# Investigating the role of sleep fragmentation in declarative memory and affect

Thomas James Hunter

Doctor of Philosophy

University of York

Psychology

December 2022

# Abstract

Sleep plays an active role in the formation and storage of declarative memories. These processes are thought to depend both on the duration and continuity of sleep. This thesis investigated the proposition that sleep fragmentation uniquely contributes to the variance in encoding error and overnight forgetting whilst controlling for sleep duration. In Experiment 1, closely related word pairs had a general advantage over more distal word pairs at encoding but there were no group differences in overnight retention between the two conditions in a novel word learning task. In Experiment 2, results suggested that interference does not occur between spatial and verbal declarative memory tasks during sleep. Two large naturalistic pre-sleep/post-sleep online memory studies (Experiments 3 and 4) went on to use hierarchical multilevel modelling to control for the duration of sleep statistically. In Experiment 3, increased awakenings were associated with increased encoding error and increased overnight forgetting in a sample of new parents and healthy controls, but only when the level of encoding error was controlled for. In Experiment 4, having Restless Legs Syndrome, characterised by sleep fragmentation, was also associated with increased encoding error, and overnight forgetting, again only when the level of encoding error was controlled for. A series of mixed-effects mega-analyses were carried out in Chapter 5 to better understand the degree to which subjective and more objective sleep measures are related to one another (e.g., sunshine and happiness) and agree with one another (e.g., a sundial and a clock). Chapter 5 showed that subjective and objective measures are related to and in agreement with one another, albeit weakly, and even less so among those with sleep disorders. It was argued that continuity is important for the formation and storage of declarative memories independently of time slept, and implications arising out of these insights are discussed.

# Table of Contents

# List of Tables

# List of Figures

# Acknowledgements

This journey was not easy. It forced me to evolve and become a better person and I want to thank everyone who was there for me along the way. Firstly, I'd like to thank my supervisors Lisa and Gareth. You have both been beacons of patience and professionalism over the years and have always been there to show me a better way. I also want to thank my loving partner Maeve, who has held my hand no matter what challenges there were to face and what sacrifices they required.

I would also like to thank my mum, who is a pure legend. Finally, I would like to thank my boys Vincent and Jimmy. Your smiles and energy are an inspiration, and I cannot wait to watch you take your own journeys.

# Declaration

I declare that this thesis is a presentation of original work carried out under the joint supervision of Dr Lisa Henderson and Prof. Gareth Gaskell and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

# Chapter 1: Literature review

## Introduction

Sleep is clearly highly adaptive or else it would not have been selected for. Its purpose, however, has for the longest time remained unclear (Frieberg, 2020). It has been thought to contribute to health, longevity, and restoration (see Ezenwanne, 2011, for review). Coinciding with technological advances including the dawn of polysomnography and functional neuroimaging respectively, sleep has been implicated in the formation and storage of memories (e.g., Diekelmann and Born, 2010). Arising out of this surge of research interest is the increasing understanding that sleep benefits from both sufficient duration and continuity (Liu, Meng, Wiggin, Zhang, Rosbash, and Griffith, 2019). Given that the potential for disturbed sleep as part of normal modern life is at an unprecedented high (Smith, Croy, Gren, and Waye, 2013), gaining a better understanding of the influence of disturbed sleep on cognitive functioning is a societal and economic imperative. This review focuses on the degree to which sleep duration and fragmentation independently contribute to disruption in the encoding and consolidation of declarative memories.

Converging evidence suggests that sleep actively participates in memory consolidation (the stabilization of a memory trace after an initial acquisition; Rasch and Born, 2013). Consolidation has been described as a “dynamic, generative, transformative, and lingering process that is posited to balance maintenance of useful experience-dependent internal representations of the world with the need to adapt these representations to the changing world” (Dudai, Karni, and Born, 2015). Since learning and memory deficits are found in many disorders in which sleep is fragmented (Lui et al., 2019), a rich literature exists examining the influence of sleep fragmentation on cognition in such populations (for example, see Bucks, Olaithe, and Eastwood, for a 2013 meta-review on neurocognitive function in Obstructive Sleep Apnoea Syndrome which considered 33 reviews on the topic). However, far fewer investigations exist specifically dedicated to better understanding the influence of sleep fragmentation on memory consolidation cutting across both typical and atypical populations, with fewer still testing the independent impact of sleep fragmentation and duration in the same paradigm. Furthermore, in studies that have examined sleep fragmentation alongside sleep duration, the emphasis has tended to have been placed

on the deleterious effects of impaired sleep duration; see Laharnar et al., 2020, for a review).

The current review seeks to provide an overview of the literature on sleep and declarative memory in typical adult populations before evaluating the influence of sleep fragmentation on encoding and consolidaiton. The latter will be achieved by reviewing studies of fragmentation-induced declarative memory deficits from the sleep pathologies literature. When few studies consider declarative memory specifically, I will consider cognitive function and hippocampus-dependent memory more broadly. As such, evidence will be considered from investigations focused on the deleterious effects of Obstructive Sleep Apnoea Syndrome and Restless Legs Syndrome. As well as this, evidence will also be considered from other relevant fields, such as those focussing on postpartum sleep disturbance, and from the animal literature. I will then complete the discussion by applying what was learned in the main body to the upcoming chapters in this thesis.

