research demonstrated that if participants undertake CPAP treatment, executive function difficulties across these five sub-domains are reduced.

This meta-analysis was unable to assess the impact of CPAP compliance on improvement in executive function. Articles assessed in the present review excluded individuals who were not compliant with treatment, or reported only the group mean number of hours CPAP was used, except in one instance where participants were divided into compliant or non-compliant, hence we cannot make any concluding statement on improvements in executive function in individuals who do not follow their treatment regime. However, exploration of the impact of months of CPAP use revealed no additional gain with extended use (6 months or more).

A recent review (Bucks et al., 2012) summarised the current understanding of cognitive function across a number of domains including executive function. The authors found that in 2 reviews (Aloia et al., 2004; Saunamäki & Jehkonen, 2007) and 2 meta-analyses (Beebe et al., 2003; Fulda & Schulz, 2003) meeting inclusion criteria, executive function was impaired by comparison with controls and norms. The present results provide further support that executive function is impaired in people with OSA. Furthermore this review (Bucks et al., 2012) examined
improvement in executive function following CPAP treatment in 1 meta-analysis (Aloia et al., 2004) and 1 literature review. (Saunamäki & Jehkonen, 2007) The reviews examined came to opposing conclusions, hence the summary was inconclusive. The present meta-analysis suggests that, overall, CPAP treatment is successful in improving executive function difficulties caused by OSA, and adds to the available evidence supporting the benefits of CPAP treatment.

Past reviews have grouped executive function into one combined domain, or collapsed them by test, making it impossible to delineate the subcomponents of executive function that are impaired. The present paper views executive dysfunction within current neuropsychological understanding of executive function. Such a framework provides a possible explanation for relationship or work difficulties seen in OSA. Individuals with OSA may experience relationship (Engleman & Douglas, 2004; Kales et al., 1985) or work productivity difficulties (Hillman et al., 2006b; Mulgrew et al., 2007) as they may not be able to inhibit inappropriate responses to aggravating social situations, or solve novel problems in a work place with Fluid Reasoning.

The effect of moderators

The present study was not able to examine the impact of disease severity on OSA across the full range of AHI, as there were insufficient numbers of mild (AHI 5-14) and moderate samples (AHI 15-29). However, comparison of severe (AHI 30-50) and very severe (AHI 50+) OSA samples revealed no impact of severity on the deficits found, and no impact on executive consequences of CPAP treatment. The literature is divided with regard to whether there is a relationship between disease severity as measured by AHI and cognitive dysfunction. (Aloia et al., 2004; Carlson, Neelon, Carlson, Hartman, & Bliwise, 2011; Tsai, 2010) In a recent meta-analysis of
episodic memory function, Wallace and Bucks (Wallace & Bucks, 2012) found no relationship with disease severity and in a systematic review Aloia et al. (Aloia et al., 2004) found no relationship between disease severity and executive function however they did find a positive relationship between disease severity and global cognitive function and attention/vigilance. As yet, the link between OSA disease severity and cognition is unclear. This is most likely due to the complex picture of co-morbidity, and as yet, no definitive way of measuring disease onset in OSA. (Valencia-Flores, Bliwise, Guilleminault, Cilveti, & Clerk, 1996)

Previous studies have demonstrated that age and OSA results in a double burden, with older individuals exhibiting poorer cognition. (Antonelli-Incalzi et al., 2004; Ayalon, Ancoli-Israel, & Drummond, 2010) The present meta-analysis did not find this same relationship as age did not significantly moderate the results. In studies comparing older and younger participants with OSA, older adults have been found to be more impaired on tests of executive functioning. (Antonelli-Incalzi et al., 2004; Ayalon et al., 2010) The lack of effect in the current meta-analytic review may be due to assessing the effect of age using group averages, which resulted in similar age distributions across samples. Primary comparison studies which explore the interaction of age and OSA on these executive function subcomponents will be important for clarifying this relationship.

Furthermore, selection of controls with PSG or questionnaires did not significantly moderate the findings. This result may seem surprising given that estimates of undiagnosed OSA are high. (Al-Lawati et al., 2009; Butkov & Lee-Chiong, 2007; Young et al., 1997) However, this finding is consistent with that recently reported by Wallace and Bucks. (Wallace & Bucks, 2012) Whilst it is still the case that some control participants in primary studies may have undiagnosed
OSA, including studies which have used questionnaire screening procedures does not appear to confer a risk of failing to find an effect in meta-analyses in OSA.

**Heterogeneity among results**

Many of the domains demonstrated a high level of heterogeneity in the test results, indicating that there may be other factors in each domain influencing the observed mean. However, further analysis demonstrated this heterogeneity did not obscure real differences in each subdomain of EF. The domain that exhibited the most heterogeneity was the domain with the largest number of different tests, Fluid Reasoning, in which there were 13 different tests. Neurocognitive tests, especially executive function tests, even purportedly measuring the same domain or sub-domain, will also capture facets of other cognitive or motor skills. For example Trails B, a measure of Shifting, also requires attention and taps into psycho-motor speed. (Lezak et al., 2004) Furthermore, few tests currently used to assess executive function were originally designed for the specific purpose of measuring executive function. (Banich, 2009; Suchy, 2009)

Future research exploring the executive dysfunction of OSA may benefit from selecting measures more closely targeted at executive functions and designed to fractionate performance into theoretically-driven and dissociable subcomponents. One such measure is the Random Number Generation (RNG) task. (Maes, Eling, Reelick, & Kessels, 2011) This task takes only a few minutes, is easily administered with a laptop computer and provides measures of Shifting and Inhibition. (Maes et al., 2011; Miyake et al., 2000)

One other factor that might lead to heterogeneity is pre-morbid IQ or intelligence. This is because greater IQ and/or education appears to provide ‘protection’ against cognitive decline because of greater cognitive reserve.
(Alchanatis et al., 2005; Stern, 2002) Unfortunately, not all studies provided a measure of academic achievement or pre-morbid intelligence (IQ). Given that age decreases reserve and IQ increases it, primary studies that stratify the sample into age and IQ groups when examining the impact of OSA on executive function are needed.

**Limitations**

An issue that cannot be addressed by the current review is whether the deficits evidenced in the literature are primary or secondary. That is to say, whether the deficits found in OSA in executive function are due to neurological damage (primary effect) or to impairments in attention which themselves are the result of sleep fragmentation, sleep deprivation and the associated excessive daytime sleepiness. (Verstraeten & Cluydts, 2004; Verstraeten et al., 2004). Given evidence of frontal activation in participants completing these tasks (Zhang et al., 2011) and the presence of structural abnormalities in the frontal lobes of individuals with OSA (Zimmerman & Aloia, 2006), it seems likely that these executive function deficits are a primary and direct consequence of OSA (see Beebe and Gozal (Beebe & Gozal, 2002) for a review; but see Durmer and Dinges (Durmer & Dinges, 2005b) for evidence that similar frontal changes are also seen post sleep loss). Indeed, one study, (Verstraeten et al., 2004) which controlled for attention deficits while exploring executive dysfunction differences between OSA and controls, concluded that most of the executive difficulties were secondary to attention problems. This study demonstrated that the one area with deficits remaining after controlling for attention, was Shifting. More studies of this nature and the development of tasks that can tease apart the contributions of attention and executive cognitive processes to task performance are required.
Conclusions

People with OSA have difficulty with the executive facets of Shifting, Updating, Inhibiting, Generativity, and with Fluid Reasoning. Further, the present research indicates that all these difficulties improve with CPAP treatment. Age and disease severity did not moderate the effects found, however, further studies are needed exploring the extent of primary and secondary effects, and the impact of age and pre-morbid ability (cognitive reserve).
CHAPTER 3:
RECRUITMENT, PARTICIPANTS AND MATERIALS

This chapter presents a detailed description of recruitment processes, participant details, and the materials used in Chapters 4 and 5 of this thesis, given strict word limits of the journals in which the papers were published.

The present thesis collected two samples; one sample of individuals with diagnosed sleep disorders from the West Australian Sleep Disorders Research Institute, Sir Charles Gairdner Hospital (the clinical sample), and another from the Western Australian Participant Pool, University of Western Australia (the community sample). The studies presented in this thesis received approval from the Human Research Ethics Committees of Sir Charles Gairdner Hospital and the University of Western Australia. Data are used in Chapters 4 and 5 of this thesis.

Chapter 4 uses the questionnaire data from the clinical and community samples (356 individuals from the community group, and 679 individuals from the clinical group) to examine the factor structure of the Epworth Sleepiness Scale (the ESS).

Chapter 5 uses the sleep study data from the clinical and community samples to examine the relationships between cognition, sleep fragmentation and hypoxia. This chapter use the sample of participants who underwent polysomnography; 134 individuals from the clinical sample and 16 individuals from the community group.

Further details of these samples, recruitment processes and materials are provided below.

Recruitment

Participants

Participants were 18 years of age or over, spoke English as a first language, did not have any sleep disorders other than OSA, and gave written, informed consent.
Community Sample (Control sample). Between January 2011 and September 2011, 356 participants were recruited from the wider community through community and University notices, community talks, radio announcements, fundraising events, the University of Western Australian psychology student pool, and the Western Australian Participant Pool (a panel of community volunteers aged 50, Director RS Bucks). Participants were aged 44±22 years (range 18 to 100 years), BMI was 25.3±6.3 (range 15.5 to 62.1) and 248 (70%) were female. Data from these participants were later used in Chapter 4 to examine the factor structure of the ESS.

From this wider sample, participants were selected through an algorithm developed by Marshall, Dawson, and Bucks (2009) which identifies people at low risk of sleep disordered breathing based on sleep disordered symptoms (e.g., snoring and/or high BMI). A brief description of this algorithm is detailed below (See, Marshall et al., 2009). At the time the questionnaire was administered, participants were also asked to indicate if they would agree to participate in an overnight sleep study. Fifty participants who were low risk also indicated they would like to participate in an overnight study. Of these 50, 24 attended an overnight sleep study at the University of Western Australia, Centre for Sleep Research. The other 27 individuals declined when phoned, could not be contacted with the details provided, or did not arrive on the night of their arranged study. The final sample was \( N = 16 \), as used in Chapter 5 to examine the relationship between cognition and nocturnal disturbance (Table 1). AHI in these individuals ranged from 0 to 16 breathing events per hour. Typically, individuals assessed in a sleep clinic are ‘flagged’ for treatment if they score an AHI of 15 or more. Thus, the community recruitment process allowed us to recruit to the lower end of the OSA severity spectrum (from none to
very mild OSA). Should hypoxia and sleep fragmentation affect cognition on a continuum from no harm to great harm, these community participants provide the data for the lower end of the spectrum.

**Clinical Sample (OSA sample).** Between March 2009 and July 2011, 679 participants were recruited consecutively through the West Australian Sleep Disorders Research Institute, at Sir Charles Gairdner Hospital, as they came into the sleep clinic for overnight assessment of a sleep disorder. Participants were aged 50.3±22.3 years (range 17 to 82 years), BMI was 33.0±7.4 kg.m² (range 16.2 to 61.4), AHI was 34.8±29.7 events/hr (range 0 to 193) and 305 (44%) were female. Data from these participants were later used in Chapter 4 to examine the factor structure of the ESS. Demographic details are presented in Table 3.
Table 3.

Demographic data for the clinical and community participants.

<table>
<thead>
<tr>
<th></th>
<th>Community sample</th>
<th></th>
<th>Clinical sample</th>
<th></th>
<th>Whole sample</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$N = 16$</td>
<td>$N = 134$</td>
<td></td>
<td>$N = 150$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>40.3</td>
<td>4.7</td>
<td>9</td>
<td>7</td>
<td>53.9</td>
<td>12.1</td>
</tr>
<tr>
<td>BMI</td>
<td>24.4</td>
<td>4.6</td>
<td>16.5</td>
<td>34.3</td>
<td>34.1</td>
<td>7.5</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>44% male</td>
<td>7 males</td>
</tr>
<tr>
<td>Education</td>
<td>12.7</td>
<td>2.9</td>
<td>10.0</td>
<td>20.0</td>
<td>11.6</td>
<td>3.0</td>
</tr>
<tr>
<td>Premorbid IQ</td>
<td>112.9</td>
<td>5.5</td>
<td>100.0</td>
<td>121.6</td>
<td>110.4</td>
<td>8.0</td>
</tr>
<tr>
<td>AHI</td>
<td>2.2</td>
<td>3.9</td>
<td>0</td>
<td>2.8</td>
<td>41.3</td>
<td>26.1</td>
</tr>
</tbody>
</table>

Note: Premorbid IQ as estimated using the NART; AHI = Apnoea Hypopnoea Index; BMI = Body Mass Index.
Participants were individuals identified on overnight polysomnography (PSG) as having clinically significant OSA, and who chose to complete the questionnaires outlined below. Questionnaires were placed by the participants’ bedside by sleep technicians, and participants were invited to complete them, if they so chose. Those who completed questionnaires were then invited to complete neuropsychological evaluation at the University of Western Australia.

Participants were scheduled for neurocognitive assessment after receiving a diagnosis of OSA, but before beginning CPAP treatment. These participants also completed the set of neurocognitive assessments detailed below.

Of 296 participants approached for cognitive testing, 126 declined to participate, 32 did not attend their appointment or cancelled, 2 discontinued testing during the appointment, and 2 had begun CPAP treatment prior to their appointment. Thus, the final sample of clinical (OSA) participants who completed pre-treatment assessment was $N = 134$: for demographic details, see Table 3. Data from these participants were used in Chapter 5 to examine the relationship between cognition and nocturnal disturbance.

Materials

Questionnaires

**Demographic questionnaire.**

**Body Mass Index (BMI).** Participant BMI (Weight kilograms / Height in metres$^2$) was calculated based on medical records from the hospital, or taken when the participant visited UWA for their overnight sleep study. If the participant was only involved in the questionnaire portion of the study self-report data was used.

**Epworth Sleepiness Scale (ESS)** (Johns, 1991). The ESS is an eight-item, self-administered, subjective scale in which respondents rate their perceived
likelihood of ‘dozing’ in a variety of everyday situations, e.g. ‘list an item here’.
Participants respond from 0 - ‘would never doze’ to 3 - ‘high chance of dozing’, with
an ESS score >10 representing excessive sleepiness (Johns & Hocking, 1997).
Strong internal consistency has been reported (Cronbach’s alpha 0.74 – 0.88; (Izci et
al., 2008; Johns, 1994; Smith et al., 2008), as well as good test-retest reliability ≥ .81(Izci et al., 2008). Use of the ESS is widespread in both clinical and research
settings (Bloch, Schoch, Zhang, & Russi, 1999; Chen et al., 2002; Izci et al., 2008;
Shahid, Shen, & Shapiro, 2010).

**Depression, Anxiety, and Stress Scale-short form (DASS-21)**(Lovibond &
Lovibond, 1995). The DASS-21 is a 21 item, self-administered questionnaire
designed to measure depressive, anxiety, and stress symptoms. The 21 item scale
was developed from the highest loading items on all 3 negative affect subscales
(Lovibond & Lovibond, 1995). Respondents indicate the degree to which they felt
negative affect symptoms over the past week, from 0 – ‘did not apply to me at all’ to
3 – ‘applied to me very much, or most of the time’. Scores are doubled to be
consistent with the original 42-item DASS and range from 0 to 126, with higher
scores indicating greater depression, anxiety and stress, respectively. The scale has
strong internal consistency on all three subscales (Depression 0.82 – 0.84; Anxiety
0.74 – 0.80; Stress 0.84 – 0.88; (Norton, 2007), good test-retest reliability .90 – .99
(Norton, 2007), and a reliable factor structure in clinical and non clinical samples
(Antony, Bieling, Cox, Enns, & Swinson, 1998; Henry & Crawford, 2005; Szabó,
2010).

**Insomnia severity scale (ISI)** (Bastien, Vallieres, & Morin, 1999). The ISI is a
7-item scale evaluating an individual’s satisfaction with their sleep patterns, daily
functioning due to current sleep habits and level of distress caused by their sleep.
The items are rated from 0 – ‘not at all’ to 4 – ‘extremely’. Scores range from 0 through 28, with higher scores indicating greater dissatisfaction with current sleep patterns. The scale has good internal consistency as measured by Cronbach’s alpha, 0.76 - 0.78 (Bastien et al., 1999), and adequate concurrent validity with other assessments of non-restorative sleep, .65 (sleep diaries, wake after sleep onset) (Morin & Azrin, 1985).

**Pittsburgh sleep quality index (PSQI)** (Buysse, Reynolds, & Monk, 1989). The PSQI is a 19-item, self-rated questionnaire assessing subjective sleep quality over the preceding month. These 19 questions are subdivided into 7, equally weighted subscales, each assessing a different facet of sleep quality: Sleep duration, Sleep disturbance, Sleep latency, Days of dysfunction due to sleepiness, Sleep efficiency, Overall sleep quality, Need of medication to sleep. These subscale scores are summed for a global PSQI score between 0 and 21, with higher scores indicating poorer sleep quality. The scale has good internal consistency as measured by Cronbach’s alpha, 0.72 - 0.83 (Buysse et al., 1989; Suleiman & Yates, 2011), and excellent test-retest reliability, 0.85 (Buysse et al., 1989).

**Berlin Sleep Questionnaire (Berlin)** (Netzer, Strohs, Netzer, Clark, & Strohl, 1999). The Berlin is a 10-item, self-rated questionnaire assessing an individual’s risk of sleep-disordered breathing. The questions load onto 3 categories (snoring, fatigue, obesity and hypertension). An individual is considered at high risk of Sleep Disordered Breathing (SDB) if they score positively on 2 or more categories. SDB is a meta-category for all nocturnal breathing disorders including OSA. The scale has good internal consistency as measured by Cronbach’s alpha, 0.86 - 0.92 (Netzer et al., 1999).
Functional Outcomes of Sleep Questionnaire (FOSQ-10) (Weaver et al., 1997). The FOSQ-10 is a 10-item, self-report scale evaluating functional outcomes as a result of sleep quality, developed from the highest loading items for each of the 5 subscales from the original 30-item questionnaire. Items are rated from 1 – ‘no’ to 4 – ‘yes, extreme’ and for questions 3, 4, 5 & 7, an option for 0 – ‘Don’t do this activity for other reasons’. These items are divided into 5 subscales and a total score. The mean of the subscale scores is multiplied by 5 to give a score of between 5 and 20, with higher scores indicating poorer functional outcomes. The scale has good internal consistency, 0.86 – 0.91 (Weaver et al., 1997), and excellent test-retest reliability, 0.81 – 0.90 (Weaver et al., 1997).

Questionnaire Data Processing.

Algorithm. The Algorithm selected for use in this study has been used to identify high and low risk of sleep disordered breathing in a truck driving population (Marshall et al., 2009). The algorithm uses the scores from the Berlin, ESS, PSQI and FOSQ-10 questionnaires outlined above. This combination of questionnaires has greater sensitivity (78.1) and specificity (93.5) than any of the individual questionnaires alone (Marshall et al., 2009). For further detail about scoring this algorithm, permission must be sought from the creators (Marshall et al., 2009).

The present study modified the algorithm to exclude participants if they were at high risk of restless legs syndrome (answered yes on the PSQI to ‘Have you had trouble sleeping due to your legs twitching or jerking while you sleep?’), or those with high levels of depression on the DASS (scores of 20 or more), or those with high scores on the ISI (scores of 10 or more). This was done to exclude participants who were at high risk of comorbid sleep disorders (restless legs syndrome and insomnia), and individuals with high levels of depression symptoms. These disorders
carry their own cognitive dysfunction profile that may have interfered with the investigation of the cognitive profile of OSA (Burt, Zembar, & Niederehe, 1995; Fulda & Schulz, 2001; Pearson et al., 2006). **Neurocognitive Assessment**

Participants were assessed by a graduate psychology student trained to administer the tasks. Assessments were scored twice, the first time immediately by the assessor, and a second time by another trained assessor. In the event there were any scoring discrepancies, these issues were discussed with a third person (RSB). Assessors followed a scripted protocol that outlined when and how tasks were to be administered, and what instructions were to be given. Timed tasks were assessed using a stopwatch. Participants completed the following neurocognitive assessments:

**Cognitive Drug Research System (CDR)** (Corporation, 2009; Wesnes, 2000). The CDR is a 30 minute, computerized battery of cognitive assessments measuring attention, short-term and episodic long-term memory. The tasks were administered on a 15-inch screen laptop computer. Word tasks were all presented in white, capitalized print on a dark blue background. All tasks have alternate versions, which are presented in random order. Alternate versions of the tasks were presented to OSA participants who returned for follow up testing (after 3 months of CPAP treatment) to control for practice effects. Participants respond verbally or using a two-button response box, marked YES and NO.