## Sleep and Memory Consolidation

Sleep is defined as a “natural and reversible state of reduced responsiveness to external stimuli and relative inactivity, accompanied by a loss of consciousness [which] occurs in regular intervals and is homeostatically-regulated” (Borbely and Achermann, 1999). This review focuses on sleep in young adults and as such most of the following discussion will be related to this age group.

Human sleep begins with NREM stage 1 sleep before alternating between the following three core sleep stages in a cyclic manner: NREM stage 2 sleep; NREM stage 3 sleep (slow-wave sleep; SWS) and; rapid eye-movement (REM) sleep. A typical night of sleep would be dominated by SWS in the first half of the night and give way to increasing bouts of REM sleep in the second half the night (Rasch and Born, 2013). More specifically, a typical young adult might expect to spend a short amount of time in Stage 1 (2-5%), the bulk of the evening in Stage 2 (45-55%), and the remainder of the evening in either SWS (13-23%) or REM (20-25%; Carskadon & Dement, 2011). Although sleep duration is highly variable and therefore difficult to generalize, around 7.5 hours/night during the week and around 8.5 hours/night during

weekends closely resembles what might be considered a “typical” sleep schedule in young adults (see Carskadon & Dennet, 2011). Concerning fragmentation, in studies ran in sleep laboratories with at least one night of habituation, control samples of participants have shown considerable variability in mean sleep fragmentation measures, for example ranging from around 50 to 75 arousals each night (Waye, Elmenhorst, Cory, and Pederson, 2019; Smith, Orgen, Morsing, and Waye, 2019, and Smith, Croy, Orgen, and Waye, 2013). This natural variability in the degree of sleep fragmentation in adults can also be sensitive to several well-documented external factors including noise, light, ambient temperature, vibration, and humidity (see Caddick et al, 2018, for a comprehensive review of the optimal conditions for human sleep).

***Measuring sleep parameters***

The most accurate way to measure human sleep is using full night polysomnography (PSG), with most memory studies focusing on EEG signals which paint a clear picture of sleep macro- and microstructure (see Mendonca, Mostafa, Morgado-Dias, Ravelo-Garcia, and Penzel, 2019, for a review of approaches for sleep quality analysis). This approach can provide sleep researchers with data on the number of arousals, nocturnal awakenings, wake after sleep onset (WASO) and sleep efficiency (% of time in bed spent sleeping), all of which are informative in assessing the continuity of a night’s sleep. Full night PSG is not always possible, however, as it can be costly and time-consuming. For this reason, several other approaches are also used in sleep fragmentation research.

A less costly, time-consuming, and burdensome (both for the participant in terms of invasiveness and the researcher in terms of set-up and analysis) way of collecting objective sleep data is through use of the wrist worn actigraphy. Here, an actigraphy watch is typically placed on the participant’s wrist and worn overnight (ideally for at least 2 weeks before any behavioural testing, although some recent evidence suggests a week is sufficient; Briscoe, Hardy, Pengo, Kosky, Williams, Hart, & Steier, 2014) and is sensitive to body movements which are used to estimate sleep parameters using computer algorithms (Martin and Hakim, 2011). Notable output includes total sleep time (TST), sleep efficiency, WASO, number of awakenings (one

common measure of sleep fragmentation), and a sleep fragmentation index (SFI; the sleep fragmentation index is an actigraphy measure of restlessness which takes into account nighttime motor activity and time spent asleep with time spent not asleep whilst in bed, Aubert-Tulkens et al., 1987). The consensus of use of wrist actigraphy is that it is:(a) less accurate than PSG (e.g., Conley, Knies, Batten, Ash, Miner, Hwang, Jeon, and Redeker, 2019), yet (b) a valid measure of way of capturing objective data on sleep parameters (e.g., Conley et al., 2019; Withrow, Roth, Koshorek, and Roehrs, 2019; Marino, Rueschman, Winkelman, Ellenbogen, Solet, Dulin, Berkman, and Buxton, 2013), within which: (i) it is highly sensitive (*detects sleep in agreement with PSG*, e.g., 0.97 sensitivity in a general estimation model created by Marino et al 2013) and accurate (*total proportion correct*, 0.86, Marino et al., 2019), but (ii) not very specific (*detects wake in agreement with PSG,* e.g., 0.33, Marino et al., 2019).

However, these general statements come with several caveats. The first of these is that actigraphy is less accurate when used in clinical populations than healthy ones (e.g., a systematic review and meta-analysis comparing healthy and individuals with chronic conditions such as depression, insomnia, and diabetes, generally found overestimated TST, sleep efficiency, and underestimated WASO; Conley et al., 2019; see also specific investigations with poorer performance of actigraphy in those who experience periodic limb movement disorder; Smith, McCrae, Cheung, Martin, Harrod, Heald, and Carden, 2018; cerebral palsy; Licis, Xue, Boyd, Hoyt, Yo-El, and Ju, 2020; Autism Spectrum Disorder; Yavuz-Kodat, Reynaud, Geoffray, Limousin, Franco, Bourgin, and Schroder, 2019). A second caveat, especially for sleep fragmentation researchers, is that actigraphy seems to perform more poorly overall on measures of sleep fragmentation (arousals, awakenings, WASO, SE) than deprivation (which relies on TST estimates). For example, Wang et al (2008, in Mendonca et al., 2019) suggest that actigraphy underestimate arousals, Withrow et al (2019) found underestimated WASO with actigraphy with no difference in TST estimation, and Conley et al (2019) found underestimated SE and WASO with actigraphy vs PSG.

And finally, an even less costly, time-consuming, and burdensome approach to capturing data on sleep parameters is with self-report estimates. Despite the obvious limitation that individuals are being asked to opine about a period for which they were not awake, self-report data can be very useful in sleep research. Sleep fragmentation

can be measured subjectively by having the individual estimate their number of nocturnal awakenings, by using their estimated sleep and wake time and estimated time spent in bed to calculate sleep efficiency, and by simply asking participants how fragmented they felt their sleep was, typically on a scale of 0 (extremely) to 10 (not at all). Croy, Smith, Gidlof-Gunnarsson, and Persson-Waye (2017) retrospectively assessed 25 items capturing sleep parameters and compared them with PSG (N=47) and found a moderate correlation between objective and subjective number of awakenings (*r*=.31). As well as this, when comparing subjective estimates of sleep with PSG data in young men and women across 3 nights, Baker, Maloney, and Driver (1999; N=20) found that participants reported significantly fewer awakenings than PSG on each night. Thurman and colleagues (2018) explored statistical relatedness and agreement between subjective and objective sleep measures and found moderate relatedness and agreement for sleep duration and poor relatedness and no agreement for the number of awakenings. Taken together, several studies already exist which have to some degree explored the relatedness and agreement between subjective and objective measures of sleep. They are at times disparate, contradictory and inconsistent with regard to methods used.

## Sleep and declarative memory consolidation

Human memory is typically conceptualized as entailing the encoding, consolidation, and retrieval of information. Encoding is, broadly speaking, the initial learning of information and first generation of a new representation of memory (Melton 1963). This activity strengthens synapses (*synaptic long-term potentiation;* Collinridge, Peineau, Howland, and Wang, 2010). Consolidation can be defined as “the transformation over time of experience-dependent internal representations and their neurobiological underpinnings” (Dudai et al., 2015, p.20). Consolidation reorganizes neuronal circuitry at both the cellular and synaptic level and relies upon repeated reactivations to wed isolated representations harmoniously with the individual’s acquired network of integrated knowledge (Dudai et al., 2015). Retrieval is the ability to access the information when you need it (e.g., McDermott, LaHoste, Chen, Musto, Bazan, and Magee, 2003). The sleeping brain is thought to provide optimal conditions for consolidation processes and to actively facilitate them.

This review focuses on the influence of sleep fragmentation on declarative memory – a separate memory system dedicated to facilitating episodic memories including those pertaining to facts, words, and objects, stored alongside contextual information (Rasch and Born, 2013). Implicating sleep in human memory processing extends back to Ebbinghaus (1885), who observed reduced forgetting when the delay period between encoding and retrieval featured sleep, and since then there have been numerous studies consistently demonstrating benefits of memories form periods of sleep (e.g., Plihal and Born, 1997; 1999). Strong evidence for the role of sleep in declarative memory consolidation has been found more recently using reactivation paradigms, in both human and animal studies. The key idea here is that memory traces are reactivated multiple times in sleep making them both hippocampus-independent over time and less vulnerable to interference (see Anthony and Paller, 2017, for a review). For example, Lehmann and McNamara (2011) conditioned fear responses in rodents and then re-exposed only half of the sample to the fear trigger during wake for five days, after which half the sample received hippocampal damage. They found clear evidence that if the rodents had not been exposed to the fear trigger in the 5-day period between initial exposure and hippocampal damage, the damage would greatly affect memory. As well as this, importantly, if the rodents had been subject to repeated exposures of the fear trigger in the 5-day interval, memory was not affected by hippocampal damage. In a very influential study from the human literature, Rudoy, Voss, Westerberg, and Paller (2009), triggered reactivations in half of their sample by playing sounds in NREM sleep associated with learning during wake (the authors used an object location task and paired each object with a sound) and found improved recall among participants who were exposed to the characteristic sounds for the object stimuli. This approach (TMR: Targeted Memory Reactivation) has become very popular in modern sleep and memory research, with multiple demonstrations of TMR-enhanced consolidation effects (e.g., Cairney, Lindsay, Sobcsak, Paller, and Gaskell, 2016). Taken together, compelling evidence exists for the role of sleep in declarative memory consolidation.