Participants were instructed to answer as quickly and accurately as possible during each of the tasks. These tests have good reliability and validity (van den Goor et al., 2008; UBC, 2010). The CDR has previously been used to assess subtle cognitive changes in OSA (Hirshkowitz et al., 2007; Roth, Rippon, & Arora, 2008; Roth et al., 2006) and other chronic illnesses, such as depression and mild cognitive impairment (Newhouse et al., 2012; Vasudev et al., 2012), with good reliability and
validity (Van Den Goor et al., 2008), and has been described in detail elsewhere (Wesnes, 2000).

**Assessment of Attention.**

**CDR Digit vigilance.** A single digit number, the target number, between zero and nine (inclusive) was presented to the right of the screen. A series of single digit numbers was also presented in the centre of the screen, and scrolled though randomly at a rate of 150 numbers per minute. The target number was presented a total of 45 times at random intervals. This task yielded a measure of the % of targets detected accurately (min = 0, and max = 100).

**CDR Choice reaction time.** The word ‘YES’ or ‘NO’ was randomly presented on screen, for a total of 30 presentations, with an inter-stimulus interval ranging from one to three and a half seconds. The participant used the response box to respond using the ‘yes’ button every time they saw ‘YES’ or the ‘no’ button if they saw ‘NO’ as quickly as possible. This task yielded a measure of the % of targets detected accurately (min = 0, and max = 100).

**Assessment of Long Term Memory.**

**CDR Picture presentation and delayed recognition.** Twenty full colour, high resolution images were presented on screen, one at a time for one second each with a three second space between images. Participants were instructed to remember the image as later they would be asked to recall the images. Twenty minutes later, the original twenty pictures were presented one at a time, in random order, amongst twenty distracter pictures. The participant was instructed to respond using the response box by pressing ‘yes’ if they had seen the picture before or ‘no’ if the picture was new. This measure gave ‘% greater than chance’ accuracy (% original
targets correctly identified + % novel targets correctly identified – 100; min = 0, and max = 100).

**CDR Delayed word learning.** Fifteen words are presented for one second in the centre of the monitor, with a two second space between words. The participant is instructed to remember as many of these 15 words as they are able, as they will be asked to recall them at a later stage. Both immediately post learning, and after a delay, the participant was asked to recall as many of the words from the original list of 15 words as possible. This task gave a measure of the amount of new information the participant had recalled (T1_IRCL - T1_DRCL; min = -15, and a max = 15).

**CDR Delayed word recognition.** The 15 words from the original word presentation task were presented one at a time, in random order, amongst fifteen distracter words. The participant was instructed to respond using the response box by pressing ‘yes’ if they had seen the word before or ‘no’ if the word was new, for each of the 30 words as quickly as possible. This task yielded a measure of the ‘% greater than chance’ performance (% original targets correctly identified + % novel targets correctly identified – 100; min = 0, and a max = 100).

**Assessment of Short Term Memory.**

**Short term numeric memory.** The participant was instructed to remember five single digits between zero to nine (inclusive) presented in series in the centre of the screen at the rate of one digit per second. The instructor checked that the participant recalled the numbers correctly by asking them to recite the numbers aloud. A second presentation of the numbers was administered in the event the participant could not recite the numbers correctly. Following this, a series of single digit numbers was presented, to which the participant responded ‘yes’ if it was one of the original they were remembering and ‘no’ if it was not one of the original five. This task yielded a
measure of ‘% greater than chance performance’ (% original targets correctly identified + % novel targets correctly identified – 100; (min = 0, and a max = 100).

**CDR Spatial working memory.** The participant was shown a laminated card with a pictorial representation of a house with nine windows and instructed that the same image would be presented on screen with three windows lit. A different configuration of ‘lit’ and ‘dark’ windows was presented a total of 36 times. The participant was instructed to remember the placement of the lit windows for the activity to follow. Then an image of the same house with one lit window was presented. Using the response box the participant was instructed to respond ‘yes’ if the lit window was one of the original 3 lit windows, or ‘no’ if the window was dark in the original house. This task yielded a measure of ‘% greater than chance performance’ score (% original targets correctly identified + % novel targets correctly identified – 100; min = 0, and a max = 100).

**Assessment of Executive Function.**

**Clock Drawing Task 1 & 2 (CLOX 1 & 2)** (Royall, Cordes, & Polk, 1998). CLOX is a brief, clock drawing task that screens for visuospatial/construction and executive function difficulties, specifically the strategic control and planning of behaviour (Strauss et al., 2006). CLOX scores correlate highly with other executive function tasks (Strauss et al., 2006). Participants draw a clock without assistance (CLOX 1) and then copy the examiner’s clock (CLOX 2). An executive control function score (ECF) is calculated by subtracting the CLOX 1 score from the CLOX 2 score. Higher ECF scores indicate better executive function. The scale has excellent internal consistency, 0.82 (Royall et al., 1998), and inter-rater reliability, 0.93 (Royall et al., 1998).
**FAS Controlled Oral Word Association task (COWA)** (Lezak et al., 2004). The COWA is a 3-minute phonemic fluency task (Strauss et al., 2006), that taps into executive function, specifically efficiency of access to long term memory stores (Fisk & Sharp, 2004). The FAS has high internal consistency, 0.83 (Tombaugh, Kozk, & Rees, 1999). Participants generate as many words as possible in 1 minute, beginning with a given letter (F, A and S). The total repetitions, non-words and intrusions are subtracted from the total number of words recalled to give the final score. Higher scores indicate better performance.

**Trail Making Test (TMT; Trails A and B)** (Partington & Leiter, 1949; Reitan, 1958). The Trail Making Test is a five minute assessment in which the participant connects circles numbered from 1 to 25 (Part A), or alternates between numbers and letters in order (1-A-2-B, etc: Part B). The task measures attention, speed, and mental flexibility. Test-retest reliability is good to excellent (Part A, 0.55 - 0.79; Part B, 0.75 - 0.89), and inter-rater reliability is, likewise, excellent 0.89-0.92 (Strauss et al., 2006). Construct validity is good: trails correlates well with other measures of attention, and part B is related to other measures of executive function (e.g. WCST (Kortte, Horner, & Windham, 2002)). Longer times and more errors indicate poorer performance. A ratio score (B/A), indexing the relative increase in time taken to complete Part B as a function of Part A, was calculated (Lamberty, Putnam, Chatel, Bieliauskas, & Adams, 1994) and used as a measure of set shifting/mental flexibility.

**Assessment of Pre-morbid Intelligence.**

**National Adult Reading Test - 2 (NART-2)** (Nelson & Willison, 1991). The NART-2 is a 5-10 minute reading task providing an estimate of pre-morbid IQ. NART scores are highly correlated with IQ in healthy participants (Schretlen, Buffington, Meyer, & Pearlson, 2005) and with premorbid IQ and general cognitive
Neurocognitive Disturbance in OSA

ability (Crawford, Deary, Starr, & Whalley, 2001; Crawford, Stewart, Cochrane, Parker, & Besson, 1989). Participants read aloud 50 irregular English nouns. Error scores were combined with demographic variables to estimate pre-morbid intelligence (Crawford et al., 2001; Crawford et al., 1989). The NART is robust to declines in other cognitive skills, hence it has utility as a measure of premorbid intelligence (Crawford et al., 2001).

\[
\text{Predicted WAIS FSIQ} = 135.96 - 0.789 (\text{NART errors}) - 4.6 (\text{sex}) - 2.15 (\text{social class}) + 0.122 (\text{age})
\]

For social class, 1 = professional, 2 = intermediate, 3 = skilled, 4 = semi-skilled, 5 = unskilled. People who were currently retired, unemployed were coded by their main lifetime occupation. Participants who had never worked were coded as unskilled, traditionally wives were coded by their husbands occupation (Crawford et al., 2001).

Sleep Metrics

Polysomnography (PSG)

PSG is the gold standard method of OSA diagnosis (Kryger, 2010). Measurements of brain activity (sleep state and arousal) (EEG), eye movements (EOG), skeletal muscle activation (EMG), heart rhythm (ECG), air flow (thermistor and nasal pressure), breathing related effort (inductance plethysmography), blood oxygen levels (pulse oximetry), snoring (vibration sensor), and patient position (position monitor) are recorded. These measurements allow calculation of a variety of measures of disease severity including level of sleep fragmentation and oxygen desaturation. Sleep metrics were recorded using the Compumedics Grael HD-PSG Sleep Diagnostic Amplifier System at the Centre for Sleep Research, University of Western Australia and the Compumedics E-Series PSG/EEG at WASDRI, Sir
Charles Gairdner Hospital. All sleep studies were scored using the Compumedics PSG 3 software.

For the clinical participants, the sleep study was performed at the West Australian Sleep Disorders Research Institute (WASDRI) at Sir Charles Gairdner Hospital.

**Sleep study scoring and staging**

All sleep studies, whether recorded at WASDRI or UWA, were scored by the researcher (MO) from the original, overnight polysomnography recordings. All equipment placement, sleep staging and event scoring was completed according to The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications (AASM, 2012). A combination of EEG, EOG and EMG are used for staging sleep, whilst the other sensors described below are used for detecting and grading the severity of sleep disrupting abnormalities/events.

**Equipment Placement.**

Table 4 provides a list of the measurements obtained. These measurements were used for marking key events that may disturb the participants’ sleep and for staging sleep.
<p>| Devices used and measurements obtained during overnight polysomnograph. |
|---------------------------------------------------|-----------------|----------------|----------------|
| Number of electrodes                     | Placement        | Measures          | Purpose                  |
| <strong>Electroencephalogram (EEG)</strong>       | Six exploring    | Scalp in accordance with international 10-20 system | Brain activity | Sleep staging |
|                                    | and three reference electrodes | | | |
| <strong>Electro-oculogram (EOG)</strong>          | Two exploring    | Outer canthus of the left and right eyes | Eye movements | Sleep staging (especially REM sleep) |
|                                    | and one reference electrode(s) | | | |
| <strong>Electromyogram (EMG)</strong>             | Two exploring    | Chin and anterior tibialis of the left and right leg | Muscle tension | Sleep staging (especially REM sleep) or diagnosing muscle disorders |
|                                    | and one reference electrode | | | |
| <strong>Electrocardiogram (ECG)</strong>          | Two exploring    | 5th intercostal space below the left nipple and right clavicle | Heart rate and rhythm | Heart rate and rhythm |
|                                    | electrodes | | | |
| <strong>Pulse Oximetry</strong>                   | One electrode   | A finger | Blood oxygen saturation | Blood oxygen saturation |
|                                    | | | | |
| <strong>Nasal Pressure</strong>                   | One nasal cannula | Inside the nostrils | Nasal pressure | Nasal airflow |
|                                    | | | | |
| <strong>Thermistor</strong>                       | One thermocouple | The base of the nose and over the mouth | Air temperature | Nasal and oral airflow |
|                                    | | | | |</p>
<table>
<thead>
<tr>
<th>Respiratory Inductance</th>
<th>Two exploring electrodes</th>
<th>Two bands – one across the abdomen and one across the thorax</th>
<th>Movement of chest and abdomen walls</th>
<th>Breathing rhythm and effort</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plethysmography</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Microphone</td>
<td>One electrode</td>
<td>Placed over the bed or on the throat</td>
<td>Sound</td>
<td>To capture snoring</td>
</tr>
<tr>
<td>Video</td>
<td>Video camera</td>
<td>Above the bed</td>
<td>Video</td>
<td>Capture gross body movements</td>
</tr>
</tbody>
</table>

**Electroencephalogram electrode placement (EEG).** EEG electrodes were placed according to the international 10-20 system at the recommended placements; Frontal 4 (F4), Central 4 (C4) and Occipital 2 (O2), referenced to Mastoid 1 (M1). Backup electrodes were placed over Frontal 3 (F3), Central 3 (C3) and Occipital 1 (O1), referenced to Mastoid 2 (M2). Figure 4 shows the placement of the EEG. The EEG records brain activity and is used for staging the presence or absence and the depth of sleep.

**Electrooculogram (EOG).** EOG were placed according to the recommended settings at Eye 1 (E1) referenced to M2 and Eye 2 (E2) referenced to M1. These were placed 1 cm out and 1 cm up of the outer canthus of the right eye, and 1 cm out and 1 cm down of the outer canthus for the left eye. The EOG records eye movements.

**Electromyogram (EMG).** Three electrodes were used to record chin EMG; 1 in the midline 1 cm above the inferior ridge of the mandible, and 2 placed 2 cm
below the inferior ridge of the mandible, 1 was placed 2 cm to the right of the
midline and 1 placed 2 cm to the left of the midline. Four electrodes were used for
measuring leg EMG. Two were placed on the body of each anterior tibialis, between
2 and 5 cm apart. Both Chin and Leg EMG record muscle activity in the region they
are applied.

**Electrocardiogram electrode placement (ECG).** ECG electrodes were
placed according to the recommended II lead torso electrode placement. One
electrode was placed on the middle of the left clavicle, the other electrode was
placed at intercostal space IV directly below the nipple. The ECG recorded the rate
and regularity of the person’s heartbeat.

**Snore sensor/microphone.** A snore microphone was placed on the
individual’s neck to the side of the participant’s laryngeal prominence at which the
Grael headbox sat (i.e. if the headbox was to the left of the bed, the sensor was
placed to the left of the laryngeal prominence). This placement was for patient
comfort. The snore microphone recorded the presence or absence of snoring and
other vocalizations.

**Piezoelectric abdomen and thoracic bands.** Two piezoelectric bands were
used for monitoring breathing rate and rhythm. One band was placed around the
abdomen at the navel, and another around the thorax at or below the nipples and
under the arms. The bands are used for detecting effort during breathing. Figure 8
shows the placement of the piezoelectric bands.

**Thermistor.** The thermistor is placed at the base of the nose, with a sensor
over the mouth and in each nostril. The sensor records the rate and rhythm of
breathing through temperature. The thermistor is used for detection of apnoea.
Nasal cannula. The nasal cannula is placed at the base of the nose with a sensor in each nostril. The sensor records the rate and rhythm of breathing through nasal pressure. The nasal cannula is used for detection of hypopnoea.

Oximeter. The oximeter was placed on a finger on the hand closest to the Grael headbox. This placement was for patient comfort. The oximeter samples the percentage of blood oxygen every 3 seconds.

Position sensor. The position sensor was placed around the thoracic piezoelectric band at the centre of the person’s chest. This sensor registered the position of the sleeping participant either as supine, prone or lateral right or left.

Video and audio. Video and audio were recorded with an infrared video camera in the ceiling of each room. This captured the individual’s activity and assisted with the definition of sleep stages and events when signals were ambiguous.

Sleep stage scoring (AASM, 2012)

Sleep was scored in 30-second, sequential epochs from the start of the study to the end of the study. In the event that there were 2 or more stages in 1 epoch, the epoch was assigned the sleep staging comprising the greater portion of the epoch.

Wake (W). Stage wake was scored if there were trains of alpha rhythm (8-13 Hz) over O1 and/or O2 when the participants eyes were closed, that attenuated with eye opening; eyeblinks were present at a frequency of 0.5-2 Hz; reading eye movements were evident, or if irregular eye movements were present with high chin EMG.

Non-REM Stage 1(N1). N1 was scored if the reasonably regular slow eye movements were present for more than 500 msec; if a low amplitude mixed frequency (4-7Hz) EEG activity was present; if vertex sharp waves were present
over C3 or C4, for less than .5sec. Sleep onset was scored as the first epoch with greater than 50% sleep, this stage is usually N1.

**Non-REM Stage 2 (N2).** N2 was scored if K complexes unassociated with arousals, or sleep spindles were evident and then continuously scored until the presence of an arousal, major body movement or the person transitioned to N3 or REM sleep. K-complexes are sharp negative waves immediately followed by a positive component for between 0.5 to 1 second. A sleep spindle is a train of distinct waves between 11-16Hz for longer than .5seconds.

**Non-REM Stage 3 (N3).** N3 was scored when the epoch exhibited greater than 20% slow wave activity. Slow wave activity are waves of 0.5 to 2 Hz with peak to peak amplitude of greater than 75 micro-volts.

**Rapid Eye Movement Sleep (REM).** REM was scored when low amplitude, mixed frequency EEG; low chin EMG; saw tooth waves and/or rapid eye movements were present. Rapid eye movements are irregular sharply peaked movements lasting for less than 500msec. Saw tooth waves are serrated waves of 2 to 6 Hz often preceding a burst of rapid eye movement.

Table 5 provides a summary of the key markers of sleep stages as defined by the AASM manual (AASM, 2012).
Table 5.
The neurological and physiological attributes of the different sleep stages in healthy individuals (AASM, 2012).

<table>
<thead>
<tr>
<th>Stage</th>
<th>Brain activity</th>
<th>Muscle activity</th>
<th>Eye activity</th>
<th>Breathing activity</th>
<th>Heart activity</th>
<th>Other features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wake</td>
<td>Alpha brain waves (8-13 Hz)</td>
<td>High muscle tone</td>
<td>Reading eye movements</td>
<td>Irregular, sighs, yawns</td>
<td>Irregular</td>
<td>Aware, talking and moving</td>
</tr>
<tr>
<td>Stage 1 (N1)</td>
<td>Theta brain waves (4-7 Hz)</td>
<td>Some loss of muscle tone</td>
<td>Slow rolling eye movements</td>
<td>Breathing becomes regular</td>
<td>Heart beat becomes regular</td>
<td>Shallow sleep stage, can jerk awake readily</td>
</tr>
<tr>
<td>Stage 2 (N2)</td>
<td>Theta waves with sleep spindles (11-16 Hz) and K-complexes</td>
<td>Decrease in muscle tone movements</td>
<td>No eye movements</td>
<td>Breathing becomes regulated</td>
<td>Heart beat becomes regulated</td>
<td>Decrease in environment awareness</td>
</tr>
<tr>
<td>Stage 3 (N3)</td>
<td>Delta waves (0.5 – 2 Hz, &gt;75 μV amplitude)</td>
<td>Decrease in muscle tone movements</td>
<td>No eye movements</td>
<td>Breathing becomes very regulated</td>
<td>Heart beat becomes very regulated</td>
<td>Deep or slow wave sleep</td>
</tr>
<tr>
<td>REM</td>
<td>Rapid low voltage EEG</td>
<td>Loss of all muscle tone movements</td>
<td>Rapid eye movements</td>
<td>Breathing becomes regular</td>
<td>Heart rate becomes regular</td>
<td>Dreaming sleep</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>quite irregular</td>
<td>quite irregular</td>
<td></td>
</tr>
</tbody>
</table>
Event scoring (AASM, 2012)

Event scoring was completed in 5-minute epochs from the start of the study to the end of the study.

Respiratory events. An apnoea was scored when there was a drop in the peak thermistor sensor equal to or more than 90% for 10 seconds or more for at least 90% of the event. The event is classified as obstructive if it is associated with increased respiratory effort measured throughout the event. The apnoea is classified as central if there is no inspiratory effort during the event. The event is classified as a mixed apnoea if it demonstrates a lack of inspiratory effort during the first portion of the event and effort is resumed in the last half of the event. A hypopnoea was scored according to the alternative criteria, which state that a drop in nasal pressure equal to or greater than 50% of baseline, lasting equal to or greater than 10 seconds for equal or more than 90% of the event, accompanied with a 3% drop in oxygen saturation from baseline (AASM, 2012).

Arousals. Arousals can be scored through any stage of sleep if there is an abrupt shift in EEG frequency greater than 16Hz for more than 3 seconds, with a minimum of 10 seconds of stable sleep preceding the arousal. During REM sleep, this disturbance in the EEG must be accompanied by an increase in chin EMG.

Periodic Leg Movements (PLM). PLM was scored if the leg movement lasted between .5 to 10 seconds with a greater than 8 micro-volt increase in baseline leg EMG amplitude. There must be 4 such events in a row with between 5 and 90 seconds between each movement for them to be included in the same series.

Bruxism. Bruxism or tooth grinding was scored if the individual demonstrated brief or sustained periods of elevated chin EMG of at least twice the baseline chin
EMG for between .25 to 2 seconds in duration, for a minimum of 3 elevations. Audio was also used to listen for the presence of tooth grinding sounds.

**Measures of Disease Severity**

**Apnoea Hypopnoea Index (AHI).** OSA severity is conventionally scored using the AHI which is the average number of times a person exhibits an apnoea or hypopnoea per hour of sleep (Kryger, 2010). According to published cut-offs, an AHI of 0-4.9 is classified as ‘No OSA’, 5-14.9 as ‘Mild OSA’, 15-29.9 as ‘Moderate OSA’, and 30+ as ‘Severe OSA’. People diagnosed and treated with mild OSA must, additionally, have daytime complications not better explained by another illness (e.g. depression, poor quality of life, high blood pressure) (Kryger, 2010).