### *Synaptic Homeostasis*

The dominant modern theory for the encoding of declarative information is the Synaptic Homeostasis Hypothesis (Tononi & Cirelli, 2018). For Tononi and Cirelli

(2018), sleep is the price we pay for the capacity to absorb information during wake. Specifically, for the authors, the Synaptic Homeostasis Hypothesis suggests that numerous populations of interconnected neurons are strengthened through wakeful exposure to encountered stimuli perceived as salient. This process is costly and has a limit before saturation is reached, at which point the need to reset the balance of available resources is experienced by the individual as increased sleep pressure. This balance is primarily restored during periods of uninterrupted sleep. According to Dudai (2012), synaptic downscaling is best understood as a sub-process occurring within systems consolidation (below), involving repeated bursts lasting only hours after encoding. Considerable evidence has been found which is consistent with synaptic homeostasis, for example field excitatory postsynaptic potentials (shifts in membrane potential caused by a neurotransmitter binding to a receptor on a receiving cell; fEPSPs; Pereda, 2014) gradually become more abundant across sustained periods of wakefulness and drop off with sleep in cortex and hippocampus in both rats and humans (Huber et al., Norimoto et al., 2018; Vyazovskiy et al., 2008; in Tononi and Cirelli, 2018, review). To put it simply - there are limits to the amount of material that can be encoded in each wakeful period before the pressure to reset during sleep becomes intolerable.

### *Systems Consolidation*

Systems consolidation involves reorganizing long-term memory representations after encoding so that they can be integrated with existing knowledge (Dudai and Morris, 2000). There is considerable variation with regards to how long this process lasts, ranging from days to months (Dudai et al., 2015; Wang and Morris, 2010). the key idea here is that hippocampal system and neocortical networks are anatomically separated, with the hippocampal system specializing in rapid, temporary acquisition of new memories to be stored in isolation for a short term, and neocortical networks specializing in the much more gradual integration of the old with the new (Rasch & Born, 2013). Considerable evidence has been found for this dual memory systems model, for example from lesion studies which suggest that hippocampal lesions prevent new declarative memories from being formed (see Corkin, 2002, for discussion).

A core assumption of the active system consolidation hypothesis is that recurrent reactivation (replay of firing patterns in neuron assemblies) of recently formed memory traces facilitate memory consolidation (see Rasch & Born, 2013, for a review). This account goes a step further than distinguishing between key features of the macrostructure of sleep (e.g., sleep stages) and seeks to conceptualize the neural underpinnings mediating the relationship between sleep and memory consolidation. These reactivations occur during SWS (a stage of deep sleep dominated by slow wave activity: EEG activity characterized by waves in the 0.5-4.0Hz band and oscillations <1-Hz; Achermann and Borbely, 1998) and bring about neuronal reorganization described above, namely the gradual transfer of newly acquired memory representations from short- to long-term integrated storage. Reactivations occur during *sharp-wave ripples* (large amplitude patterns of oscillation which only occur in the hippocampus and nearby brain regions; Buzsaki, 1998) and are driven by the synchronous activity of slow oscillations and thalamo-cortical spindles (oscillations which take place in the 10-15 Hz range in all mammals, predominantly found in Stage 2 in humans Gennaro, Ferrara, and Bertini, 2000). Staresina and colleagues (2015) used direct intracranial electroencephalogram recording among individuals with epilepsy to provide further evidence of the workflow of the key oscillations involved in systems consolidation. Specifically, current consensus is that slow oscillations control spindles, which in turn gather sharp wave ripples in their troughs in a manner dedicated towards triggering reactivations, which create the conditions in the intended neocortical transfer sites to integrate the to-be-remembered representation among long term storage networks.

Multiple sources of evidence using a variety of methods support active systems consolidation. For example, neurobiological evidence exists from single-cell recordings demonstrating that hippocampal replays occur in SWS (Ji & Wilson, 2007). As well as this, there is also an abundance of evidence that slow oscillations at the heart of facilitating sleep's serviceable effect on memory consolidation (notable animal literature: Kattler, Dijk, & Borbely, 1994; Vyazovskiy, Borbely, & Tobler, 2000; causal evidence from tDCS in humans: Marshall, Molle, Hallschmidt, & Born, 2004), and that both sleep spindles and sharp-wave ripples are robustly associated with memory processing during sleep (Spindles: Gais, Molle, Helms, & Born, 2002; Berner, Schabus, Wienerroither, & Klimesch, 2006; Tamminen, Payne, Stickgold, Wamsley,

& Gaskell, 2010; Sharp-wave ripples: Axmacher, Helmstaedter, Elger, & Fell, 2012; Eschenko, and Sara, 2008; SO-Spindle-coupling-events: Staresina et al., 2015, Schreiner et al., 2021).

### *The Interaction of Synaptic Homeostasis and Active Systems*

This review mostly focusses on Synaptic Homeostasis as pertaining to the encoding of declarative memories and Active Systems as pertaining to the consolidation of declarative memories. It is important to note however that these theories are complementary of each other and more interactive than the narrative above and throughout might suggest at times. Whilst it is broadly thought that Synaptic Homeostasis ensures capacity for future encoding, and Active Systems integrates memories into long-term storage, processes associated with Synaptic Homeostasis may also benefit consolidation and process associated with Active Systems may benefit encoding. For example, global downscaling during sleep, where less important synapses are pruned, may get rid of unwanted noise during slow wave sleep by pruning weak synapses and boost the signal of to-be-remembered representations, making it easier and more efficient to identify them, reactivate them and integrate them into the neo-cortex (see Tononi and Cirelli, 2014, for discussion). Active Sytems consolidation can also influence encoding, for example by shaping future encoding. This could occur, for example, by shaping schemas within which future information might be encoded more efficiently, (e.g., Tse et al., 2007).

In sum, a robust and informative literature exists suggesting a transformational process of rapid acquisition of isolated memory representations in the hippocampus to gradual integration with existing knowledge in the neocortex mediated by a delicate and intricate synchrony of oscillatory activity.

### *Alternative Accounts of Memory Consolidation: Contextual Binding*

Despite the popularity and convincing evidence base supporting systems consolidation, there are alternative accounts that are worth considering. At its core, Yonelinas, Ranganath, Ekstrom and Wiltgen (2019) propose a Contextual Binding account of the consolidation of episodic memories in which the hippocampus plays a

necessary and not just temporary role in episodic memory. Importantly, the role of sleep within a Contextual Binding account is to protect against interference. For Contextual Binding, therefore, forgetting is interference occurring before or after the study event. This is compared to forgetting as reflecting a failure of systems consolidation within Active Systems Consolidation. Furthermore, SWS within Contextual Binding represents the sleep stage which is best at reducing interference both because of its deep electrophysiological properties and its place in the typical architecture of sleep. It is argued by Contextual Binding theorists that the part of the sleep phase occupied by SWS is the point at which risk of contextual interference is at its greatest. For the authors, systems consolidation cannot account for phenomena including proactive interference, item similarity, context re-instatement, pre-encoding sleep benefits, and normal forgetting rates in amnesia. Contextual Binding accounts of memory consolidation are in their infancy, but clearly pose some fresh and interesting challenges to more traditional accounts of sleep and declarative memory consolidation.

Sleep fragmentation may provide a useful lens through which to test some of the claims of Contextual Binding theory. For example, if the core function of SWS is to protect to-be-remembered representation from contextual interference, then it would hard to account for sleep-fragmentation induced forgetting where the awakenings involved are very brief, intermittent, and outside the recollection of the individual experiencing them (see study by Winser et al., 2013 in which it was found that awakenings need to be on average 4 minutes and 19 seconds long to be remembered). Finding that unremembered awakenings were associated with forgetting would be difficult but not impossible to account for and would involve description of how very brief episodes of partial consciousness outside of the reach of recollection could alter context.

## The Influence of Sleep Fragmentation on Declarative Memory

Typically sleep accounts for approximately one third of the human experience, is circadian driven, and sleep pressure intensifies proportionally with time just spent awake (Reutrakul and Van Cauter, 2014). However, these control mechanisms can be overridden (e.g., by human behavior such as working a high demanding job;

Laharner et al., 2020), or disrupted, either by a sleep disorder (such as Obstructive Sleep Apnoea Syndrome, Carreras, Zhang, Peris, Qiao, Gileles-Hillel, Li, Wang, and Gozal, 2014; and Restless Legs Syndrome, Trenwalder and Paulus, 2010), or by a suboptimal sleeping environment (such as living near a railway station, Aasvang, Overland, Ursin, and Moum, 2011; or when caring for a newborn, Insana, Williams, and Montgomery-Downs, 2013).

Sleep fragmentation is a hallmark of poor sleep quality, and negatively affects cognition, performance, and health (Laharna et al., 2020). Moving away from studies focusing on “normal” sleep , this section reviews studies from populations with rich natural variation in sleep fragmentation and tries to provide a brief but comprehensive summary of the findings with regards to declarative memory. In cases where the evidence relating to declarative memory (encoding and/or consolidation) is sparse, I will consider evidence relating to the association between sleep fragmentation and cognition more generally. Finally, I will consider some highly relevant findings from the animal literature.