**Arousal Index (ArI).** Sleep Fragmentation was operationalised using the Arousal Index (ArI). The ArI is the total number of arousals divided by total sleep time.

**Oxygen Desaturation Index ≤3% of baseline (ODI3).** ODI3 is the total number of times the individual’s blood oxygen saturation dropped 3% below their baseline oxygen saturation level.

**Sleep Efficiency (SE).** SE is the number of minutes of sleep divided by the number of minutes in bed. Good SE is approximately 85 to 90%.

**Mean and Lowest Sp02.** Oxygen saturation is defined as the ratio of oxyhemoglobin to the total concentration of hemoglobin present in the blood (ie Oxyhemoglobin + reduced hemoglobin). When arterial oxyhemoglobin saturation is measured non-invasively by pulse oximetry, it is called SpO2. Mean SpO2 is the average SpO2. Lowest SpO2 is the lowest point SpO2 reached during sleep.
Conclusions

The present chapter detailed the participants, recruitment, and materials used in chapters 4 and 5 of the present thesis. These methods were used in part or full for both samples; the clinical and community samples.
CHAPTER 4: EXAMINING THE UTILITY OF THE ESS

This chapter was published in the journal *Sleep and Breathing*: Olaithè M., Skinner T.C., Clarke J., Eastwood P., Bucks R.S. (2012). Can we get more from the Epworth Sleepiness Scale than just a single score? A confirmatory factor analysis of the ESS. *Sleep and Breathing*; 17(2); 763-9.

Preface

Excessive daytime sleepiness (EDS) is both a common symptom of OSA and impacts on cognitive function. EDS can be measured subjectively with questionnaires such as the Epworth Sleepiness Scale (ESS), or objectively with tasks such as the Multiple Sleep Latency Test (MSLT). There is debate as to which method is best for ascertaining daytime sleepiness. Considered in this debate is that the ESS far outweighs the MSLT for time and monetary cost, and may be more ecologically valid, as it is a simple paper-and-pencil task taking approximately five minutes to complete, and asks about sleepiness in a range of circumstances most people are exposed to daily. Possibly due to ease and low cost of administration and high ecological validity, the ESS is used extensively (Shahid et al., 2010).

The present thesis uses the ESS as a measure of sleepiness in order to understand the impact of sleepiness on cognition. Prior to doing so, the authors evaluated the factor structure of the ESS as a measure of sleepiness in both a community and clinical sample.
Abstract

Purpose: The Epworth Sleepiness Scale (ESS) is a widely used tool for measuring sleepiness. In addition to providing a single measure of sleepiness (a one factor structure), the ESS also has the capacity to provide additional information about specific factors that facilitate sleep-onset, including a person’s posture, activity and environment. These features of sleepiness are referred to as somnificity. This study evaluates and compares the fit of a 1-factor structure (sleepiness) and 3-factor structure (reflecting low, medium and high levels of somnificity) for the ESS.

Methods: All participants (a community sample \(N = 356\) and a clinical sample \(N = 679\)) were administered the ESS. Confirmatory Factor Analysis was used to evaluate and compare the fit of 1-factor and 3-factor models of the ESS.

Results: In both samples, a 3-factor structure (community sample adjusted \(\chi^2 = 2.95, \text{RMSEA} = 0.07, \text{CFI} = 0.95\); clinical sample adjusted \(\chi^2 = 3.98, \text{RMSEA} = 0.07, \text{CFI} = 0.98\)) provided a level of model fit that was at least as good as the 1-factor structure (community sample adjusted \(\chi^2 = 5.01, \text{RMSEA} = 0.11, \text{CFI} = 0.87\); clinical sample adjusted \(\chi^2 = 8.87, \text{RMSEA} = 0.11, \text{CFI} = 0.92\)).

Conclusions: In addition to a single measure of sleepiness, the ESS can provide subscale scores which relate to three underlying levels of somnificity. These findings suggest that the ESS can be used to measure an individual’s overall sleep propensity as well as more specific measures of sleep propensity in low, moderate and high levels of situational somnificity.
Can we get more from the Epworth Sleepiness Scale than just a single score? A confirmatory factor analysis of the ESS

Excessive Daytime Sleepiness (EDS) is a subjective report of a lack of energy during the day and an amplified desire to fall asleep when sedentary, despite a full nights’ sleep (Sauter et al., 2010; Tsai, 2010). EDS is associated with an increased risk of motor vehicle accidents (Lyznicki, Doege, Davis, & Williams, 1998; Webb, 1995), decreased quality of life (Takegami et al., 2011) and significant financial costs within the community (Webb, 1995).

The Epworth Sleepiness Scale (ESS) (Johns, 1991, 2002) is a simple, commonly used measure of sleepiness among a variety of populations in both clinical and research settings (Bloch et al., 1999; Chen et al., 2002; Izci et al., 2008; Shahid et al., 2010). Its widespread use has resulted in the ESS becoming highly influential in shaping our understanding of the concept of sleepiness and its response to treatment (Johns, 2002). Developed by Johns in 1991 (Johns, 1991) the ESS generates a single score from estimates of the level of sleepiness an individual experiences in eight commonly encountered situations, each question reflecting a low, medium or high somniforic (sleep-inducing) situation (Johns, 1994). While generally used as a single score, previous research demonstrates that the ESS has the potential to provide additional information about an individual’s sleepiness by analysing it in terms of 3 factors, reflecting low, medium and high somniforic situations (Johns, 1994).

Confirmatory factor analysis (CFA) is a flexible statistical technique designed for construct validation and scale refinement (MacCallum & Austin, 2010).
In the present study we have applied it to examine the hypothesised relationships between the construct (sleepiness/somnificity), and the items used to measure this construct (the 8 ESS items) (MacCallum & Austin, 2010). CFA is particularly useful when examining hypothesised constructs such as sleepiness as it can account for measurement error that naturally exists between the ‘pure’ construct and the measurement of the construct (Iaccobucci, 2010), providing a stringent test of this latent structure. Because CFA allows the comparison of alternative models, and identifies the model of ‘best fit’ to the data (Byrne, 2010; Floyd & Widaman, 1995), it is an ideal technique for assessing the fit of different models of the ESS.

A number of studies have investigated the factor structure of the ESS (Chen et al., 2002; Heaton & Anderson, 2007; Izci et al., 2008; Johns, 1992, 1994; Kingshott, Engleman, Deary, & Douglas, 1998; Smith et al., 2008), however none has compared the original 1-factor design to a 3-factor structure. Therefore, the aim of the current study was to use CFA to compare a 1-factor to a 3-factor structure (reflecting low, medium and high somnificity) in community and clinical participant samples. Study of these populations is important as the ESS is commonly used in both clinical and community settings.

**Method**

The ESS was completed by a community sample and a clinical sample, either online or in hardcopy, as part of a set of questionnaires investigating their subjective sleep, mood and, memory functioning. The study was approved by the Human Research Ethics Committees of Sir Charles Gairdner Hospital, and the University of
Participants

**Community Sample.** Between January 2011 and September 2011, 356 participants were recruited from the wider community through community and University notices, community talks, radio announcements, fundraising events, the University of Western Australian and the West Australian Participant Pool (a panel of community volunteers aged 50+). Participants were 44.4±22.3 years (range 17.4 to 100.9 years), BMI was 25.3±6.3 (range 15.5 to 62.1) and 248 (70%) were female. Of these participants 28 reported a diagnosed sleep condition, 11 of which reported OSA (only 3 reported this to be treated), 3 reported Restless Legs Syndrome and 1 reported Sleep Paralysis.

**Clinical Sample.** Between March 2009 and July 2011, 679 participants were recruited through the West Australian Sleep Disorders Institute as they came into the sleep clinic for overnight assessment of a sleep disorder. Participants were 50.3±22.3 years (range 17.0 to 82.5 years), BMI was 33.03±7.42 kg.m² (range 16.2 to 61.4), AHI was 34.8±29.7 events/hr (range 0 to 193) and 305 (44%) were female. Of these participants 665 later received a diagnosis of OSA, none of whom were treated at the time of completing this questionnaire.

Questionnaire

The ESS (Johns, 1991) is an eight-item, self-report scale in which respondents rate their perceived likelihood of dozing in a variety of everyday
Neurocognitive Disturbance in OSA situations. The scale response options range from 0 – “would never doze” to 3 – “high chance of dozing”, with an ESS score >10 representing excessive sleepiness (Johns & Hocking, 1997). The measure has strong internal consistency (Cronbach’s alpha 0.74 – 0.88) (Johns, 1992, 1994; Smith et al., 2008), indicating that the items of the ESS measure a similar construct. The ESS also exhibits good test-retest reliability ≥0.8218, indicating that without treatment an individual score will remain stable when retested.

**Statistical analyses**

Data screening and descriptive analysis were performed using SPSS 19.0 for Windows (IBM, 2012b). The factor structure of the ESS was evaluated using AMOS 18.0 for Windows Version 2.0 (IBM, 2012a).

**Data screening.** Data were screened for missing values. Little’s MCAR test demonstrated that missing data were missing completely at random in both the community, $X^2(21) = 27.02, p = .17$, and the clinical sample, $X^2(80) = 67.70, p = .84$. The proportion of missing values was less than 5% in both samples, thus, Expectation Maximisation in SPSS was used to replace missing values, before conducting factor analysis in AMOS.

**Confirmatory Factor Analysis.** Confirmatory Factor Analysis was used to compare the 1-factor (Figure 2) and 3-factor (Figure 3) ESS structures. The models were estimated using Maximum Likelihood (ML) estimation. The fit of the models was compared through examination of Chi square difference tests and incremental model fit indices (Normed Fit index, NFI; Incremental Fit Index, IFI; Comparative Fit Index, CFI; Goodness-of-fit index, GFI; Adjusted Goodness of Fit Index, AGFI;
Standardized Root Mean Squared Residual, SRMR; and Root Mean Square Error of Approximation, RMSEA). Particular emphasis was placed on X2 adjusted for degrees of freedom (X2 / df), the RMSEA and associated confidence intervals, the SRMR and the CFI, as the power and robustness of these particular indices has previously been demonstrated (Hu & Bentler, 1999; Iaccobucci, 2010; MacCallum & Austin, 2010). As fit indices cannot be compared across models (Iaccobucci, 2010), model cross-validation indices (Akaike’s Inclusion Criterion, AIC; Consistent Akaike’s Inclusion Criterion, CAIC; and Expected Cross Validation Index, ECVI) were calculated in order to compare models, as determined by published cut-offs (See Table 7) (Hu & Bentler, 1999).
Figure 2.

Proposed 1-factor structure of the ESS
Figure 3.

Proposed 3-factor structure of the ESS
Results

Mean scale scores for ESS total score and each of the 3 factors (e.g. low somnificity score = average of Item 6 + Item 8 scores) for each sample are shown in Table 6. As expected, higher scores were obtained from the clinical sample for total ESS, low and moderate somnificity scores ($p < .001$) but not for high somnificity scores, which were similar for both samples. These results indicate that the items most endorsed were the high and moderately sleepy items, and that of these items the clinical group endorsed these more than the community group. However, there was no significant difference between how often the low somnificity items were endorsed in the clinical and community groups. This indicates that these items are endorsed as often in the community group as they are in the clinical group. The relationships between the scores for each ESS item from the clinical and community samples are presented as a correlation matrix in Appendix 2.

Table 6.

ESS total score, subscale scores for each of the 3 somnificity constructs and t-test results between the community and clinical samples.

<table>
<thead>
<tr>
<th>Scale total</th>
<th>Community sample M±SD</th>
<th>Clinical sample M±SD</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>N = 356</td>
<td>N = 693</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ESS Total score (max. 23)</td>
<td>7.2±4.2</td>
<td>10.0±53</td>
<td>0.001</td>
</tr>
<tr>
<td>Low somnificity score (max. 3)</td>
<td>0.1±0.3</td>
<td>0.3±0.6</td>
<td>0.001</td>
</tr>
<tr>
<td>Medium somnificity score (max. 3)</td>
<td>0.8±0.7</td>
<td>1.2±0.9</td>
<td>0.001</td>
</tr>
<tr>
<td>High somnificity score (max. 3)</td>
<td>1.5±0.7</td>
<td>1.9±0.8</td>
<td>0.724</td>
</tr>
</tbody>
</table>
Comparisons of the 1 and 3 factor solutions were conducted using adjusted 
$\chi^2$, incremental fit indices and model comparison statistics. The Chi square statistic 
for the community sample was significant for the 1-factor, $\chi^2(20) = 100.06, p < 
.001$, and 3-factor, $\chi^2(17) = 50.23, p < .001$ models. Likewise, the Chi square 
statistic for the clinical sample was significant for the 1-factor, $\chi^2(20) = 177.30, p < 
.001$, and 3-factor, $\chi^2(17) = 67.61, p < .001$, models. However, chi square is highly 
affected by sample size and, as such, becomes unstable with moderate to large 
samples (Byrne, 2010; Hu & Bentler, 1999; Iaccobucci, 2010; MacCallum & Austin, 
2010). Due to the sensitivity of $\chi^2$ to sample size it has been argued that a model 
demonstrates reasonable fit if $\chi^2$ is $\leq 3$ when adjusted by degrees of freedom ($\chi^2 
/df$) (Hu & Bentler, 1999; Iaccobucci, 2010). For both the community and clinical 
samples the 3-factor model adjusted $\chi^2$ was close to or less than 3 (2.95 and 3.98 
respectively) whilst the 1-factor model exceeded 3 in both the community and 
clinical samples (5.01 and 8.87 respectively). Accordingly, and as recommended, we 
report incremental fit statistics (Table 7) (Byrne, 2010; Hu & Bentler, 1999;
Iaccobucci, 2010).
Table 7.

Fit indices and model statistics for the 1- and 3 factor models in the community and clinical samples

<table>
<thead>
<tr>
<th>Fit Index</th>
<th>Community sample (n=356)</th>
<th>Clinical sample (n=679)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normed fit index (NFI)</td>
<td>0.85</td>
<td>0.92</td>
</tr>
<tr>
<td>Incremental fit index (IFI)</td>
<td>0.88</td>
<td>0.95</td>
</tr>
<tr>
<td>Comparative fit index (CFI)</td>
<td>0.87</td>
<td>0.95</td>
</tr>
<tr>
<td>Goodness-of-fit index (GFI)</td>
<td>0.93</td>
<td>0.97</td>
</tr>
<tr>
<td>Adjusted goodness-of-fit index (AGFI)</td>
<td>0.88</td>
<td>0.93</td>
</tr>
<tr>
<td>Standardized root mean squared residual (SRMR)</td>
<td>0.06</td>
<td>0.04</td>
</tr>
<tr>
<td>Root mean square error of approximation (RMSEA)</td>
<td>0.11 (0.09 to 0.13)</td>
<td>0.11 (0.09 to 0.12)</td>
</tr>
<tr>
<td>Akaike’s Inclusion Criterion (AIC)</td>
<td>132.06</td>
<td>88.20</td>
</tr>
<tr>
<td>Consistent Akaike’s Inclusion Criterion (CAIC)</td>
<td>210.06</td>
<td>180.86</td>
</tr>
<tr>
<td>Expected Cross Validation Index</td>
<td>0.37 (0.29 to 0.47)</td>
<td>0.30 (0.25 to 0.37)</td>
</tr>
<tr>
<td>(ECVI)</td>
<td>0.15 (0.12 to 0.19)</td>
<td>0.15 (0.12 to 0.19)</td>
</tr>
</tbody>
</table>

Note. 1 Ideal value ≥ .95; 2 Ideal value ≤ .08; 3 Ideal value ≤ .06; 4 lowest value is the best model (Hu & Bentler, 1999).
Results presented in Table 7 demonstrate that the 3-factor model accounts for at least as much variance as the 1-factor model. Firstly, the magnitude of all fit indices presented in Table 7 was greater for the 1-factor model than the 3-factor model in both the clinical and population samples. Secondly, for the majority of fit indices the 3-factor model met or exceeded recommended cut off values in both samples, whereas the 1-factor model did not. Finally, the validation indices (AIC, CIAC, ECVI) indicated that the 3-factor model had the lowest values in all instances indicating this to be the best fit to the data.

**Discussion**

This study shows that the ESS can provide measurements of situational somnificity at three different levels, in addition to its commonly used single measure of sleepiness. The term somnificity reflects the facets of sleepiness encompassing the broad features of a person’s posture, activity and environment that facilitate sleep-onset for a majority of people, the majority of the time (Johns, 2002). In the context of the ESS, a highly somniforic situation is one that induces sleep easily (e.g. lying down in the afternoon), whilst a low somniforic situation does not normally do so (e.g. waiting at traffic lights) (Johns, 2002).

An example of the clinical relevance of such a finding is seen in those individuals who do not achieve a clinically sleepy total ESS score but concede to falling asleep under conditions that do not normally induce sleepiness. In the clinical sample, of those who obtained an ESS total score of <10, 5.9% conceded that there was a slight chance they would doze “whilst sitting and talking to someone” (Item 6) and 4.4% said there was a slight chance of dozing “whilst in a car stopped in traffic”
(Item 8). Based on the total ESS score alone, these participants would not normally be highlighted as having a problem with sleepiness. However, should the construct scores be available, a high score on the low somnificity scale could raise concern about the individual’s sleepiness, as they are falling asleep under unusual and potentially hazardous circumstances.

In addition to its clinical applications, the concept of somnificity is important for developing an accurate model of sleep processes. Understanding whether or not the feeling of sleepiness, or even sleep, can be achieved by a particular person in a particular situation depends upon the characteristics of both the person and the environment. Understanding the influence of the environment upon our wake and sleep drives is critical in the diagnosis and treatment of sleep disorders. For example, cognitive-behavioural therapy for insomnia acknowledges the potentially important role of environmental sleep cues and incorporates environmental changes into sleep hygiene practices (Perlis, Jungquist, Smith, & Posner, 2005).

The capacity of the ESS to provide potentially useful information on multiple levels of somnificity is due to its design, as it contains a number of items that exist on a continuum of most somniforic to least somniforic (Johns, 1994). Using normalized factor loadings for four different samples, Johns (Johns, 1994) demonstrated that items 1, 2 & 5 were highly sleep inducing, 3, 4 & 7 were moderately sleep inducing and 6 & 8 the least somniforic. The findings of the present study provide convergent evidence that the ESS can be conceptualised as measuring three levels of somnificity.
Previous research has examined the factor structure of the ESS. Many have found, as did we in both samples, that the fit of the one-factor model was not ideal (Heaton & Anderson, 2007; Smith et al., 2008). Whilst psychometric evaluation of a scale aids development and refinement of the measurement of theorized constructs (Floyd & Widaman, 1995), previous evaluations of the ESS have taken, in our view, problematic approaches to dealing with model miss-fit. Either they have proposed the abolition of items from an already very small scale (Smith et al., 2008), utilized niche populations (Heaton & Anderson, 2007), or utilised less than optimal statistical techniques (Heaton & Anderson, 2007).

For example, Smith et al (Smith et al., 2008) examined the validity of a single-factor ESS solution and found that a re-specified model, omitting Items 6 and 8 provided a superior fit to the data. Although deletion of problematic items is a commonly accepted technique for attaining construct unity (Floyd & Widaman, 1995), this can lead to a loss of information. In particular, the loss of Items 6 (sitting and talking to someone) & 8 (whilst stopped at traffic lights) would lead to a loss of information about situations that a majority of people do not find sleep inducing. Identifying those individuals who do find these situations sleep inducing may be crucial, however, for recognizing those people for whom sleepiness is particularly hazardous or problematic. Results from the present paper demonstrate that all items contribute to an individual’s sleepiness score and, furthermore, provide information on the type of situations an individual finds most sleep inducing.

A similar analysis conducted by Heaton and Anderson (Heaton & Anderson, 2007) examined workplace risks for long-haul truck drivers. Instead of abolishing Items 6 & 8, these authors demonstrated that a 2-factor structure with items 6 & 8
falling onto a second factor offered a superior fit to the 1-factor solution. To extract factors, Heaton and Anderson (2007) utilized principle components analysis (PCA) and eigenvalues greater than 1. Heaton and Anderson (2007) defined factor 1 as “representative of activities that may be solitary and encompass little interaction or variation in interaction with environment or other people” (pp. 185) and factor 2 as “activities that may involve more cognitive stimulation and interaction with environment or other people” (pp. 185). However, these factors did not cleanly separate into 2 distinct factors and there were problematic cross loadings for Items 3 (sitting in a public place) and 7 (sitting quietly after lunch). Heaton and Anderson (2007) grouped these items with Factor 1, basing their decision on logic rather than factor loadings. The present findings may explain some of the cross-loadings in the 2-factor design of the ESS proposed by Heaton and Anderson (2007).