### *Measuring Sleep Fragmentation*

Sleep fragmentation is easy to describe at a surface level but becomes more difficult with increasing depth of analysis. At a surface level, sleep fragmentation is just the collective term for repeated instances of waking up whilst trying to sleep, something most if not all individuals experience with some degree of regularity (e.g., on average 5 times per night, Winser et al., 2013). According to O’Hayon and colleagues (2010), for example, approximately one third of the population report being bothered by awakening during the night at least three times per week. Despite this high prevalence, however, there is no one research grade definition for sleep fragmentation. This is perhaps due to how intuitive the concept seems at a surface level. This is also arguably partially the case since many techniques exist purporting to capture the umbrella term sleep fragmentation. This section explores these in detail before outlining which measures are most suited for better understanding the relationship between sleep fragmentation and declarative memory.

Some forms of sleep fragmentation can be measured by all of the domains of sleep measurements described in greater detail above: polysomnography, actigraphy,

and self-report. The most common way to measure sleep fragmentation is to simply define what constitutes being awake, and to calculate the number of times an awakening occurs within the sleep phase. The American Academy of Sleep Medicine defines wake, subject to a few exceptions, as a 30-second epoch characterised by a greater than 50% occurrence of alpha waves over the occipital region (Stage W; AASM, 2020). The number of awakenings, therefore, within a polysomnography framework, is the number of times after sleep onset and before sleep offset in which the criteria for Stage W occurs. With actrigraphy, wakefulness is typically inferred from increased physical activity for a sufficiently extended period of time (e.g., Morgenthaler et al., 2007). The degree of activity and length of time can vary depending on the specific device and scoring system used, but the same general pattern exists as is the case for polysomnography applies – define what it means to be awake, and then count those. Despite both being demonstratively sensitive and specific measures, there is a clear consensus among sleep researchers that polysomnography is the gold standard objective measurement tool for sleep (Iber et al., 2007). Despite being significantly more practical and cost-effective (see Ancoli-Israel et al., 2003, for discussion), a shortcoming of actigraphically recorded sleep is a greater likelihood than polysomnography of misclassifying sleep as wake (Marino et al., 2013), especially among those with clinical disorders (Sadeh et al., 2011). Among those seeking to measure sleep fragmentation objectively, therefore, it seems clear that if need for precision is high and concern for time and resources are low, then polysomnography is a more attractive option. If research is being carried out at a scale and pace, however, actigraphy is a more practical resource to draw upon.

A third option exists, of course – just asking (see Buysse, 2014, for a discussion of self-report measures in sleep research). Measuring sleep fragmentation using self-report can follow the same pattern as above, i.e., by defining wake and calculating the number of awakenings which occurred during the intended sleep phase. It is clear, however, that capturing sleep fragmentation using this approach is quite distinct from its more objective counterparts. It is also clear that this approach is clearly lacking in some regards compared to more objective measures yet potentially more informative in others. Considering the former, self-reported sleep relies on the individual to remember and calculate this metric by themselves, using their own criteria for calculation. Winser and colleagues (2013) showed for example that the recall

threshold for a nocturnal awakening is on average 4 minutes and 19 seconds. Many awakenings, therefore, are beyond the reach of recall. The self-reporter must therefore rely on other faculties to report how many times they woke up the night before. What is left at the individual's disposal is subjective experience – the relationship between sleep fragmentation and sleep quality is intuitively assumed and most individuals will report higher number of awakenings in accordance with poorer perceived sleep quality (Harvey et al., 2008).

This element of subjective experience clearly captures information that polysomnography and actigraphy alone cannot but is also clearly a confound to take into account for those seeking to better understand the relationship between sleep fragmentation and declarative memory. There is at least some evidence, however, to suggest that the subjective experience of sleep quality *is* the subjective experience of sleep fragmentation and its daytime consequences. Harvey and colleages (2008) used a "Speak Freely" approach to investigating subjective sleep quality and analysed participants' responses when asked to describe what a good or poor-quality night of sleep was. What the researchers found, among healthy sleepers and those with insomnia, was that a poor-quality night of sleep was defined as waking up many times during the night and feeling fatigued the next day. If this were to be the case, then subjectively measured number of awakenings can capture aspects of sleep fragmentation that objectively measured approaches cannot and are therefore an important tool. Whether or not this is the case, however, is still unclear. Where practical it seems clear that combining subjective and objective measures within the same research paradigm is optimal (as discussed in Harvey and Tang, 2012).

Finally, it is possible to infer an effect of sleep fragmentation from group differences, whereby one group's experience is characterised by sleep fragmentation and the others is not (as in Lim, Kowgier, Yu, Buchman, and Bennet, 2013). The logic is straightforward – seek out groups rich in nocturnal awakenings, compare them to healthy sleepers, and attribute memory differences observed between the two groups to the nocturnal awakenings. Of course, it is not that simple, however, as care has to be taken to control for the confounding factors associated with each target group rich in sleep fragmentation. These include age-related health conditions in older adults (Vitello et al., 2004); obesity and cardiovascular disease in Obstructive Sleep Apnoea

Syndrome; mental health conditions in insomnia (Morin et al., 2006); job stress in shift workers (Akerstedt, 2003); iron deficiency in Restless Legs Syndrome (Allen et al., 2003); and diuretic medications in individuals with urinary tract problems (Tikkinen et al., 2010).