With respect to statistical techniques, many previous factor analytic studies of the ESS cite the Kaiser criterion of eigenvalues greater than 1.0 as the primary reason to extract a single factor ESS solution (Chen et al., 2002; Johns, 1994). Unfortunately, this method frequently results in over and under factoring; resulting in the extraction of too many or too few constructs (Green, Lissitz, & Mulaik, 1977). Previous studies have also cited the use of Cronbach’s alpha in support of the homogeneity of the scale (Chen et al., 2002; Johns, 1992). However, high internal consistency does not necessarily equal high scale homogeneity.

Of the factor analytic studies published, all have demonstrated some low factor loadings and, to date, no studies have attempted to contrast the traditional 1-factor model of sleepiness with a 3-factor model of somnificity using confirmatory factor analysis. The present research expands upon our previous understanding of the
ESS by utilizing optimal statistical techniques, utilising the ESS scale in its entirety, and studying the scale in two large samples.

There are two common pitfalls of using CFA, namely confirmation bias and the generalizability of results (MacCallum & Austin, 2010). Confirmation bias is an unfair liking for the model the researcher believes accounts for the most variance in the data, leading a researcher to evaluate only the fit of the preferred model (MacCallum & Austin, 2010). Utilizing a comparative model strategy by evaluating the fit of the traditional model (1-factor) to the hypothesised model (3-factors), as used in the present paper, lends some protection against confirmation bias (Iaccobucci, 2010; MacCallum & Austin, 2010). Additionally, often the results produced in CFA are only generalizable to the population tested (Iaccobucci, 2010). In order to address this issue the present research assessed the fit of the proposed models in two distinct samples, making the results more widely generalizable (Iaccobucci, 2010).

**Conclusions**

The results of this study support a multidimensional conceptualisation of sleepiness. Specifically, they indicate that the ESS can be used both as a measure of overall sleep propensity and as a combination of more specific measures assessing an individual’s sleep propensity under given situations. This second conceptualisation of the ESS scale may be useful for more precisely identifying sources of excessive sleepiness in non-clinical samples. The immediate concerns of an individual experiencing disordered sleep are likely to relate to the day-to-day consequences of their condition. As such, it is extremely important to understand the various factors that contribute to daytime sleepiness, both in terms of ensuring the accurate
assessment and diagnosis of sleep conditions and addressing the patient’s primary concerns.
CHAPTER 5: RELATIONSHIP BETWEEN OSA SEVERITY AND COGNITION

This chapter was submitted to Sleep and Breathing, presently, the authors are replying to reviewer comments: Olaitehe M, Skinner T., Hillman D., Eastwood P., Bucks R. The relationship between cognition and sleep apnea in OSA: a call to arms. Sleep and Breathing (Submitted and under review, December, 2013).

Preface

Studies 1 and 2 (presented in Chapters 2 and 4) examined the pattern of deficits in executive function, and the factor structure of the Epworth Sleepiness Scale (ESS), respectively. Chapter 2 revealed that individuals with OSA experience widespread executive dysfunction that improves with treatment. Furthermore, published literature indicates that attention and memory are likewise impaired in individuals with OSA. The findings from Chapter 4 revealed that sleepiness, as measured by the ESS, was best captured as a 3-factor model measuring low, medium and high situational somnificity.

This chapter uses this information to examine the relationship between the domains of attention, memory and executive function and nocturnal disturbance (hypoxia and sleep fragmentation), whilst controlling for the inter-individual factors of sleepiness, age and premorbid IQ.
Abstract

Introduction: Obstructive Sleep Apnea (OSA) is a common disorder that is associated with impaired attention, memory and executive function. However, the mechanisms underlying such dysfunction are unclear. To determine the influence of sleep fragmentation and hypoxia, this study examined the effect of sleep fragmentation and hypoxia on cognition in OSA, while controlling for potentially confounding variables including sleepiness, age and premorbid intelligence.

Method: Participants with and without OSA (N = 150) were recruited from the general community and a tertiary hospital sleep clinic. All underwent comprehensive, laboratory-based polysomnography (PSG) and completed assessments of cognition including attention, short-term & long-term memory and executive function. Structural Equation Modelling (SEM) was used to construct a theoretically-driven model to examine the relationships between hypoxia and sleep fragmentation, and cognitive function.

Results: Although after controlling for IQ, increased sleep disturbance was a significant predictor of decreased attention (p=0.04) and decreased executive function (p=0.05), controlling for age removes these significant relationships. No significant predictors of memory function were found.

Conclusions: The mechanisms underlying the effects of OSA on cognition remain to be defined. Implications are discussed in light of these findings.
Cognition and nocturnal disturbance in OSA: the importance of accounting for age and premorbid intelligence

Obstructive Sleep Apnea (OSA) is a common sleep disorder (T. Young et al., 2002) characterized by repeated upper airway collapse resulting in intermittent hypoxia and arousals from sleep (Kryger, 2010). Previous studies, comparing individuals with and without OSA, have shown that OSA is associated with impaired cognitive function, particularly in the domains of attention, memory, and executive function. In the attention domain, vigilance appears to be most affected (Beebe et al., 2003; Bucks et al., 2012; Fulda & Schulz, 2003), while in the memory domain, impairment is in delayed visual and verbal memory (Aloia et al., 2004; Bucks et al., 2012; Wallace & Bucks, 2012), and in visuospatial/constructional memory (Bucks et al., 2012). Executive function impairments have been demonstrated in Shifting, Updating, Inhibition, Generativity, and Fluid reasoning (Beebe et al., 2003; Bucks et al., 2012; Fulda & Schulz, 2003; Olaithé & Bucks, 2013; Saunamäki & Jehkonen, 2007). Several domains of cognition appear unaffected by OSA, including language abilities (Beebe et al., 2003; Bucks et al., 2012), immediate visual and verbal memory (Wallace & Bucks, 2012), visuospatial learning (Wallace & Bucks, 2012), and psychomotor functions (Bucks et al., 2012).

The mechanisms underlying cognitive dysfunction in OSA remain undefined. In 2002, Beebe and Gozal (Beebe & Gozal, 2002) proposed a conceptual framework based around critical roles for sleep fragmentation and nocturnal hypoxia in the development of cognitive dysfunction in individuals with OSA. In their model, sleep is viewed as a necessary restorative process, regulating processes including reinforcing foundations for learning and memory (Poe et al., 2010) and modulating...
neuroendocrine demands (Everson, 1995). Disruption of these processes due to sleep fragmentation leads to an inability of the body to return to a balanced state, impairing neural function. The model also postulates that blood gas abnormalities (i.e., hypoxia) due to periodic obstructed breathing compounds this damage. A proposal of this model is that hypoxia and sleep fragmentation contribute equally to impaired cognition in OSA, however it remains unclear whether or not this is correct.

Substantial evidence exists for a negative impact on cognition by sleep fragmentation and/or sleep deprivation (Alhola & Polo-Kantola, 2007; Durmer & Dinges, 2005b; Stepanski, 2002; Verstraeten et al., 2004). Indeed, some authors argue that sleep fragmentation is the key mechanism underlying impaired cognition in OSA (Verstraeten & Cluydts, 2004; Verstraeten et al., 2004). Of note, is the finding that the nature of cognitive deficits in sleep deprived individuals is similar to those individuals with OSA (Alhola & Polo-Kantola, 2007; Durmer & Dinges, 2005b; Verstraeten et al., 2004) Particular impairment in attention and executive function are prominent amongst them (Alhola & Polo-Kantola, 2007; Durmer & Dinges, 2005b; Verstraeten et al., 2004). It has also been noted that deficits in attention in individuals with OSA are associated with indices of sleep fragmentation (Ayalon, Ancoli-Israel, Aka, McKenna, & Drummond, 2009).

Other research suggests that intermittent hypoxia experienced throughout the night by individuals with OSA is primarily responsible for irreparable neural damage and lasting cognitive impairments (Beebe & Gozal, 2002; Beebe et al., 2003). Such permanent damage might explain the finding that, even with effective treatment (viz. abolition of hypoxia and sleep fragmentation), individuals with OSA retain some
Neurocognitive Disturbance in OSA

decrements in memory and executive functioning (Bucks et al., 2012; Olaithe & Bucks, 2013; Wallace & Bucks, 2012). Thus, while there is considerable evidence that both sleep fragmentation (Bartlett et al., 2004; Chiang, 2006; Gale & Hopkins, 2004; Hopkins, Kesner, & Goldstein, 1995; Naismith, Winter, Gotsopoulous, Hickie, & Cistulli, 2011) and hypoxia (Aloia et al., 2004; Bucks et al., 2012; Saunamäki et al., 2010) lead to cognitive dysfunction in OSA, their relative importance remains unclear.

In addition to hypoxia and sleep fragmentation, other inter-individual factors including age (Naismith et al., 2004), premorbid intelligence and sleepiness should be considered in any assessment of cognitive function in individuals with OSA. This is because those individuals with OSA who have high intelligence may exhibit intact performance on neuropsychological assessments, as they can recruit additional cognitive resources or utilise cognitive strategies that aid in performance, acting as a ‘buffer’ for brain insult and injury (i.e., cognitive reserve) (Alchanatis et al., 2005). Similarly, individual differences in age and subjective sleepiness in OSA could modify performance on neurocognitive tests because age independently affects cognition (Alchanatis, Zias, Deligiorgis, Chroneou, et al., 2008), and excessive sleepiness lessens the ability to direct cognitive resources to attend to the task at hand (Durmer & Dinges, 2005a, 2005b; Guilleminault & Brooks, 2001; Naismith et al., 2004).

Despite the wealth of published evidence, most studies of the cognitive burden of OSA have compared OSA to non-OSA participants (Bucks et al., 2012). By contrast, few studies have explored the relationship between the severity of OSA (e.g. as quantified by the Apnoea Hypopnea Index, AHI) and the degree of deficits
experienced, and none has explored the relative contribution of the two proposed mechanisms (that is sleep fragmentation AND hypoxia) within the same sample, whilst controlling for premorbid intelligence, age and daytime sleepiness (Naismith et al., 2004). Thus, this study is the first comprehensively to test the relationship between severity of OSA and cognitive deficits. To do this we utilized Structural Equation Modelling (SEM), a powerful analytic technique that allows testing and estimation of causal relationships between multiple potential mechanisms of cognitive dysfunction, as well as exploration of the relative importance of inter-individual factors of age, premorbid IQ, and sleepiness.

**Methods**

**Participants and Protocol**

In total, 150 adults (18 years+) with and without OSA participated: 134 were patients diagnosed with OSA attending the Sleep Clinic of the West Australian Sleep Disorders Research Institute, Queen Elizabeth II Medical Centre (WASDRI-QEII), Western Australia, for diagnostic polysomnography between March 2009 and July 2011. Cognitive assessments were performed with these individuals after diagnosis but before starting treatment for OSA. To capture the full range of severity of OSA (from no disease to severe disease) an additional 16 community volunteers were recruited. These individuals were members of the West Australian Participant Pool of the University of Western Australia (WAPP-UWA), who completed a set of questionnaires on sleep health between October 2011 and October 2012, and who undertook cognitive assessment prior to a sleep study at the Centre for Sleep Science, at the University of Western Australia. The study was approved by the
Human Research Ethics Committees of Sir Charles Gairdner Hospital, and the University of Western Australia. All participants provided written, informed consent.

**Measurements**

*Polysomnography (PSG)* required participants to attend a full, overnight laboratory-based sleep study. Electrodes were attached according to the American Academy of Sleep Medicine (AASM) recommendations: electroencephalogram (EEG) electrode, electrooculogram (EOG) channels, electromyogram (EMG) (chin and legs), oxygen saturation via pulse oximetry (SpO$_2$), electrocardiogram (ECG), thoracic and abdominal respiratory effort via belts, oral/nasal thermistor, nasal prongs and position sensor. Studies were analysed and apnea hypopnea index (AHI) and arousal index (ArI) determined according to standard guidelines (AASM, 2007).

*Attention and Memory* were assessed using the Cognitive Drug Research System (CDR; United BioSource Corporation), a 30-minute computerized battery of cognitive assessments (Van Den Goor et al., 2008; Wesnes, 2000). All CDR tasks were administered on a 15-inch screen laptop computer, and included assessment of Attention (digit vigilance, simple and choice reaction time), Long Term Memory (delayed picture recognition, delayed word learning, delayed word recognition) and Short Term Memory (digit recall, visuo-spatial recall). Responses were provided either verbally or via a YES/NO response box.

*Executive Function* was assessed using the Clock Drawing Task (CLOX 1 & 2) (Royall et al., 1998), the Controlled Oral Word Association task (COWA), and the Trail Making Test (TMT; Trails A and B) (AASM, 1999; Partington & Leiter, 1949; Reitan, 1958).
Pre-morbid Intelligence was assessed using the National Adult Reading Test - 2 (NART-2), a widely accepted method of estimating premorbid IQ from current capacity, by testing reading capacity of irregular nouns (Nelson & Willison, 1991). The number of errors is used to estimate IQ using a regression equation, with fewer errors being associated with higher premorbid IQ. Irregular word reading ability has been shown both to correlate highly with full scale IQ (Nelson & Willison, 1991) and to be resistant to decline due to dementia (McGurn et al., 2004).

Subjective Sleepiness was assessed using the Epworth Sleepiness Scale (ESS) (Johns, 1991) including its separation into three factors, of low, moderate, and high situational somnificity (Olaithe, Skinner, Clarke, Eastwood, & Bucks, 2012).

All cognitive assessments were administered by graduate psychologists at the University of Western Australia, and all sleep study measurements were obtained and scored by trained PSG Technicians.

Analyses

Data screening and descriptive analysis were performed using SPSS 20.0 for Windows (IBM, 2012b).

Characterising hypoxia and sleep fragmentation. The degree of hypoxia was operationalised in two ways: (i) mean overnight SpO$_2$; and (ii) minimum overnight SpO$_2$\(^1\). The degree of sleep fragmentation was also represented in two ways: (i) sleep

\(^1\) Other indices of sleep fragmentation (e.g., TST, AHI during REM) and the following measures of hypoxia: CT90, Lowest SaO2 during REM, Mean SaO2 during REM) were considered and tested, none accounted for sufficient variance in cognitive performance to be included.
efficiency, the ratio of time spent asleep (total sleep time) to the amount of time spent in bed; and (ii) proportion of arousals not associated with a 3% oxygen desaturation (Total ODI3) (Sulit, Siorfer-Isser, Kirchner, & Redline, 2006), the latter being the number of times an individual’s SpO₂ decreased to below 3% of baseline². The purpose of this index was to model arousals not associated with a significant drop in oxygen, as significant respiratory events generally occur with an arousal making these two factors highly correlated.

**Data screening.** Little’s MCAR revealed that data were missing completely at random, \( X^2(485) = 10.19, p = 1.00 \). The proportion of missing values was less than 10%, thus Expectation Maximisation was used to replace missing values (Cohen & Cohen, 1983). Data were assessed for normality. Many variables violated normality assumptions, hence Generalised Least Squares (GLS) was used to calculate estimates of model fit (Lei & Lomax, 2005). Furthermore, incremental fit indices are reported, as these have been shown to be less sensitive to non-normality (Lei & Lomax, 2005).

**Confirmatory Factor Analysis.** CFA is a confirmatory form of SEM that can be used to confirm a theorised model. SEM is an extension of multiple regression designed to test a set of hypothesised relationships between variables, that are estimated simultaneously (Ullman & Bentler, 2012). It provides a mechanism through which to examine relationships between hypothesized constructs, whilst controlling for individual differences, such as premorbid intelligence and sleepiness, and accounting for measurement error. SEM is particularly useful when examining
hypothesized constructs such as memory, attention, and executive function, as it can account for measurement error that naturally exists between the ‘pure’ construct and its measurement providing a stringent test of the latent structure (Byrne, 2010; Iaccobucci, 2010; MacCallum & Austin, 2010). The weights presented in the figures represent standardized regression weights, and Appendix 3 presents a correlation matrix of all variables included in the final model.

CFA was used to evaluate the theoretical models (Figure 4 and 5). The fit of the models was compared through examination of the Chi square difference test, and incremental model fit indices. ‘Good fit’ indicates that the hypothesised model generates a sample covariance matrix that is similar to the observed covariance matrix inherent in the data set; that is the model is a good explanation of the variance in the observed data (Schreiber, Stage, King, Nora, & Barlow, 2006). Good fit is indicated by a non-significant Chi square value, indicating that the observed and hypothesised models are similar, and fit indices are within recommended guidelines (Table 8 foot note). Particular emphasis is placed on the values of the Comparative Fit Index, CFI; Standardized Root Mean Squared Residual, SRMR; and Root Mean Square Error of Approximation, RMSEA, as the power and robustness of these particular indices has previously been demonstrated (Byrne, 2010; Hu & Bentler, 1999; Iaccobucci, 2010; MacCallum & Austin, 2010). The factor structure of the models was evaluated using AMOS 18.0 for Windows Version 2.0 (IBM, 2012a). The correlation matrix is provided in Appendix 3.

Four models were analysed. The first model examined the interrelationships between hypoxia/sleep fragmentation and cognitive function, controlling for premorbid IQ. The second model repeated this analysis, but controlled for premorbid...
IQ and age. A third model examined the interrelationships between AHI and cognitive function by replacing the measures of hypoxia and sleep fragmentation with AHI. This was performed because AHI is the primary measure of OSA disease severity used in current clinical practice (Kribbs et al., 1993; Tsai, 2010) and the utility of this single measure to predict cognitive dysfunction is disputed (Kribbs et al., 1993; Tsai, 2010). The fourth model assessed the role of attention in mediating cognitive deficits in the domains of short and long-term memory and executive function. This model was examined because cognitive deficits OSA have been proposed to be due to attention difficulties as a consequence of sleep fragmentation (Verstraeten et al., 2004).

Results

Descriptive data and preliminary analyses

Participants were aged 52.4±13.1 years (range 19 to 82 years), BMI was 33.0±7.8 kg/m² (range 15.5 to 57.5), AHI was 37.1±27.5 events/hr (range 0 to 154), had received 11.9±3.1 years (range 6 to 23) of education, and 69 (46%) were female. Descriptive statistics for cognitive and OSA indices are provided in Table 8.
Table 8.

Descriptive statistics for sleep and cognitive measures for the whole sample.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>ArI, /hr</td>
<td>38.3</td>
<td>21.5</td>
<td>0.0</td>
<td>114.5</td>
</tr>
<tr>
<td>Total ODI3</td>
<td>132.6</td>
<td>135.4</td>
<td>0.0</td>
<td>718.0</td>
</tr>
<tr>
<td>Arousals proportion ODI3</td>
<td>4.2</td>
<td>4.6</td>
<td>0.0</td>
<td>28.6</td>
</tr>
<tr>
<td>Sleep Efficiency, %</td>
<td>75.4</td>
<td>12.1</td>
<td>38.6</td>
<td>95.8</td>
</tr>
<tr>
<td>Mean SaO2, %</td>
<td>92.8</td>
<td>3.3</td>
<td>73.0</td>
<td>98.0</td>
</tr>
<tr>
<td>Lowest SaO2, %</td>
<td>83.6</td>
<td>8.7</td>
<td>29.0</td>
<td>96.0</td>
</tr>
<tr>
<td>ESS Total Score</td>
<td>9.5</td>
<td>4.9</td>
<td>0.0</td>
<td>21.4</td>
</tr>
<tr>
<td>Somnificity Factor Low (ESS)</td>
<td>0.0</td>
<td>1.0</td>
<td>-0.6</td>
<td>3.6</td>
</tr>
<tr>
<td>Somnificity Factor Moderate (ESS)</td>
<td>0.0</td>
<td>1.0</td>
<td>-1.3</td>
<td>2.4</td>
</tr>
<tr>
<td>Somnificity Factor High (ESS)</td>
<td>0.0</td>
<td>1.0</td>
<td>-2.5</td>
<td>1.5</td>
</tr>
<tr>
<td>CLOX Ratio</td>
<td>1.7</td>
<td>2.1</td>
<td>-4.0</td>
<td>10.0</td>
</tr>
<tr>
<td>Trails Ratio</td>
<td>2.5</td>
<td>0.8</td>
<td>1.1</td>
<td>6.0</td>
</tr>
<tr>
<td>Verbal Fluency</td>
<td>37.0</td>
<td>12.5</td>
<td>3.0</td>
<td>68.0</td>
</tr>
<tr>
<td>Immediate Spatial Recall</td>
<td>93.5</td>
<td>11.8</td>
<td>12.5</td>
<td>100.0</td>
</tr>
<tr>
<td>CDR Word Learning</td>
<td>0.9</td>
<td>1.9</td>
<td>-7.0</td>
<td>7.0</td>
</tr>
<tr>
<td>Delayed Word Recognition</td>
<td>0.6</td>
<td>0.2</td>
<td>-0.7</td>
<td>0.9</td>
</tr>
<tr>
<td>Delayed Picture Recognition</td>
<td>0.7</td>
<td>0.2</td>
<td>0.2</td>
<td>1.1</td>
</tr>
<tr>
<td>Choice Reaction Time</td>
<td>96.9</td>
<td>2.6</td>
<td>84.0</td>
<td>100.0</td>
</tr>
<tr>
<td>Digit Vigilance Score</td>
<td>96.5</td>
<td>5.1</td>
<td>75.6</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Note: Estimated Premorbid IQ, based on the NART-2, years education, age and occupation (population Mean 100±15); Arousals proportion ODI3 = The number of arousals as a proportion of the number of times an individual desaturated 3% below baseline; Sleep Efficiency = is the ratio of time spent asleep (total sleep time) to the amount of time spent in bed.
Compared to the clinical sample (n = 134), the community sample (n = 16) was younger ($p = 0.001$), and had more years of education ($p = 0.003$). As anticipated they had lower BMI ($p = 0.003$) and lower AHI ($p = 0.001$). There were no differences between the groups in estimated premorbid intelligence ($p = 0.23$). For descriptive statistics on these variables, and for the sample as a whole see Table 9. Age and education were accounted for within the SEM analysis.
Table 9.