Overall, this review takes a multimethod stance and advocates for investigating the relationship between sleep fragmentation and declarative memory from a variety of contrasting and complementary perspectives, both within and across studies (see Brewer and Hunter, 2006, for discussion of multimethod research). It seems clear that polysomnography is at the forefront of better understanding topics such as declarative memory consolidation since it gives researchers insight into the oscillatory patterns thought to be at the heart of this phenomena. Self-report tools can complement and enrich this picture by offering real-world insight at the behaviour level. The following section considers individual difference groups rich in sleep fragmentation in more detail.

### *Obstructive Sleep Apnoea Syndrome*

Sleep fragmentation is a problem for individuals who experience sleep disordered breathing (Carreras et al., 2014). For example, obstructive Sleep Apnea Syndrome is characterized by “complete upper airway occlusion (absent airflow, tongue falling backwards) in the face of continued activity of inspiratory thoracic pump muscles” (Jahaveri et al., 2017). This pattern of activity is further associated with pathologies including intermittent hypoxia (persistent bouts of low oxygen, e.g., Navarette-Opazo and Mitchell, 2014) and sleep fragmentation (Ahuja, Chen, Korey, Pettibone, Osorio, and Varga, 2018; Bucks, Olaithe, Rosenweig, and Morrell, 2017). SHY predicts that the repeated interruptions to sleep experienced in OSAS will impair encoding, and systems consolidation predicts that repeated interruptions in slow wave sleep will impair consolidation.

OSAS occurs with varying degrees of severity (Mild-Moderate-Severe), determined by what is known as the Apnea-Hypopnea Index (the AHI reflects “the sum of all the apneas and hypopneas divided by the total sleep time in hours”; Ahuja et al., 2018). Untreated OSAS is associated with sleepiness, fatigue, depressed mood,

impaired memory, and/or poor concentration (Dempsey et. al., 2010). The key points to take from the available epidemiological literature on OSAS are that: (1) it is highly prevalent disorder (for example, Bucks and colleagues estimated in 2017 that 10% of men and 3% of women aged 30-49 years are diagnosed, rising to 17% and 9% respectively for 50-70-year-olds); (2) rates are increasing significantly around the world (at least, in part, owing to global increases in obesity and lifespan, Benjafield, Ayas, Eastwood, Heinzer, Mary, Morrell, Nunez, Patel, Penzel, Pepin, Peppard, Sinha, Tufik, Valentine, and Malhorta, 2019); and (3) it is widely agreed to be highly underdiagnosed (Bucks et. al., 2017; Abrishami, Khajehdehi, and Chung, 2010; Chen, Wang, Zee, Lutsey, Jahaveri, Alcantara, Jackson, Williams, and Redline, 2014).

Despite a clearer case being made in the literature for procedural memory deficits in OSAS adults (Landry, Anderson, Andrewartha, Sasse, and Conduit, 2014; Csabi, Vargezi-shulz, Janacsek, Malecek, and Nemeth, 2014; Medeiros, Carvalhedo de Bruin, Ponte e Silva, Coutinho, and Sales de Bruin, 2012; Kloepfer, Riemann, Nofzinger, Feige, Unterrainer, O'Hara, Sorichter, and Nissen, 2009), the available evidence for the potential declarative memory deficits in OSAS adults is less clear. The strongest evidence for impaired declarative memory amongst untreated OSAS individuals compared to healthy controls is from a meta-analysis conducted by Wallace and Bucks (2013) in which 42 studies (*n* = 2294 adults with OSA vs *n* = 1364 matched controls) including an episodic memory test were analyzed. The authors reported significant negative effect sizes for immediate verbal recall, delayed verbal recall, and visuo-spatial learning (with visuo-spatial learning, e.g., object-location tasks having the largest negative effect. Specific and notable reports of declarative memory deficits in OSAS samples include work conducted by Kloepfer and colleagues (2009). In this matched-groups (IQ, Sex, Education) study ( *n* =35, Mean Age = 46.4, SD = 5.9 yrs.) participants were asked to remember visual components of a map and contextual details of a building in which significantly reduced verbal retention was found (around 8% after a retention period which included sleep) compared to healthy controls. Further examples include a study carried out by Kheirandish-Gozal, de Long, Spruyt, Chamleua, and Gozal (2010), who trained children with OSAS on a visual memory task and found impaired encoding and next-day retrieval in this group relative to healthy controls. These findings broadly give

support to the hypothesis that declarative memory deficits are observable at the group level in OSAS patients under certain experimental conditions.