Descriptive statistics provided for the clinical and community samples, and for the sample as a whole.

<table>
<thead>
<tr>
<th></th>
<th>Community sample</th>
<th>Clinical sample</th>
<th>Whole sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>40.3</td>
<td>14.7</td>
<td>19</td>
</tr>
<tr>
<td>BMI</td>
<td>24.4</td>
<td>4.6</td>
<td>16.5</td>
</tr>
<tr>
<td>Gender</td>
<td>44% male</td>
<td>7 male</td>
<td>55% male</td>
</tr>
<tr>
<td>Education</td>
<td>12.7</td>
<td>2.9</td>
<td>10.0</td>
</tr>
<tr>
<td>Estimated Premorbid IQ</td>
<td>112.9</td>
<td>5.5</td>
<td>100.0</td>
</tr>
<tr>
<td>AHI</td>
<td>2.2</td>
<td>3.9</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Note. AHI = Apnoea Hyponoea Index; BMI = Body Mass Index; Estimated Premorbid IQ based on Crawfords equation using the NART.
Model I: Hypoxia/sleep fragmentation vs cognitive function

Preliminary assessment of model weights showed that neither sleepiness, nor somnificity factors (Olaithe et al., 2012) were related to any cognitive domain ($p > .05$ for all) (Appendix 3, Table 2), whilst premorbid intelligence was related to the cognitive domains ($p < .05$) (Figure 4). Thus, sleepiness/somnificity was removed from subsequent models whilst premorbid intelligence was retained.
Figure 4.
Model I - The Sleep fragmentation (Sleep Fragment) and Oxygen Desaturation (Hypoxia) Model showing the relationship between nocturnal disturbance and cognitive constructs; Attention, Long Term Memory, Short Term Memory, and Executive Function, controlling for IQ. All weights are presented as standardised regression weights. Note. This model was also explored controlling for daytime sleepiness (using the ESS). ESS scores produced poor fit and were removed.
The final model (Figure 4) was a good account of the relationships between sleep fragmentation (*i.e.*, proportion non-ODI3 arousals and sleep efficiency) and hypoxia (*i.e.*, mean and lowest SpO\textsubscript{2}) and cognitive function (Figure 4). Specifically, good model fit was indicated by a non-significant Chi square statistic, $\chi^2(79) = 91.52, p = 0.16$, and incremental fit indices were within recommended guidelines (Table 10).

The strength of the interrelationships between sleep fragmentation, hypoxia and each cognitive factor (short and long-term memory, attention and executive function) was assessed by examination of the beta weights (Figure 4). Because the sleep fragmentation and hypoxia factors were highly correlated ($r = 0.88$ some paths from these factors to cognition had beta weights > 1.0 (Deegan, 1978; Joreskog, 1999), the model fit was assessed when sleep fragmentation and hypoxia were combined into a single construct, ‘sleep disturbance’ (Appendix 3, Figure 1). This model included measures of ArI as a proportion of total ODI3, Sleep efficiency, Mean and Lowest SpO\textsubscript{2} loading onto a latent construct termed ‘sleep disturbance’, and examined the relationship between these constructs and cognition. Although Chi square model fit was non-significant, the Akaike Information Criterion (AIC) indicated that the one factor model (*i.e.*, sleep disturbance) provided a poorer fit to the data than the original model. Thus, we retained the model separating sleep fragmentation and hypoxia.

The beta weights of Model I revealed significant relationships between sleep fragmentation and both attention ($r = 1.16, p = 0.04$) and executive function ($r = 1.08, p = 0.03$), after controlling for premorbid intelligence (Figure 4), indicating that increased sleep fragmentation is significantly and positively related to increased difficulty with attention and with executive function.
Model II: Controlling for age

The model above was also examined controlling for age (Figure 5). This analysis was separately in order to allow examination of the relationships prior to the removal of age from the model, as age was likely to be strongly associated with disease duration.
Figure 5.

Model II – The Sleep fragmentation (Sleep Fragment) and Oxygen Desaturation (Hypoxia) Model showing the relationship between nocturnal disturbance and cognitive constructs; Attention, Long Term Memory, Short Term Memory, and Executive Function, controlling for IQ, and age. All weights are presented as standardised regression weights. Note. This model was also explored controlling for daytime sleepiness (using the ESS). ESS scores produced poor fit and were removed.
This model was a good account of the relationships between sleep fragmentation 
(i.e., arousals as a proportion of total ODI3 and sleep efficiency) and hypoxia (i.e., 
mean and lowest SpO2) to cognitive function (Figure 5). Specifically, good model fit 
was indicated by a non-significant Chi square statistic, $\chi^2(90) = 103.72$, $p = 0.15$ and 
incremental fit indices were within recommended guidelines (Table 10).
Table 10.

Fit indices for both the AHI and sleep fragmentation/hypoxia models are presented.

<table>
<thead>
<tr>
<th>Fit Index</th>
<th>Sleep fragmentation/hypoxia with Premorbid IQ (Model I)</th>
<th>Sleep fragmentation/hypoxia with Premorbid IQ, and age (Model II)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normed fit index (NFI)¹</td>
<td>0.54</td>
<td>0.54</td>
</tr>
<tr>
<td>Incremental fit index (IFI)¹</td>
<td>0.88</td>
<td>0.90</td>
</tr>
<tr>
<td>Comparative fit index (CFI)¹</td>
<td>0.84</td>
<td>0.87</td>
</tr>
<tr>
<td>Goodness-of-fit index (GFI)¹</td>
<td>0.92</td>
<td>0.91</td>
</tr>
<tr>
<td>Adjusted goodness-of-fit index (AGFI)¹</td>
<td>0.87</td>
<td>0.87</td>
</tr>
<tr>
<td>Standardized root mean squared residual (SRMR)²</td>
<td>0.08</td>
<td>0.08</td>
</tr>
<tr>
<td>Root mean square error of approximation (RMSEA)³</td>
<td>(0 to 0.06)</td>
<td>(0 to 0.06)</td>
</tr>
</tbody>
</table>

Note. ¹ Ideal value ≥ 0.95; ² Ideal value ≤ 0.08; ³ Ideal value ≤ 0.06; ⁴ Lowest value is indicative of best model (Hu & Bentler, 1999).

When age was added, however, there was no longer a significant relationship between either of the sleep disturbance factors and any cognitive factor. That is, the relationship between sleep fragmentation, attention and executive function was no longer found.
Model III: AHI

The model fit was also assessed when hypoxia and sleep fragmentation were replaced by AHI. For this model (Appendix 3, Figure 2) Chi square was statistically significant, $X^2(57) = 86.31, p = 0.007$, and incremental fit indices did not meet required cut points (CFI = 0.39, RMSEA = 0.06, SRMR = 0.08). Such findings indicate that severity of AHI does not predict level of cognitive performance in OSA.

Model IV: Cognitive dysfunction mediated by attention

A model was also developed to test the hypothesis that cognitive deficits in executive function and memory are mediated by attention (Appendix 3, Figure 3). This model would not converge, hence this model provided a poor account of the data and could not be explored further.

Discussion

This study used structural equation modeling to examine the effects on cognition of sleep fragmentation and hypoxia in OSA, while controlling for the potential confounding influences of age, premorbid intelligence, mood and sleepiness.

Before controlling for age, greater sleep fragmentation significantly predicted poorer executive function and attention, whereas hypoxia was unrelated to cognition. After controlling for age, sleep fragmentation joined hypoxia in failing to predict level of cognitive performance in any domain (attention, language, executive function, or episodic memory).
This is not the first study to report no relationship between the severity of cognitive deficits and the severity of OSA (see studies reviewed by Aloia et al. (2004), and presented in Table 3 of their publication). Previous studies, using the AHI, have explained their finding by arguing that AHI is a poor index of OSA severity, at least as it relates to the cognitive burden of the disorder (Tsai, 2010). Other authors have suggested that the lack of a relationship is due to the lack of control of the inter-individual factors of age, premorbid IQ, daytime sleepiness or attention (Alchanatis et al., 2005; Alchanatis, Zias, Deligiorgis, Liappas, et al., 2008; Tsai, 2010; Verstraeten et al., 2004).

This study addressed all of these methodological criticisms and still failed to find a relationship between OSA severity and cognition. What should be made of such findings?

The simplest explanation for the results of the present and other similar studies (Aloia et al., 2004; Wallace & Bucks, 2012), is that there is no dose-response relationship in OSA. That is to say, over a critical threshold there will be neural damage and cognitive dysfunction, but this will not increase with increasing severity of sleep fragmentation and hypoxia.

However, given that cognitive performance is related to time spent without oxygen in other instances of hypoxia (e.g., climbing at altitude (Regard, Oelz, Brugger, & Landis, 1989), carbon monoxide poisoning (Hopkins & Woon, 2006)), and longer time exposed to sleep deprivation results in greater cognitive dysfunction (Durmer & Dinges, 2005a), this account seems unlikely. An alternative hypothesis is
that hypoxia or sleep fragmentation are not captured appropriately, and/or that the correct mechanisms of harm are not being captured.

An individual’s level of hypercapnia is closely linked to their level of hypoxia, and excessive levels of carbon dioxide are found in individuals with OSA (AASM, 2007; Kaw, Hernandez, Walker, Aboussouan, & Mohlesi, 2009). Yet hypercapnia is not routinely explored in OSA-cognition research, despite being included as a potential mechanism of harm in the dominant model of cognitive harm in OSA (Beebe & Gozal, 2002). Further, there is evidence that the degree of hypercapnia correlates with overall cognitive impairment (Findley et al., 1986).

Alternatively, routine indices of sleep fragmentation (ArI) and hypoxia (SpO2) may be insufficiently sensitive to the extent of sleep disturbance, or the length and depth of hypoxia (Asano et al., 2009). An individual with OSA may exhibit many small fragmentations of sleep architecture, or shallow, lengthy oxygen desaturations, not captured by standard indices. Novel measures that capture more detail about sleep fragmentation (Pepin et al., 2005), or describe the length and depth of desaturation (e.g. Integrated Area of Desaturation) (Asano et al., 2009) may demonstrate greater sensitivity to a dose-response relationship in OSA.

The impact of inter-individual differences; Age and IQ

The final model showed that, when premorbid IQ was accounted for, more severe sleep fragmentation, but not hypoxia, was significantly associated with poorer attention and executive dysfunction, but not with poorer memory. However, adding age resulted in these relationships becoming non-significant. A simple explanation for this finding is that age accounts for all of the cognitive deficits seen in OSA. This
Neurocognitive Disturbance in OSA

is unlikely, as case-control studies (with participants matched for age) reveal clear evidence of cognitive deficits in OSA (Bucks et al., 2012). Furthermore, not all individuals experience age-related change in cognition individuals experience age-related cognitive decline, nor are all individuals affected equally. Indeed, many older adults stay cognitively healthy (Hultsch, Hertzog, Small, & Dixon, 1999; Kramar, Erikson, & Colcombe, 2006; Norbury, Craig, Cutter, Whitehead, & Murphy, 2004). Given that there is no good measure of disease duration in OSA, and that older individuals are likely to have experienced OSA for longer, age itself is confounded with disease duration. By controlling for age, it is possible that the variance of interest in an exploration of the mechanisms of cognitive harm in OSA has been discarded. One of the challenges for the field of sleep medicine, therefore, is to develop sound methods of estimating or measuring disease duration.

The choice of measures of sleep fragmentation and hypoxia in the present study was guided by indices that are routinely produced by most sleep software, and that have previously been shown to be related to aspects of cognition (Aloia et al., 2004). Statistical modelling showed that measures of hypoxia and sleep fragmentation exhibited a large degree of co-variation, indicating shared underlying variance, even after calculating an index of arousals independent of 3% oxygen desaturations. The importance of this covariance was investigated by modifying the original model such that sleep fragmentation and hypoxia were considered as a single factor by: (i) loading all measured variables from the latent factors ‘sleep fragmentation’ and ‘hypoxia’ to a single factor (termed sleep disturbance); and (ii) replacing sleep fragmentation and hypoxia with AHI. In both cases, the resultant model fits were not improved relative to the model that utilised separate constructs.
of hypoxia and sleep fragmentation, suggesting that separating out these components of potential harm was the correct approach.

The present study was also careful to consider the potentially important confounding influences of premorbid intelligence and sleepiness in assessing cognitive function in individuals with OSA. The finding that premorbid intelligence was positively related to cognition was expected, and underscores the importance of controlling for this when evaluating cognition in OSA (AASM, 1999; Alchanatis et al., 2005; Ayalon, Ancoli-Israel, Klemfuss, Shalauta, & Drummond, 2006; Castronovo et al., 2009). In contrast, analysis of subjective sleepiness revealed that neither ESS total, nor factors of situational somnificity, were related to cognition. Such findings support studies reporting that subjectively measured sleepiness and somnificity are not related to cognition in OSA (Incalzi et al., 2004). One reason for this might be that individuals with chronic sleep restriction, as occurs in OSA, can become desensitized to the feeling of sleepiness (Durmer & Dinges, 2005b).

The role of attention in mediating cognitive deficits in OSA warranted a separate analysis, given the proposal that sleep fragmentation might cause impaired attention which, of itself, could lead to impaired memory and executive function (Verstraeten, 2007; Verstraeten & Cluydts, 2004; Verstraeten et al., 2004). However, the resultant model fit was poor, suggesting that attention deficits did not explain the effect of sleep fragmentation on attention and executive function.

**Cognitive damage, but no relationship to disease severity**

Despite growing evidence that OSA is associated with cognitive impairment (Bucks et al., 2012), the findings of the present study, and those of meta-analyses
and systematic reviews have found no consistent evidence that the degree of sleep fragmentation or hypoxia is related to the degree of cognitive dysfunction in OSA (Aloia et al., 2004; Wallace & Bucks, 2012). There are several possible reasons for this lack of association.

First, it is possible that the cognitive dysfunction in OSA is due to a mechanism not captured in the models presented in the current study. One such mechanism could be recruitment of additional brain regions to compensate for poor cognitive performance. Consistent with this hypothesis, Ayalon et al. (Ayalon et al., 2006) reported that individuals with OSA recruit additional brain areas, not typically recruited during a verbal learning task and Castronovo et al. (Castronovo et al., 2009) reported increased activation in the left frontal cortex, medial precuneus, and hippocampus in OSA patients whilst performance was at the same level as controls.

Second, individual differences might play a more critical role than previously credited (Durmer & Dinges, 2005b). Such a concept arises from emerging research indicating that the relationship between disease severity and cognitive dysfunction is the product of a multitude of vulnerability and protective factors (Beebe, 2006; Beebe & Gozal, 2002), of which sleep fragmentation, hypoxia and cognitive reserve are only three aspects (Casseli, 2008; Durmer & Dinges, 2005b; Zimmerman & Aloia, 2012). These other facets could include duration of disease (as noted above) (Casseli, 2008), genetic vulnerability (e.g., apolipoprotein e4 genotype) (Casseli, 2008; Zimmerman & Aloia, 2012), the role of the blood brain barrier (Lim & Pack, 2013), and cerebral blood flow (Kiralti, Demir, Volkan-Salanci, Demir, & Sahin, 2010).
Third, the present study examined attention, memory and executive function as whole domains, rather than exploring separate facets of these domains (Olaithe & Bucks, 2013; Wallace & Bucks, 2012). Cognitive domains, such as executive function are not unitary constructs, and should be considered umbrella terms for a range of different cognitive capacities (Fisk & Sharp, 2004; Lezak et al., 2004; Miyake et al., 2000). For example, executive function can be divided into Shifting, Updating, Inhibition, Generativity, and Fluid reasoning (Fisk & Sharp, 2004; Lezak et al., 2004; Miyake et al., 2000). Decomposing each cognitive domain, and then exploring the contribution of a range of risk and protective factors specific to OSA, could enable a more targeted analysis of the impact of OSA on cognition.

Last, as mentioned briefly before, measurement of hypoxia through SpO\textsubscript{2} may be a poor proxy measure of the underlying mechanism by which hypoxia causes harm, namely oxidative stress (Guo, Yan, Qing, Min, & Huan, 2013; Nair, Dayyat, Zhang, Wang, & Gozal, 2011), particularly given that oximetry measured below 80% SpO\textsubscript{2} is not particularly reliable or accurate (Razi & Akbari, 2006). Animal models suggest that oxidative stress and apoptosis-related neural injury might be measured more directly by Brain Derived Neurotrophic Factor (BDNF) (Xie & Yung, 2012), thioredoxin (Nair et al., 2011), or NADPH Oxidase (Guo et al., 2013). It remains unclear whether such measurements would be of value in individuals with OSA.

**Conclusions**

Our findings demonstrate that the relationship of cognitive deficit to sleep fragmentation and hypoxia in OSA is still unclear. That OSA causes cognitive
impairment is not in doubt. However, until disease duration can be measured and specific mechanisms of harm, the nature of protective and risk factors, and their relationship to specific aspects of cognitive ability are established, there is no basis on which to identify who is most at risk of cognitive decline, and how to intervene most effectively.
CHAPTER 6:
GENERAL DISCUSSION

The present chapter was reformatted for submission as an invited paper to *Translational Issues in Psychological Science*; Olaithe, M., Bucks, R.S. & Eastwood, P. *Obstructive Sleep Apnoea (OSA), affecting cognition while we’re in the dark: Implications for practitioners* (Submitted 15.04.2014). This chapter has been formatted for consistency with the rest of the thesis.

Overview of findings

This thesis explored the relationship between attention, memory, executive function, and nocturnal disturbance (hypoxia and sleep fragmentation), in individuals with OSA.

Chapter 2 (Study 1) used meta-analytic techniques to examine executive function in individuals with OSA, before and after treatment. It is the first meta-analysis to divide executive function by its theoretical components and examine the effects of OSA before and after treatment on these components. The results from this chapter show that individuals with OSA exhibit widespread impairments across all facets of executive function, and that these impairments are somewhat ameliorated with CPAP treatment. It remains to be understood whether the executive impairments seen are mediated by inter-individual differences in premorbid IQ, sleepiness, age, and disease severity. Furthermore, it is uncertain if these impairments represent primary or secondary effects of OSA, that is to say whether the impairments seen in any one aspect of cognition are direct consequences of OSA.
or due to impairment in another domain, such as attention, and whether functioning returns to baseline with perfectly adherent treatment.

Chapter 4 (Study 2) used Structural Equation Modelling (SEM) to examine the factor structure of the ESS and to assess the utility of a model with one factor of sleepiness using all items, versus a model with 3 levels of somnificity. In both community and sleep clinic samples, a 3-factor structure provided a level of model fit that was at least as good as the 1-factor structure. These findings suggest that the ESS can be used to measure an individual’s overall sleep propensity as well as more specific measures of sleep propensity in low, moderate and high levels of situational somnificity.

Chapter 5 (Study 3) sought to clarify the relationships between the domains of cognition that are reported to be impaired in OSA (attention, memory and executive function) and the proposed mechanisms that cause such impairment (sleep fragmentation and hypoxia). This study is the first comprehensively to assess the contributions of the major mechanisms currently thought to be responsible for cognitive dysfunction in OSA; namely hypoxia and sleep fragmentation (Beebe & Gozal, 2002; Verstraeten et al., 2004). When controlling for premorbid IQ alone, more severe sleep fragmentation, but not hypoxia, was significantly associated with poorer attention, and executive dysfunction, but not with memory dysfunction. However, when age was accounted for, these effects became non-significant.

Taken together, the studies in this thesis show that OSA has a negative impact on cognition, but that, after accounting for inter-individual differences in premorbid IQ and age there is no relationship between the severity of hypoxia and sleep fragmentation and cognitive impairment.
Implications of these findings will be discussed with reference to the wider literature, future directions, and recommendations for practicing psychology clinicians.

**Discussion of findings**

OSA is associated with activation in (Zhang et al., 2011) and structural abnormality of (Zimmerman & Aloia, 2006) areas of the brain associated with attention, memory and executive function in OSA. OSA also leads to grey matter loss in frontal (associated with executive function) and temporo-parieto-occipital cortices (broadly associated with attention and visuospatial processing), the thalamus (largely involved in relaying information, and regulating sleep-wake cycles) and the hippocampal region (largely associated with memory) (Yahoui et al., 2009). The results from Chapter 2 provide further support that OSA leads to dysfunction in the cognitive facets associated with these areas.

There is additional evidence through examination of chronic diseases with similar mechanisms of harm (Incalzi et al., 2004), animal models (Row, 2007), imaging studies (Thomas, Rosen, Stern, Weiss, & Kwong, 2005) and treatment studies (Ferini-Strambi, Baietto, & Di Gioia, 2003) that sleep fragmentation and hypoxia negatively impact on the brain and on cognitive functioning. Indeed, the dominant model of neural harm and cognitive dysfunction accentuates sleep fragmentation and hypoxia as the main mechanisms of harm in OSA (Beebe & Gozal, 2002).

However, despite evidence that OSA is associated with cognitive impairment, and similar to the findings reported in Chapter 5, other papers have not consistently found a relationship between the degree of nocturnal harm, operationalized by sleep
fragmentation and hypoxia, and daytime cognitive dysfunction. For example, Aloia et al. (2004) undertook a review examining evidence of the relationship between cognitive dysfunction in OSA, sleep fragmentation and hypoxia. This review concluded that, at present, the evidence is equivocal. More recent reviews focusing on memory (Wallace & Bucks, 2012), together with the meta-analysis in this thesis, similarly support the lack of evidence of a relationship between sleep fragmentation, hypoxia and cognition.

It has previously been suggested that no relationship between has been found between these due to small sample sizes, use of less than optimal statistical techniques, and failure to control for inter-individual variables such as cognitive reserve or sleepiness (Tsai, 2010). These points are discussed with reference to the present thesis.

One possibility for a lack of consistent association between cognitive dysfunction, sleep fragmentation and hypoxia is that previous studies have been underpowered. Examining this line of inquiry, the present thesis used large samples. Chapter 2 (Study 1) synthesised 24 years of research, reporting combined samples of 551 healthy controls and 1010 individuals with OSA. Chapter 4 (Study 2) obtained samples of 356 community individuals, and 693 individuals visiting a sleep clinic. Similarly, Chapter 5 reports cognitive performance in one of the largest samples to date of individuals with OSA ($N = 150$). Of the papers reviewed by Beebe et al. (2003), which examined cognition in OSA, the median sample size was 19 (Range = 7 - 199). Furthermore, the samples in all thesis chapters had a wide range of ages, and used male and female individuals, from a wide range of occupational settings.
Therefore, it seems unlikely that a dose-response relationship between hypoxia, sleep fragmentation and cognition was not found due to lack of power.

Another possibility for a lack of association between cognition and severity of OSA is that statistical techniques used in previous studies do not account for measurement error. Many papers have used correlation or traditional regression which do not allow the modelling of measurement error, nor can multiple measures of the same hypothesised construct be used to capture cognitive performance (see Aloia et al., 2004, Table 3, for a summary of papers that found or did not find a relationship between domains of cognition, sleep fragmentation and/or hypoxia). To address this concern, the present thesis used Structural Equation Modelling (SEM), which provides a better fit to the data by accounting for measurement error. The use of SEM makes it unlikely that a dose-response relationship was not found for statistical reasons.

Last, a number of key inter-individual variables including age and premorbid intelligence can impact upon the expression of cognitive impairment. These variables are not routinely accounted for in many studies, and some cannot yet be quantified, e.g., disease duration. This thesis carefully examined a range of inter-individual confounds, including age, premorbid intelligence and sleepiness.

Premorbid intelligence was positively related to cognition, which underscores the importance of controlling for premorbid ability when evaluating cognition in OSA. This is in line with findings by Alchanatis et al. (2005), who reported that people with OSA and high intelligence performed equally well to people without OSA, whilst individuals with OSA and normal intelligence demonstrated cognitive deficits, compared to controls. Cognitive reserve has been shown to be influential in
understanding changes in cognitive evidenced in other chronic disorders, including Alzheimer’s disease (Querbes et al., 2009).

In contrast to premorbid IQ, analysis of subjective sleepiness revealed that neither ESS total, nor factors of situational somnificity (low, moderate, high situational somnificity), were related to cognition. The finding that subjectively measured sleepiness and somnificity are not related to cognition in OSA has been reported elsewhere (Incalzi et al., 2004). One reason for this might be that individuals with chronic sleep deprivation, as occurs in OSA, can become desensitized to the feeling of sleepiness leading to under or over reporting (Durmer & Dinges, 2005b). In order to take into account such factors, future research could use objective measures of sleepiness such as the Maintenance of Wakefulness Test (MWT), and/or the Multiple Sleep Latency Test (MSLT). However, these tests are not routinely used in research due to their high monetary and time costs. Novel markers of sleepiness and neurocognitive impairment, for example the Detrended Fluctuation Analysis (DFA) scaling exponent of the awake electroencephalogram, may be particularly useful in providing relatively quick and objective measures of sleepiness (D’Rozario et al., 2013).

This raises the question of why, despite responding to criticisms levelled at earlier studies regarding sample size, inter-individual differences, or statistical methods, the present analysis did not find a relationship between sleep fragmentation and hypoxia and cognitive dysfunction. The results presented here, and those of past reviews, indicate the very real possibility that there is no relationship to be found. That is, that OSA causes cognitive impairment but that the degree of such impairment is a function of age, disease duration, premorbid cognitive reserve and
reaching some critical threshold of OSA severity above which no additional damage to cognition is accrued. An alternative explanation is that we need to examine other possible mechanisms of harm such as hypercapnia, or develop novel metrics that capture inter-individual differences, and subtle changes in cognition, sleep fragmentation and hypoxia.

These factors are discussed below under the headings of 1. Incorrect or imprecise indices of harm and, 2. Incorrect or imprecise measurement of cognition in OSA and, 3. Additional inter-individual differences.

1. Incorrect or imprecise measurement of nocturnal disturbance

It is possible that sleep fragmentation and hypoxia are not the main mechanisms of harm in OSA. One poorly investigated but destructive influence on cerebral tissue is hypercapnia. Indeed, excessive levels of carbon dioxide are found in OSA (AASM, 2007; Kaw et al., 2009), yet hypercapnia is not routinely measured in overnight sleep studies nor explored in cognitive research in OSA. What limited research there is suggests that the degree of hypercapnia correlates with overall cognitive impairment (Findley et al., 1986).

Alternatively, sleep fragmentation and hypoxia may be the mechanisms of harm, but routine indices may be insufficiently sensitive to the extent of sleep disturbance, or the length and depth of hypoxia (Asano et al., 2009). Routine, clinical sleep fragmentation measures (such as Arousal Index: ArI) capture gross awakenings. Likewise, routine measures of oxygenation (such as the SpO2: peripheral capillary oxygen saturation) report only the depth or time spent below a certain oxygen threshold, or the average number of events per hour. An individual with OSA may exhibit many shallow but long lasting oxygen desaturations, or
multiple short complete apnoeas, or they may experience ‘micro-arousals’, a small fragmentation of sleep architecture, not captured by standard indices. Measures that capture microarousals may provide more intimate detail of sleep fragmentation (Pepin et al., 2005), likewise, new measures may describe the length and depth of desaturation for each individual (e.g. Integrated Area of Desaturation) (Asano et al., 2009).

2. Incorrect or imprecise measurement of cognition

Research examining the cognitive profile of individuals with OSA has focused on examining whole domains of cognition, such as ‘attention’, ‘memory’ and ‘executive function’. Such categorisation may simply be too broad. Neuropsychological tasks typically measure a discrete facet of a whole domain. For example, a task such as Trails B (the participant is instructed to connect a set of 25 letters and numbers as fast as possible, while maintaining accuracy) is typically used in OSA research to examine ‘executive function’, however this task most likely quantifies the capacity of an individual to ‘shift’ their attention between information sets. This Shifting capacity is only a subset of executive function, and does not reflect the whole theoretical domain (Fisk & Sharp, 2004; Miyake et al., 2000).

3. Additional inter-individual differences

Recent research suggests that perception of sleep quality affects daytime functioning and neurocognitive task performance (Draganich & Erdal, 2014). In a novel experiment Draganich and Erdal (2014) demonstrated that individuals labelled as having poor sleep quality, regardless of actual sleep quality, performed less well on a task of attention (Paced Auditory Serial Attention Task, which assesses auditory information processing speed and flexibility by asking the individual to add up a
series of numbers) and executive function (Controlled Oral Word Association Task, which measures verbal fluency by asking individuals to pronounce a set of irregular verbs) than individuals instructed they had slept well. These authors were able to show that this was not due to task demands, and instead was solely attributable to assigned sleep quality. Other researchers have reported that self-perceived sleep quality can affect cognitive performance (Querbes et al., 2009). Similarly, individuals with OSA may perceive their sleep quality to be affected by their new diagnosis of OSA and this may, in turn, affect their performance on neuropsychological assessment.

The studies in this thesis showed that, having accounted for participant age, relationships between sleep fragmentation, attention and executive function were no longer significant. This may be, in part, due to age being confounded with disease duration. At present the best estimate of disease duration in OSA comes from subjective reports of how long the individual has exhibited snoring. However, the individual (and/or partner) may not be aware of their snoring, nor does it conclusively indicate OSA (Cirignotta et al., 2009). Through examination of other chronic diseases where hypoxia is a symptom, longer untreated exposure has been shown to result in greater cumulative damage (Langan, Deary, Hepburn, & Frier, 1991). As such, it is likely that individuals with longer disease duration may exhibit more severe cognitive dysfunction with less responsiveness to CPAP treatment due to longer exposure to hypoxia, sleep fragmentation, or some other mechanism of harm (e.g., hypercapnia). Support for such a notion can be found in studies showing persistent cognitive difficulties following CPAP therapy (Bardwell, Ancoli-Israel, Berry, & Dimsdale, 2001; Bedard, Montplaisir, Malo, Richer, & Rouleau, 1993).
Measurement of disease duration remains a challenge for the field of sleep medicine and presents a substantial barrier to our capacity to interpret studies related to OSA. However, it may be possible to develop a method to ascertain disease duration.

In other chronic diseases such as diabetes, long term disease duration can be measured by the degree of impairment in vascular reactivity of systemic arteries (Clarkson et al., 1996). Given that endothelial and cerebral tissue damage and loss is evident in individuals with OSA (Budhiraja, Parthasarathy, & Quan, 2007; Yahoui et al., 2009), such strategies hold promise for the estimation of disease duration in OSA. Longitudinal studies would be required to catalogue tissue damage in OSA, and develop a way of quantifying disease duration with respect to cerebral tissue damage. Such an examination is already being carried out within the realm of cardiovascular research (Wang, Wu, Feng, & Sun, 2013).

**Implications for clinicians**

There is a need for greater awareness of OSA which should assist with earlier detection of this condition. The negative cognitive and social implications of OSA are substantial (Al-Ghanim et al., 2008; Hillman et al., 2006a), and to some degree irreversible (Bedard et al., 1993). Hence, developing systems for early detection is necessary. Due to the high co-morbidity between depression (Schroder & O'Hara, 2005), cognitive disturbance (Beebe, 2006) and OSA, clinical psychologists and neuropsychologists are likely to be visited by a high proportion of individuals with OSA. Furthermore, psychologists receive intimate information about their clients, and could extend this to asking about sleep quality and habits.

Clinical psychologists and clinical neuropsychologists could potentially be a well-informed point of contact for individuals with sleep disturbances. Sleep
physicians are already a well informed source of information about sleep disorders and treatments, however, due to the overlap of sleep disorders and psychological disorders (Schroder & O'Hara, 2005), not all sufferers of sleep disorders will be referred to a sleep physician. There are presently discrete roles for assessing and addressing sleep disorders and psychological disorders, however due to the overlap in OSA and psychological disorders, greater communication and collaboration between fields is required.

Although not all OSA can be picked up by subjective report, or identifying demographic risk factors, there are a number of features to which clinical psychologists and clinical neuropsychologists should attend. First, the foremost demographic risk factors for OSA are older age, obesity, and being male (Young et al., 1993; T. Young et al., 2002). Second, should an individual report snoring, frequent nocturnal awakenings with a feeling of dread or gasping for air, feeling unrefreshed upon awakening, falling asleep or feeling tired during the day, these individuals are at risk of OSA (Netzer et al., 1999). There are a number of paper-and-pencil tasks that can be utilised to assess these features. Of particular note is the Berlin Questionnaire, which is a 10 minute, self-administered questionnaire that identifies the level of risk that an individual has sleep disordered breathing (Netzer et al., 1999). The clinician can then recommend the individual visit their general practitioner, with the Berlin Questionnaire scores.

At present, the psychologist must refer the client to their GP, who can assess further if a sleep study is warranted; a psychologist cannot directly refer to a sleep physician. However, this is possibly another point where revision of the health care system is required.
Recommending a client has their sleep assessed may assist in the treatment of the psychological disorder for which the person was first referred. Research demonstrates that some individuals with co-morbid OSA and depression experience improvements in depression symptoms with OSA treatment (Schroder & O’Hara, 2005).

In addition, psychologists may have an important role to play to improve CPAP treatment adherence. Treatment with CPAP improves quality of life, reduces car accidents and, to some extent, assists with deficits in cognition (Weaver & Grunstein, 2008). However, between 46-83% of people with OSA are not adherent to the minimum 4 hours a night of recommended CPAP use (Weaver & Grunstein, 2008).

Illness beliefs may play an important role in this non-adherence. Individuals who believe that their OSA symptoms or diagnosis is due to a psychological stressor rather than a physical cause or condition are less likely to adopt a physiological treatment, such as CPAP (Skinner et al., 2013). Psychologists can assess and address the impact such beliefs may have on treatment adherence.

Additionally, individuals with little social support demonstrate lower rates of CPAP adherence (Stepnowsky, Marler, Palau, & Brooks, 2006). Providing social support or assisting clients to rebuild their social support, may assist with treatment compliance (Richards, Bartlett, Wong, Malouff, & Grunstein, 2007).

In order to accomplish greater detection and support for individuals with OSA, psychology clinicians need to be taught about OSA and other sleep disorders. However, the impact of sleep disturbance on psychological and cognitive health are not routinely taught in clinical or neuro psychology courses (Waters & Bucks, 2011). Presently, the Australian Psychological Society (APS) does not list sleep disorders as...
a necessary facet of psychology teaching programs, and most neuropsychological and clinical training text books do not include information on sleep disorders (Waters & Bucks, 2011). Despite this, sleep disorders have notable effects on cognition, psychological health and quality of life (Butkov & Lee-Chiong, 2007). It is time training institutes routinely teach the impact of, and how to screen for, sleep disorders in the clinic.

Conclusions

This thesis examined the impact of OSA upon cognition, demonstrating that OSA impacts upon attention and executive function. As yet, it is unclear which aspects of the nocturnal disturbance of OSA result in the cognitive dysfunction seen in OSA, and the precise picture of inter-individual differences that moderate this harm. It is clear that age and premorbid IQ play a role in moderating cognitive dysfunction in OSA, however future research is required to identify novel metrics of nocturnal disturbance, ‘best practice’ ways to measure cognition, and how to capture disease duration.
References


Antic, N. (September, 2012). [Data acquired from the author through email correspondence].


Bartlett, D., Rae, C., Thompson, C., Byth, K., Joffe, D., Enright, T., & Grunstein, R. (2004). Hippocampal area metabolites relate to severity and cognitive function in obstructive sleep apnea. Sleep Medicine, 5(6), 593-596.


Guo, Q., Yan, W., Qing, L., Min, L., & Huan, W. (2013). Levels of thioredoxin are related to the severity of obstructive sleep apnea: based on oxidative stress concept. *Sleep and Breathing*, 17, 311-316.


Neurocognitive Disturbance in OSA 135


Neurocognitive Disturbance in OSA


Olaith, M., Skinner, T. C., Clarke, J., Eastwood, P., & Bucks, R. S. (2012). Can we get more from the Epworth Sleepiness Scale than just one score?: A confirmatory factor analysis of the ESS. *Sleep and Breathing, 17*(2), 763-769.


Neurocognitive Disturbance in OSA


undergoing primary CABG and other CPB-assisted cardiac procedures.

*Perfusion, 23, 267-273.*


Young, T., Peppard, P. E., & Gottlieb, D. J. (2002). Epidemiology of obstructive sleep apnea. *American Journal of Respiratory and Critical Care Medicine, 165*.


Appendix 1

The below material was provided as additional online material for chapter 2.

Table 1

Participant characteristics for each study for controls, OSA and Treatment samples.

<table>
<thead>
<tr>
<th>First Author</th>
<th>Year</th>
<th>Source</th>
<th>Type</th>
<th>Controls</th>
<th>OSA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>N</td>
<td>Age</td>
</tr>
<tr>
<td>Antic</td>
<td>2010</td>
<td>UP</td>
<td>T</td>
<td>xx</td>
<td>xx</td>
</tr>
<tr>
<td>Aloia ±</td>
<td>2010</td>
<td>P</td>
<td>T</td>
<td>xx</td>
<td>xx</td>
</tr>
<tr>
<td>Ayalon (older age group)</td>
<td>2010</td>
<td>P</td>
<td>C</td>
<td>7</td>
<td>49.4</td>
</tr>
<tr>
<td>Ayalon (younger age group)</td>
<td>2010</td>
<td>P</td>
<td>C</td>
<td>7</td>
<td>37.9</td>
</tr>
<tr>
<td>First Author</td>
<td>Year</td>
<td>Source</td>
<td>Type</td>
<td>Controls</td>
<td>OSA</td>
</tr>
<tr>
<td>--------------</td>
<td>------</td>
<td>--------</td>
<td>------</td>
<td>----------</td>
<td>-----</td>
</tr>
<tr>
<td>Bailey ±</td>
<td>1993</td>
<td>D</td>
<td>T</td>
<td>xx</td>
<td>xx</td>
</tr>
<tr>
<td>Barbé ±</td>
<td>2001</td>
<td>P</td>
<td>T</td>
<td>xx</td>
<td>xx</td>
</tr>
<tr>
<td>Bardwell ±</td>
<td>2001</td>
<td>P</td>
<td>T</td>
<td>xx</td>
<td>xx</td>
</tr>
<tr>
<td>Bédard</td>
<td>1993</td>
<td>P</td>
<td>C/T</td>
<td>10</td>
<td>50</td>
</tr>
<tr>
<td>Canessa</td>
<td>2011</td>
<td>P</td>
<td>C/T</td>
<td>15</td>
<td>42.2</td>
</tr>
<tr>
<td>Castronovo</td>
<td>2009</td>
<td>P</td>
<td>C/T</td>
<td>14</td>
<td>42.2</td>
</tr>
<tr>
<td>Daurat</td>
<td>2008</td>
<td>P</td>
<td>C</td>
<td>29</td>
<td>50</td>
</tr>
<tr>
<td>Dolan (40-60 age group)</td>
<td>2009</td>
<td>D</td>
<td>T</td>
<td>xx</td>
<td>xx</td>
</tr>
<tr>
<td>Dolan (Over 60 age group)</td>
<td>2009</td>
<td>D</td>
<td>T</td>
<td>xx</td>
<td>xx</td>
</tr>
<tr>
<td>Eaglesman ±</td>
<td>1995</td>
<td>D</td>
<td>T</td>
<td>xx</td>
<td>xx</td>
</tr>
</tbody>
</table>
Neurocognitive Disturbance in OSA

<table>
<thead>
<tr>
<th>First Author</th>
<th>Year</th>
<th>Source</th>
<th>Type</th>
<th>Controls</th>
<th>OSA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ferini-Strambi</td>
<td>2003</td>
<td>P</td>
<td>C/T</td>
<td>23</td>
<td>55.8</td>
</tr>
<tr>
<td>Findley</td>
<td>1991</td>
<td>P</td>
<td>C</td>
<td>21</td>
<td>59</td>
</tr>
<tr>
<td>Froehling</td>
<td>1991</td>
<td>D</td>
<td>T</td>
<td>xx</td>
<td>xx</td>
</tr>
<tr>
<td>Gale</td>
<td>2004</td>
<td>P</td>
<td>T</td>
<td>xx</td>
<td>xx</td>
</tr>
<tr>
<td>Gast</td>
<td>2006</td>
<td>P</td>
<td>T</td>
<td>xx</td>
<td>xx</td>
</tr>
<tr>
<td>Greenberg ±</td>
<td>1987</td>
<td>P</td>
<td>C</td>
<td>14</td>
<td>44.2</td>
</tr>
<tr>
<td>Grenèche</td>
<td>2011</td>
<td>P</td>
<td>C</td>
<td>10</td>
<td>49.6</td>
</tr>
<tr>
<td>Kribbs ±</td>
<td>1993</td>
<td>P</td>
<td>T</td>
<td>xx</td>
<td>xx</td>
</tr>
<tr>
<td>Lojander**</td>
<td>1999</td>
<td>P</td>
<td>T</td>
<td>xx</td>
<td>xx</td>
</tr>
<tr>
<td>Mathieu (Under 50 age group)</td>
<td>2008</td>
<td>P</td>
<td>C</td>
<td>12</td>
<td>38.9</td>
</tr>
<tr>
<td>First Author</td>
<td>Year</td>
<td>Source</td>
<td>Type</td>
<td>Controls</td>
<td>OSA</td>
</tr>
<tr>
<td>--------------</td>
<td>------</td>
<td>--------</td>
<td>------</td>
<td>----------</td>
<td>-----</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>N</td>
<td>Age</td>
</tr>
<tr>
<td>Mathieu (Over 50 age group)</td>
<td>2008</td>
<td>P</td>
<td>C</td>
<td>18</td>
<td>62.5</td>
</tr>
<tr>
<td>McGovern</td>
<td>2008</td>
<td>D</td>
<td>C</td>
<td>30</td>
<td>66.1</td>
</tr>
<tr>
<td>Meurice ±</td>
<td>1996</td>
<td>P</td>
<td>T</td>
<td>xx</td>
<td>xx</td>
</tr>
<tr>
<td>Naëgelé</td>
<td>1995</td>
<td>P</td>
<td>C</td>
<td>17</td>
<td>49</td>
</tr>
<tr>
<td>Neu</td>
<td>2011</td>
<td>P</td>
<td>C</td>
<td>16</td>
<td>36.9</td>
</tr>
<tr>
<td>Quan</td>
<td>2006</td>
<td>P</td>
<td>C</td>
<td>74</td>
<td>57.4</td>
</tr>
<tr>
<td>Redline</td>
<td>1997</td>
<td>P</td>
<td>C</td>
<td>20</td>
<td>48.9</td>
</tr>
<tr>
<td>Rouleau</td>
<td>2002</td>
<td>P</td>
<td>C</td>
<td>18</td>
<td>47.2</td>
</tr>
<tr>
<td>Saunamäki</td>
<td>2010</td>
<td>P</td>
<td>C/T</td>
<td>17</td>
<td>44</td>
</tr>
<tr>
<td>Sharma</td>
<td>2010</td>
<td>P</td>
<td>C</td>
<td>25</td>
<td>45.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First Author</td>
<td>Year</td>
<td>Source</td>
<td>Type</td>
<td>Controls</td>
<td>OSA</td>
</tr>
<tr>
<td>--------------</td>
<td>------</td>
<td>--------</td>
<td>------</td>
<td>----------</td>
<td>-----</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>Age</td>
<td>% Male</td>
<td>Select</td>
<td>ESS</td>
</tr>
<tr>
<td>Sloan (hypoxic group)</td>
<td>1989</td>
<td>D</td>
<td>C</td>
<td>19</td>
<td>44.8</td>
</tr>
<tr>
<td>Sloan (non-hypoxic group)</td>
<td>1989</td>
<td>D</td>
<td>C</td>
<td>19</td>
<td>44.8</td>
</tr>
<tr>
<td>Torelli</td>
<td>2011</td>
<td>P</td>
<td>C</td>
<td>14</td>
<td>57.6</td>
</tr>
<tr>
<td>Verstraeten</td>
<td>2004</td>
<td>P</td>
<td>C</td>
<td>32</td>
<td>47.4</td>
</tr>
<tr>
<td>Walker</td>
<td>1990</td>
<td>D</td>
<td>T</td>
<td>xx</td>
<td>xx</td>
</tr>
</tbody>
</table>

Note: Articles are displayed alphabetically by first author. 'Source' corresponds to the publication status of the study (P, peer reviewed published article; D, dissertation; UP, unpublished data). 'Type' indicates to whether the study was comparing people with OSA to controls or pre and post treatment (C, people with OSA to controls, T, people with OSA pre and post-treatment and C/T, controls and pre/post treatment). The variable ‘Select’ indicates how the study chose controls (PSG, hospital sleep study screening, hPSG, Home or portable sleep study screening, Q, Questionnaire, O, Oximetry screening). Disease severity in the OSA sample is represented by AHI or RDI, Sleep fragmentation index and Time spent below 90% Sa02. Length CPAP indicates how long in months the individuals with OSA were given CPAP treatment for that study. ‘#’ indicates that this CPAP value was estimated from the information given in the report. * = the IQ values given for the Alchanatis et al, 2005 study are given as percentile values. The symbol ‘**’ indicates that the values in these studies gave median and range values which were transformed into mean and standard deviation values using formulae and recommendations of Hozo et al, 2005. The symbol ‘xx’ indicates that these details were not given in or not applicable to the study. ‘Test’ indicates which individual tests were used in the study (1, etc). The symbol ‘±’ indicates that the data used in this meta analysis are only from one group presented in the original paper (Aloia, only the adherent group data are examined; Bailey, only group 1, the group treated with CPAP were used; Barbe, only the true CPAP not the Sham CPAP group is
used; Bardwell, only the CPAP not the placebo group are included; Engleman 1995, Ch 5, only patients who received CPAP treatment and had good compliance are included here not the patients who received conservative treatment; Engleman, 1998, only patients who were given CPAP treatment, not MAS were included here; Greenberg, only patients in the apnoea and healthy group were included; Kribbs, only data for before and after CPAP treatment is included, not for CPAP withdrawal; Meurice, means were combined for the auto-CPAP and constant-CPAP groups; Walker, only group 1 the treated group were used, not the treatment rejecters).
Articles selected for meta-analysis


Bailey GL. Neuropsychological function, sleep and vigilance in men with obstructive sleep apnea syndrome treated with continuous positive airway pressure: The Florida State University; 1993.


Dolan DC. Cognitive dysfunction in middle-aged adults vs. older adults with obstructive sleep apnea. Texas: University of Texas; 2009.


Froehling B. Neuropsychological dysfunction in sleep apnoea syndrome: response to nasal CPAP. Chicago: University of Health Sciences; 1991.


Walker CP. Neuropsychological changes in obstructive sleep apnoea syndrome patients following nasal CPAP treatment. Birmingham: University of Alabama; 1990
Forest plots of the effect size and confidence intervals for each study for pre-treatment OSA participants compared with controls studies in all 5 sub-domains of executive function.

(i) Generativity

Meta Analysis
(ii) **Fluid Reasoning**

<table>
<thead>
<tr>
<th>Model</th>
<th>Study Name</th>
<th>Outcome</th>
<th>Statistics for each study</th>
<th>Std diff in means and 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Std diff</td>
<td>Standard error</td>
</tr>
<tr>
<td>Fixed</td>
<td></td>
<td>Fluid Reasoning</td>
<td>2.04</td>
<td>0.57</td>
</tr>
<tr>
<td>Random</td>
<td></td>
<td>Fluid Reasoning</td>
<td>0.61</td>
<td>0.20</td>
</tr>
</tbody>
</table>

**Meta Analysis**

- Favours OSA
- Favours Controls
(iii) \textit{Inhibition}

<table>
<thead>
<tr>
<th>Model</th>
<th>Study Name</th>
<th>Outcome</th>
<th>Statistics for each study</th>
<th>Std diff in means and 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Std diff</td>
<td>Standard error</td>
</tr>
<tr>
<td>Fixed</td>
<td>Ayal, 2010</td>
<td>Combined</td>
<td>0.76</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td>Canessa, 2011</td>
<td>Combined</td>
<td>1.36</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td>Castronovo, 2009</td>
<td>Combined</td>
<td>1.36</td>
<td>0.42</td>
</tr>
<tr>
<td></td>
<td>Ferini-Staumbi, 2003</td>
<td>Combined</td>
<td>0.94</td>
<td>0.31</td>
</tr>
<tr>
<td></td>
<td>Naegeli, 1995</td>
<td>Combined</td>
<td>2.63</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>Quan, 2006</td>
<td>Inhibition, Stroop</td>
<td>0.10</td>
<td>0.17</td>
</tr>
<tr>
<td></td>
<td>Sharma, 2010</td>
<td>Inhibition, Stroop</td>
<td>2.12</td>
<td>0.30</td>
</tr>
<tr>
<td></td>
<td>Tonelli, 2011</td>
<td>Inhibition, Stroop</td>
<td>0.63</td>
<td>0.37</td>
</tr>
<tr>
<td></td>
<td>Verstraeten, 2004</td>
<td>Combined</td>
<td>0.44</td>
<td>0.26</td>
</tr>
<tr>
<td>Random</td>
<td></td>
<td></td>
<td>0.81</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1.12</td>
<td>0.29</td>
</tr>
</tbody>
</table>

Meta Analysis
(iv) **Shifting**

<table>
<thead>
<tr>
<th>Model</th>
<th>Study Name</th>
<th>Outcome</th>
<th>Std diff in means</th>
<th>Std error</th>
<th>Variance</th>
<th>Lower limit</th>
<th>Upper limit</th>
<th>Z-Value</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed</td>
<td>Bédard, 1993</td>
<td>Shifting, Trails B</td>
<td>0.87</td>
<td>0.47</td>
<td>0.20</td>
<td>-0.05</td>
<td>1.79</td>
<td>1.86</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>Canessa, 2011</td>
<td>Shifting, Trails B</td>
<td>1.16</td>
<td>0.37</td>
<td>0.14</td>
<td>0.43</td>
<td>1.88</td>
<td>3.12</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>Findley, 1991</td>
<td>Shifting, Trails B</td>
<td>1.24</td>
<td>0.31</td>
<td>0.10</td>
<td>0.63</td>
<td>1.85</td>
<td>4.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>Greenberg, 1987</td>
<td>Shifting, Trails B</td>
<td>0.34</td>
<td>0.38</td>
<td>0.14</td>
<td>-0.41</td>
<td>1.08</td>
<td>0.88</td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td>Mathieu, 2009</td>
<td>Shifting, Trails B</td>
<td>0.79</td>
<td>0.40</td>
<td>0.16</td>
<td>0.00</td>
<td>1.57</td>
<td>1.97</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>McGovern, 2008</td>
<td>Shifting, Trails B</td>
<td>0.36</td>
<td>0.29</td>
<td>0.08</td>
<td>-0.20</td>
<td>0.92</td>
<td>1.26</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>Naëgelé, 1995</td>
<td>Combined</td>
<td>2.81</td>
<td>0.48</td>
<td>0.24</td>
<td>1.86</td>
<td>3.76</td>
<td>5.80</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>Quan, 2006</td>
<td>Combined</td>
<td>0.12</td>
<td>0.17</td>
<td>0.03</td>
<td>-0.21</td>
<td>0.45</td>
<td>0.70</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>Redline, 1997</td>
<td>Combined</td>
<td>0.20</td>
<td>0.29</td>
<td>0.08</td>
<td>-0.36</td>
<td>0.76</td>
<td>0.71</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>Roux, 2002</td>
<td>Combined</td>
<td>0.57</td>
<td>0.31</td>
<td>0.03</td>
<td>-0.03</td>
<td>1.17</td>
<td>1.85</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>Saunamäki, 2010</td>
<td>Shifting, Trails B</td>
<td>0.27</td>
<td>0.27</td>
<td>0.08</td>
<td>-0.27</td>
<td>0.81</td>
<td>0.97</td>
<td>0.33</td>
</tr>
<tr>
<td></td>
<td>Sharma, 2010</td>
<td>Combined</td>
<td>0.61</td>
<td>0.25</td>
<td>0.06</td>
<td>0.12</td>
<td>1.10</td>
<td>2.43</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>Sloan, 1989</td>
<td>Combined</td>
<td>0.39</td>
<td>0.32</td>
<td>0.10</td>
<td>-0.24</td>
<td>1.01</td>
<td>1.21</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>Verstraeten, 2004</td>
<td>Combined</td>
<td>0.35</td>
<td>0.25</td>
<td>0.06</td>
<td>-0.15</td>
<td>0.85</td>
<td>1.37</td>
<td>0.17</td>
</tr>
<tr>
<td>Random</td>
<td></td>
<td></td>
<td>0.53</td>
<td>0.08</td>
<td>0.01</td>
<td>0.38</td>
<td>0.68</td>
<td>7.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Fixed</td>
<td></td>
<td></td>
<td>0.65</td>
<td>0.14</td>
<td>0.02</td>
<td>0.38</td>
<td>0.92</td>
<td>4.75</td>
<td>0.00</td>
</tr>
</tbody>
</table>

**Meta Analysis**
### Meta Analysis

#### Updating

<table>
<thead>
<tr>
<th>Model</th>
<th>Study Name</th>
<th>Outcome</th>
<th>Statistic for each study</th>
<th>Std diff in means</th>
<th>Standard error</th>
<th>Variance</th>
<th>Lower limit</th>
<th>Upper limit</th>
<th>Z-Value</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Canessa, 2011</td>
<td>Updating, Digit span WAIS-R</td>
<td>1.49</td>
<td>0.39</td>
<td>0.15</td>
<td>0.73</td>
<td>2.25</td>
<td>3.84</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Castronovo, 2009</td>
<td>Combined</td>
<td>0.27</td>
<td>0.38</td>
<td>0.14</td>
<td>-0.47</td>
<td>1.02</td>
<td>0.72</td>
<td>0.47</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Daurat, 2008</td>
<td>Updating, Digit span BW</td>
<td>1.41</td>
<td>0.30</td>
<td>0.09</td>
<td>0.83</td>
<td>1.98</td>
<td>4.76</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ferrin-Shanks, 2003</td>
<td>Updating, Digit span backwards</td>
<td>0.95</td>
<td>0.31</td>
<td>0.10</td>
<td>-0.44</td>
<td>1.98</td>
<td>3.06</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Greenberg, 1987</td>
<td>Updating, Digit span backwards</td>
<td>0.72</td>
<td>0.39</td>
<td>0.15</td>
<td>-0.44</td>
<td>1.98</td>
<td>3.06</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Greene, 2011</td>
<td>Updating, Digit span BW</td>
<td>3.20</td>
<td>0.65</td>
<td>0.42</td>
<td>-0.64</td>
<td>4.64</td>
<td>4.96</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Malheu, 2008</td>
<td>Combined</td>
<td>0.26</td>
<td>0.38</td>
<td>0.14</td>
<td>-0.44</td>
<td>1.03</td>
<td>0.70</td>
<td>0.49</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Neu, 2010</td>
<td>Updating, Digit span BW</td>
<td>0.66</td>
<td>0.38</td>
<td>0.14</td>
<td>-0.44</td>
<td>1.03</td>
<td>0.70</td>
<td>0.49</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Redline, 1997</td>
<td>Combined</td>
<td>0.38</td>
<td>0.29</td>
<td>0.09</td>
<td>-0.20</td>
<td>0.96</td>
<td>1.29</td>
<td>0.20</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rouleau, 2002</td>
<td>Updating, Arithmetic WAIS-R</td>
<td>0.35</td>
<td>0.30</td>
<td>0.09</td>
<td>-0.25</td>
<td>0.94</td>
<td>1.14</td>
<td>0.23</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Saunamäki, 2010</td>
<td>Combined</td>
<td>0.12</td>
<td>0.27</td>
<td>0.09</td>
<td>-0.25</td>
<td>0.94</td>
<td>1.14</td>
<td>0.23</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Shermes, 2010</td>
<td>Updating, Digit span BW</td>
<td>2.61</td>
<td>0.32</td>
<td>0.11</td>
<td>1.87</td>
<td>3.24</td>
<td>8.04</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Torelli, 2011</td>
<td>Updating, Digit span BW</td>
<td>0.77</td>
<td>0.38</td>
<td>0.14</td>
<td>0.03</td>
<td>1.52</td>
<td>2.04</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Verstraeten, 2004</td>
<td>Updating, Digit span backwards</td>
<td>0.31</td>
<td>0.24</td>
<td>0.06</td>
<td>-0.17</td>
<td>0.79</td>
<td>1.26</td>
<td>0.21</td>
<td></td>
</tr>
<tr>
<td>Fixed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Random</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Favours OSA** **Favours Controls**

-2.00 -1.00 0.00 1.00 2.00
Forrest plots of the effect size and confidence intervals for pre-treatment to post-treatment studies in all 5 sub-domains of executive function.

(i) Generativity

Meta Analysis
(ii) Fluid Reasoning

<table>
<thead>
<tr>
<th>Model</th>
<th>Study Name</th>
<th>Outcome</th>
<th>Statistics for each study</th>
<th>Std diff in means and 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Std diff</td>
</tr>
<tr>
<td>Fixed</td>
<td>Bailey, 1993</td>
<td>Combined</td>
<td></td>
<td>0.27</td>
</tr>
<tr>
<td>Fixed</td>
<td>Bardwell, 2001</td>
<td>Combined</td>
<td></td>
<td>1.48</td>
</tr>
<tr>
<td>Fixed</td>
<td>Canessa, 2011</td>
<td>Combined</td>
<td></td>
<td>0.94</td>
</tr>
<tr>
<td>Fixed</td>
<td>Castronovo, 2009</td>
<td>Combined</td>
<td></td>
<td>0.30</td>
</tr>
<tr>
<td>Fixed</td>
<td>Ferini-Brambilla, 2003</td>
<td>Combined</td>
<td></td>
<td>0.30</td>
</tr>
<tr>
<td>Fixed</td>
<td>Froehling, 1991</td>
<td>Inhibition, Stroop colour-word score</td>
<td></td>
<td>0.50</td>
</tr>
<tr>
<td>Fixed</td>
<td>Gaz, 2006</td>
<td>Combined</td>
<td></td>
<td>0.44</td>
</tr>
<tr>
<td>Fixed</td>
<td>Kribbs, 1993</td>
<td>Inhibition, Stroop</td>
<td></td>
<td>0.30</td>
</tr>
<tr>
<td>Fixed</td>
<td>Walker, 1990</td>
<td>Combined</td>
<td></td>
<td>0.24</td>
</tr>
<tr>
<td>Random</td>
<td>Fixed</td>
<td>Fluid Reasoning</td>
<td></td>
<td>0.58</td>
</tr>
<tr>
<td>Random</td>
<td>Random</td>
<td>Fluid Reasoning</td>
<td></td>
<td>0.59</td>
</tr>
</tbody>
</table>

Meta Analysis
### (iii) Inhibition

<table>
<thead>
<tr>
<th>Model</th>
<th>Study Name</th>
<th>Outcome</th>
<th>Std diff in means</th>
<th>Standard error</th>
<th>Variance</th>
<th>Lower limit</th>
<th>Upper limit</th>
<th>Z-Value</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed</td>
<td>-2.00</td>
<td>0.15</td>
<td>0.02</td>
<td>0.46</td>
<td>0.36</td>
<td>0.13</td>
<td>0.34</td>
<td>1.00</td>
<td>3.75</td>
</tr>
<tr>
<td>Random</td>
<td>0.00</td>
<td>0.18</td>
<td>0.31</td>
<td>1.00</td>
<td>3.75</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

**Meta Analysis**
(iv) **Shifting**

### Meta Analysis

<table>
<thead>
<tr>
<th>Study Name</th>
<th>Outcome</th>
<th>Std diff in means</th>
<th>Standard error</th>
<th>Variance</th>
<th>Lower limit</th>
<th>Upper limit</th>
<th>Z-Value</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aloia, 2010</td>
<td>Shifting, Trails B, adherent</td>
<td>0.28</td>
<td>0.15</td>
<td>0.02</td>
<td>-0.01</td>
<td>0.56</td>
<td>1.90</td>
<td>0.06</td>
</tr>
<tr>
<td>Bailey, 1993</td>
<td>Shifting, Trails B</td>
<td>0.26</td>
<td>0.34</td>
<td>0.12</td>
<td>-0.42</td>
<td>0.93</td>
<td>0.74</td>
<td>0.46</td>
</tr>
<tr>
<td>Barbe, 2001</td>
<td>Shifting, Trails B</td>
<td>2.15</td>
<td>0.33</td>
<td>0.11</td>
<td>1.51</td>
<td>2.80</td>
<td>6.52</td>
<td>0.00</td>
</tr>
<tr>
<td>Bardwell, 2001</td>
<td>Combined</td>
<td>2.03</td>
<td>0.40</td>
<td>0.16</td>
<td>1.25</td>
<td>2.81</td>
<td>5.10</td>
<td>0.00</td>
</tr>
<tr>
<td>Bedard, 1993</td>
<td>Shifting, Trails B</td>
<td>0.29</td>
<td>0.45</td>
<td>0.20</td>
<td>-0.09</td>
<td>1.17</td>
<td>0.44</td>
<td>0.64</td>
</tr>
<tr>
<td>Canessa, 2011</td>
<td>Shifting, Trails B</td>
<td>0.15</td>
<td>0.34</td>
<td>0.12</td>
<td>-0.52</td>
<td>0.82</td>
<td>0.44</td>
<td>0.66</td>
</tr>
<tr>
<td>Dolan, 2000</td>
<td>Combined</td>
<td>0.62</td>
<td>0.51</td>
<td>0.26</td>
<td>-0.58</td>
<td>1.62</td>
<td>1.22</td>
<td>0.22</td>
</tr>
<tr>
<td>Engleman, 1995</td>
<td>Shifting, Trails B</td>
<td>1.09</td>
<td>0.41</td>
<td>0.16</td>
<td>0.50</td>
<td>1.60</td>
<td>2.70</td>
<td>0.01</td>
</tr>
<tr>
<td>Ferri-Strambi, 2003</td>
<td>Shifting, Trails B</td>
<td>0.23</td>
<td>0.34</td>
<td>0.12</td>
<td>-0.44</td>
<td>0.91</td>
<td>0.68</td>
<td>0.50</td>
</tr>
<tr>
<td>Gale, 2004</td>
<td>Shifting, Trails B</td>
<td>0.43</td>
<td>0.36</td>
<td>0.19</td>
<td>-0.32</td>
<td>1.18</td>
<td>1.12</td>
<td>0.26</td>
</tr>
<tr>
<td>Gast, 2006</td>
<td>Combined</td>
<td>0.48</td>
<td>0.35</td>
<td>0.12</td>
<td>-0.21</td>
<td>1.16</td>
<td>1.37</td>
<td>0.17</td>
</tr>
<tr>
<td>Ljungdahl, 1999</td>
<td>Shifting, Trails B</td>
<td>0.72</td>
<td>0.47</td>
<td>0.22</td>
<td>-0.21</td>
<td>1.65</td>
<td>1.52</td>
<td>0.13</td>
</tr>
<tr>
<td>Meurice, 1998</td>
<td>Shifting, Trails B</td>
<td>0.36</td>
<td>0.36</td>
<td>0.13</td>
<td>-0.34</td>
<td>1.06</td>
<td>1.01</td>
<td>0.31</td>
</tr>
<tr>
<td>Saunamäki, 2010</td>
<td>Shifting, Trails B</td>
<td>0.27</td>
<td>0.37</td>
<td>0.13</td>
<td>-0.45</td>
<td>0.99</td>
<td>0.74</td>
<td>0.46</td>
</tr>
<tr>
<td>Fixed</td>
<td>0.56</td>
<td>0.09</td>
<td>0.01</td>
<td>-0.39</td>
<td>0.72</td>
<td>6.81</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Random</td>
<td>0.86</td>
<td>0.18</td>
<td>0.03</td>
<td>0.31</td>
<td>1.00</td>
<td>3.75</td>
<td>0.00</td>
<td></td>
</tr>
</tbody>
</table>

**Fixed Effects**

- 0.56
- 0.09
- 0.01
- -0.39
- 0.72
- 6.81
- 0.00

**Random Effects**

- 0.86
- 0.18
- 0.03
- 0.31
- 1.00
- 3.75
- 0.00

-2.00 -1.00 0.00 1.00 2.00

Favours Pre-Treatment Favours Post-Treatment

---

**Neurocognitive Disturbance in OSA**

- **Model**: Study Name
- **Outcome**: Statistics for each study
- **Std diff in means and 95% CI**: Std diff Standard Lower Upper Z-Value p-Value
- **Fixed**: 0.56
- **Random**: 0.86

---

**Figure**: Graph showing the distribution of differences in means for pre- and post-treatment effects, with confidence intervals for fixed and random effects.
### Core Excerpts

#### Meta Analysis

<table>
<thead>
<tr>
<th>Model</th>
<th>Study Name</th>
<th>Outcome</th>
<th>Std diff in means</th>
<th>Standard error</th>
<th>Variance</th>
<th>Lower limit</th>
<th>Upper limit</th>
<th>Z-Value</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bardwell, 2001</td>
<td>Updating, Digit span</td>
<td>WAIS-R</td>
<td>1.69</td>
<td>0.37</td>
<td>0.14</td>
<td>0.91</td>
<td>2.41</td>
<td>4.58</td>
<td>0.00</td>
</tr>
<tr>
<td>Canessa, 2011</td>
<td>Updating, Digit span</td>
<td>WAIS-R</td>
<td>1.00</td>
<td>0.36</td>
<td>0.13</td>
<td>0.29</td>
<td>1.72</td>
<td>2.76</td>
<td>0.01</td>
</tr>
<tr>
<td>Castronovo, 2009</td>
<td>Combined</td>
<td></td>
<td>0.47</td>
<td>0.38</td>
<td>0.15</td>
<td>-0.29</td>
<td>1.22</td>
<td>1.21</td>
<td>0.22</td>
</tr>
<tr>
<td>Dolan, 2009</td>
<td>Combined</td>
<td></td>
<td>0.32</td>
<td>0.50</td>
<td>0.25</td>
<td>-0.65</td>
<td>1.30</td>
<td>0.65</td>
<td>0.52</td>
</tr>
<tr>
<td>Engstrom, 1995</td>
<td>Updating, Arithmetic</td>
<td>WAIS-R</td>
<td>0.16</td>
<td>0.38</td>
<td>0.14</td>
<td>-0.58</td>
<td>0.91</td>
<td>0.43</td>
<td>0.66</td>
</tr>
<tr>
<td>Ferini-Strambi, 2003</td>
<td>Updating, Digit span</td>
<td>WAIS-R</td>
<td>0.09</td>
<td>0.34</td>
<td>0.12</td>
<td>-0.58</td>
<td>0.76</td>
<td>0.26</td>
<td>0.79</td>
</tr>
<tr>
<td>Gans, 2006</td>
<td>Updating, Digit span</td>
<td>WAIS-R</td>
<td>0.09</td>
<td>0.34</td>
<td>0.12</td>
<td>-0.58</td>
<td>0.77</td>
<td>0.27</td>
<td>0.79</td>
</tr>
<tr>
<td>Hloba, 1993</td>
<td>Updating, Digit span</td>
<td>BW</td>
<td>0.08</td>
<td>0.37</td>
<td>0.13</td>
<td>-0.64</td>
<td>0.80</td>
<td>0.22</td>
<td>0.63</td>
</tr>
<tr>
<td>Saunamäki, 2010</td>
<td>Updating, Digit Span</td>
<td>BW</td>
<td>0.06</td>
<td>0.37</td>
<td>0.13</td>
<td>-0.65</td>
<td>0.78</td>
<td>0.17</td>
<td>0.86</td>
</tr>
<tr>
<td>Fixed</td>
<td></td>
<td></td>
<td>0.44</td>
<td>0.12</td>
<td>0.02</td>
<td>0.19</td>
<td>0.68</td>
<td>3.51</td>
<td>0.00</td>
</tr>
<tr>
<td>Random</td>
<td></td>
<td></td>
<td>0.44</td>
<td>0.19</td>
<td>0.04</td>
<td>0.07</td>
<td>0.81</td>
<td>2.32</td>
<td>0.02</td>
</tr>
</tbody>
</table>

---

(v) **Updating**

---

**Meta Analysis**
Table 1.
Correlation matrix of the relationships between the scores for each ESS item from the clinical sample (n=679) and the community samples (n=356).

<table>
<thead>
<tr>
<th></th>
<th>CLINICAL SAMPLE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ESS8</td>
</tr>
<tr>
<td>ESS8</td>
<td>-</td>
</tr>
<tr>
<td>ESS7</td>
<td>0.32</td>
</tr>
<tr>
<td>ESS6</td>
<td>0.43</td>
</tr>
<tr>
<td>ESS5</td>
<td>0.22</td>
</tr>
<tr>
<td>ESS4</td>
<td>0.36</td>
</tr>
<tr>
<td>ESS3</td>
<td>0.30</td>
</tr>
<tr>
<td>ESS2</td>
<td>0.16</td>
</tr>
<tr>
<td>ESS1</td>
<td>0.19</td>
</tr>
</tbody>
</table>

Note. ESS1 = ‘Sitting and reading’, ESS2 = ‘Watching television’, ESS3 = ‘Sitting inactive in a public place’, ESS4 = ‘As a passenger in a car’, ESS5 = ‘Lying down to rest in the afternoon’, ESS6 = ‘Sitting and talking to someone’, ESS7 = ‘Sitting quietly after lunch’, ESS8 = ‘In a car, while stopped in traffic’
Appendix 3

This material was provided as additional online material for chapter 5.

Table 1.

Correlation matrix of the variables used in the sleep fragmentation/hypoxia model to model the relationships between nocturnal disturbance and cognition. Provided here for readers to replicate the results, as per publication recommendations of Schreiber et. al.45

<table>
<thead>
<tr>
<th></th>
<th>Sleep fragmentation</th>
<th>Estimated IQ</th>
<th>Age</th>
<th>Hypoxia</th>
<th>Executive Function</th>
<th>Short Term Memory</th>
<th>Long Term Memory</th>
<th>Attention</th>
<th>Arousal proportion ODI3</th>
<th>Sleep Efficiency</th>
<th>Mean SaO2</th>
<th>Lowest SaO2</th>
<th>CLOX Ratio</th>
<th>Trails Ratio</th>
<th>Verbal Fluency</th>
<th>Immediate Spatial Recall</th>
<th>CDR Word Learning</th>
<th>Delayed Word Recognition</th>
<th>Delayed Picture Recognition</th>
<th>Choice Reaction Time</th>
<th>Digit Vigilance Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sleep fragmentation</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Estimated IQ</td>
<td>0.00</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.87</td>
<td>0.78</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oxygen Desaturation</td>
<td>-0.91</td>
<td>0.00</td>
<td>0.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variable</td>
<td>Sleep Fragmentation</td>
<td>Estimated IQ</td>
<td>Age</td>
<td>Hypoxia</td>
<td>Executive Function</td>
<td>Short Term Memory</td>
<td>Long Term Memory</td>
<td>Attention</td>
<td>Arousals proportion ODI3</td>
<td>Sleep Efficiency</td>
<td>Mean SaO2</td>
<td>Lowest SaO2</td>
<td>CLOX Ratio</td>
<td>Trails Ratio</td>
<td>Verbal Fluency</td>
<td>Immediate Spatial Recall</td>
<td>CDR Word Learning</td>
<td>Delayed Word Recognition</td>
<td>Delayed Picture Recognition</td>
<td>Choice Reaction Time</td>
<td>Digi1 Vigilance Score</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>---------------------</td>
<td>--------------</td>
<td>-----</td>
<td>---------</td>
<td>-------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>-----------</td>
<td>--------------------------</td>
<td>----------------</td>
<td>-----------</td>
<td>-------------</td>
<td>------------</td>
<td>--------------</td>
<td>----------------</td>
<td>-------------------------</td>
<td>----------------------</td>
<td>--------------------------</td>
<td>-----------------------------</td>
<td>--------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>Executive Function</td>
<td>0.15</td>
<td>0.69</td>
<td>0.25</td>
<td>0.04</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short Term Memory</td>
<td>-0.09</td>
<td>0.41</td>
<td>0.15</td>
<td>0.25</td>
<td>0.44</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Long term Memory</td>
<td>0.27</td>
<td>0.14</td>
<td>0.16</td>
<td>0.15</td>
<td>0.54</td>
<td>0.41</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attention</td>
<td>0.06</td>
<td>0.35</td>
<td>0.01</td>
<td>0.16</td>
<td>0.47</td>
<td>0.34</td>
<td>0.54</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arousals proportion ODI3</td>
<td>-0.01</td>
<td>0.00</td>
<td>0.26</td>
<td>0.01</td>
<td>-0.00</td>
<td>0.00</td>
<td>-0.00</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sleep Efficiency</td>
<td>-0.28</td>
<td>0.00</td>
<td>0.98</td>
<td>0.26</td>
<td>-0.04</td>
<td>0.03</td>
<td>-0.08</td>
<td>-0.02</td>
<td>0.00</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean SaO2</td>
<td>-0.89</td>
<td>0.00</td>
<td>0.62</td>
<td>0.98</td>
<td>0.04</td>
<td>0.24</td>
<td>0.14</td>
<td>0.15</td>
<td>0.01</td>
<td>0.25</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lowest SaO2</td>
<td>-0.56</td>
<td>0.00</td>
<td>-0.01</td>
<td>0.62</td>
<td>0.03</td>
<td>0.15</td>
<td>0.09</td>
<td>0.20</td>
<td>0.01</td>
<td>0.16</td>
<td>0.60</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CLOX Ratio</td>
<td>-0.02</td>
<td>-0.11</td>
<td>-0.20</td>
<td>-0.01</td>
<td>-0.16</td>
<td>-0.06</td>
<td>-0.08</td>
<td>-0.07</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.01</td>
<td>-0.00</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trails Ratio</td>
<td>-0.08</td>
<td>-0.34</td>
<td>0.03</td>
<td>-0.20</td>
<td>-0.49</td>
<td>-0.21</td>
<td>-0.27</td>
<td>-0.23</td>
<td>0.00</td>
<td>0.02</td>
<td>-0.02</td>
<td>-0.01</td>
<td>0.08</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sleep fragmentation</td>
<td>Estimated IQ</td>
<td>Age</td>
<td>Hypoxia</td>
<td>Executive Function</td>
<td>Short Term Memory</td>
<td>Long Term Memory</td>
<td>Attention</td>
<td>Arousals proportion ODI3</td>
<td>Sleep Efficiency</td>
<td>Mean SaO2</td>
<td>Lowest SaO2</td>
<td>CLOX Ratio</td>
<td>Trails Ratio</td>
<td>Verbal Fluency</td>
<td>Immediate Spatial Recall</td>
<td>CDR Word Learning</td>
<td>Delayed Word Recognition</td>
<td>Delayed Picture Recognition</td>
<td>Choice Reaction Time</td>
<td>Digit Vigilance Score</td>
</tr>
<tr>
<td>---------------------------</td>
<td>---------------------</td>
<td>--------------</td>
<td>-----</td>
<td>---------</td>
<td>---------------------</td>
<td>-------------------</td>
<td>------------------</td>
<td>-----------</td>
<td>--------------------------</td>
<td>----------------</td>
<td>-----------</td>
<td>-------------</td>
<td>------------</td>
<td>---------------</td>
<td>----------------</td>
<td>--------------------------</td>
<td>------------------------</td>
<td>------------------------</td>
<td>--------------------------</td>
<td>----------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>Verbal Fluency</td>
<td>0.12</td>
<td>0.54</td>
<td>0.11</td>
<td>0.03</td>
<td>0.78</td>
<td>0.34</td>
<td>0.43</td>
<td>0.36</td>
<td>-0.00</td>
<td>-0.03</td>
<td>0.03</td>
<td>0.02</td>
<td>-0.12</td>
<td>-0.38</td>
<td>1.00</td>
<td>0.16</td>
<td>0.10</td>
<td>0.16</td>
<td>0.10</td>
<td>1.00</td>
<td>0.16</td>
</tr>
<tr>
<td>Immediate Spatial Recall</td>
<td>-0.04</td>
<td>0.19</td>
<td>0.08</td>
<td>0.11</td>
<td>0.20</td>
<td>0.45</td>
<td>0.19</td>
<td>0.15</td>
<td>0.00</td>
<td>0.01</td>
<td>0.11</td>
<td>0.07</td>
<td>-0.03</td>
<td>-0.10</td>
<td>0.16</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDR Word Learning</td>
<td>-0.03</td>
<td>0.14</td>
<td>-0.07</td>
<td>0.08</td>
<td>0.15</td>
<td>0.34</td>
<td>0.14</td>
<td>0.12</td>
<td>0.00</td>
<td>0.01</td>
<td>0.08</td>
<td>0.05</td>
<td>-0.02</td>
<td>-0.07</td>
<td>0.12</td>
<td>0.15</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delayed Word Recognition</td>
<td>-0.12</td>
<td>-0.07</td>
<td>0.07</td>
<td>-0.25</td>
<td>-0.19</td>
<td>-0.46</td>
<td>-0.25</td>
<td>0.00</td>
<td>0.04</td>
<td>-0.07</td>
<td>-0.04</td>
<td>0.12</td>
<td>-0.19</td>
<td>-0.08</td>
<td>-0.06</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delayed Picture Recognition</td>
<td>0.13</td>
<td>0.07</td>
<td>0.11</td>
<td>0.07</td>
<td>0.25</td>
<td>0.19</td>
<td>0.46</td>
<td>0.25</td>
<td>-0.00</td>
<td>-0.04</td>
<td>0.07</td>
<td>0.04</td>
<td>-0.04</td>
<td>-0.12</td>
<td>0.20</td>
<td>0.08</td>
<td>0.06</td>
<td>-0.21</td>
<td>1.00</td>
<td>-0.21</td>
<td>1.00</td>
</tr>
<tr>
<td>Choice Reaction Time</td>
<td>0.20</td>
<td>0.11</td>
<td>0.06</td>
<td>0.40</td>
<td>0.31</td>
<td>0.74</td>
<td>0.40</td>
<td>-0.00</td>
<td>-0.06</td>
<td>0.11</td>
<td>0.07</td>
<td>-0.06</td>
<td>0.14</td>
<td>0.34</td>
<td>0.10</td>
<td>0.10</td>
<td>-0.34</td>
<td>0.34</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Digit Vigilance Score</td>
<td>0.02</td>
<td>0.12</td>
<td>0.78</td>
<td>0.06</td>
<td>0.16</td>
<td>0.19</td>
<td>0.35</td>
<td>0.00</td>
<td>-0.01</td>
<td>0.05</td>
<td>0.03</td>
<td>-0.03</td>
<td>-0.08</td>
<td>0.13</td>
<td>0.05</td>
<td>0.04</td>
<td>-0.09</td>
<td>0.09</td>
<td>0.14</td>
<td>1.00</td>
<td>0.14</td>
</tr>
</tbody>
</table>
Table 2.

Table shows Beta weights of paths from Epworth total score and somnificity factors to the domains of cognition measured in the model. None of the weights were significant, and hence sleepiness was removed from the final model.

<table>
<thead>
<tr>
<th>Fit Index</th>
<th>Attention</th>
<th>Short Term Memory</th>
<th>Long Term Memory</th>
<th>Executive Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total ESS Score</td>
<td>-0.51</td>
<td>-0.39</td>
<td>-0.16</td>
<td>0.15</td>
</tr>
<tr>
<td>Low situational somnificity</td>
<td>-0.13</td>
<td>-0.06</td>
<td>-0.07</td>
<td>0.14</td>
</tr>
<tr>
<td>Moderate situational somnificity</td>
<td>0.03</td>
<td>-0.01</td>
<td>-0.08</td>
<td>0.12</td>
</tr>
<tr>
<td>High situational somnificity</td>
<td>-0.07</td>
<td>-0.04</td>
<td>-0.21</td>
<td>0.12</td>
</tr>
</tbody>
</table>
Figure 1. Model III - The Sleep Disturbance Model showing the relationship between nocturnal disturbance and cognitive constructs; Attention, Long Term Memory, Short Term Memory, and Executive Function, controlling for IQ. Note. This model was also explored controlling for daytime sleepiness (using the ESS). ESS scores produced poor fit and were removed. This model fit the data less well than the hypothesised model which separated out sleep fragmentation from hypoxia.
Figure 2. Model IV - The AHI Model showing the relationship between nocturnal disturbance and cognitive constructs; Attention, Long Term Memory, Short Term Memory, and Executive Function, controlling for IQ. Note. This model was also explored controlling for daytime sleepiness (using the ESS). ESS scores produced poor fit and were removed. This model provided poor fit to the data.
The Attention Mediated Model showing the relationship between nocturnal disturbance and cognitive constructs; Long Term Memory, Short Term Memory, and Executive Function, controlling for IQ. Note. This model failed to converge suggesting it was not a good account of the data, therefore no weights.