There have also, however, been notable failures to find declarative memory deficits in OSAS participants as compared to healthy controls. Among these include similar performance between OSAS and control groups on a verbal paired-associates test (Weschler Memory Scale, N=30 [20 OSA vs 10 controls matched on just age and gender], Mean Age = 57.9(5.8) yrs.; Medeiros et al., 2012), and similarly on a visual retention task (Kloepfer et al., 2009 – it should be noted however that there was a numerical difference of 8% in favor of the control group for this task). Taken together, the evidence for declarative memory deficits in OSAS is less consistent and therefore less persuasive than for non-declarative memory. It is possible that the lack of effect in some studies is better explained by methodological issues rather than the lack of a measurable effect. For example, in the Medeiros study, the sample was around 11 years older than the Kloepfer study (leading to the suggestion that perhaps the deleterious effects of advanced age can under some circumstances cancel out those of OSAS), the study was only matched on age and gender (and not, for example, education and BMI). One possibility is that declarative memory is more vulnerable to task interference than procedural memory (since most investigations which investigate the impact of OSAS on declarative memory take a cognitive battery approach). Most OSAS studies employ neuropsychological assessment in the form of a battery of tests. See Brown & Robertson, 2007, for evidence that consolidation of a primary task can be disrupted by immediately performing a second task, and that declarative and procedural tasks can interfere with each other. This line of thinking lends support to the call by Ahuja and colleagues (2018) for the field to adopt experimental paradigms which employ tasks previously and reliably used to research overnight declarative memory consolidation. One such suitable task, described above, is Rudoy and colleagues' object location task, since this task has been used with great success in the pure sleep and memory consolidation literature and since the task's core domain (visuo-spatial learning) has been highlighted in Wallace and Buck's (2013) and having the largest known negative effects in the sleep apnea literature.

### *Restless Legs Syndrome*

Restless legs syndrome (RLS) has been described as “the most common condition one has never heard of” (Natarajan, 2010), in reference to the condition not being well understood in terms of pathophysiology or symptomatology. For current purposes, however, Restless Legs Syndrome is well-situated towards helping better understand the influence of sleep fragmentation on memory consolidation. This is because sleep fragmentation is a key feature of RLS, alongside other sleep disturbances (including impaired sleep latency, insomnia, and poor sleep duration, Allen, Walters, and Montplaisir, 2005), with polysomnography studies demonstrating a higher arousal index among individuals diagnosed with RLS (Winkleman, Redline, and Baldwin, 2009). Like with OSAS, repeated interruptions to sleep in RLS ought to impair both synaptic downscaling (SHY, encoding) and systems consolidation.

RLS (or Willis-Ekbom-Disease as it is also known), is “a common neurological sensorimotor disorder which manifests in an irresistible urge to move the body to relieve uncomfortable sensations” (Guo, Huang, Jiang, Han, Li, Xu, Zhang, Lin, Xiong, and Wang, 2017). The pathogenesis remains debated, with O’Regan and Anderson suggesting that RLS should be viewed as “a multi-transmitter neurochemical disorder resulting in enhanced excitability and decreased inhibition” (2020, with the authors implicating dopamine dysfunction, deficits in brain iron metabolism and thalamic glutamate levels specifically). RLS has a prevalence rate of around 7-10% in Caucasians (but note far lower incidence among Asian populations; Ohayon, O’hara, and Vitiello, 2012), a wildly varying age of onset (across the lifespan, with the majority of clinical patients being at least middle-aged, Walters, Hickey, Maltzman, Verrico, Joseph, and Hening, 1996), and can be classified as primary (idiopathic with unknown cause, Tison, Crochard, Leger, Bouee, Lainey, and Hasnaoui, 2005), or secondary (presenting alongside one of several neurological disorders, iron deficiency, pregnancy, or chronic renal failure, e.g., Srivanitchapoom, Pandey, and Hallet, 2014).

A particularly relevant facet of the condition is that 80-90% of individuals with RLS also exhibit periodic limb movements during nocturnal sleep (PLMs; Natarajan, 2010; Montplaisir, Nicolas, and Denesle, 1997). Periodic Limb movements (PLMs) are “involuntary movements of the patient’s limb or torso during sleep, different from the

voluntary movement of the limb to relieve the discomfort in RLS patients during wake" (Hening, Walters, and Allen, 2004; Trenkwalder, Stiansy, and Pollmacher, 2004). It is important to note that both the sensory and motor aspects of the condition are closely tied to the circadian rhythm of the individual, consistently peaking at around the same time each night (with spontaneous PLMs peaking between midnight and 3 AM) and are associated with EEG or overt arousals (Natarajan, 2010). This is important because the sleep fragmentation profile of RLS therefore nicely complements that of OSAS – they are both chronic and frequent (multiple arousal-inducing events per hour of symptomatic activity) yet peak differentially in their disruption of the sleep cycle. More specifically, sleep disruption from PLMs in RLS occurs when SWS typically dominates (in an average adult) , whereas the highest rate of arousals in OSAS tends to occur in REM, the bulk of which occurs in the latter half of a typical adults' night's sleep (Alzoubaidi and Mokhlesi, 2017).

Up until very recently, research into the influence of RLS on declarative memory, or even cognition, has been sparse. Jung (2015) reviewed the cognitive profiles of adults with RLS, identifying 10 relevant studies. Half of these studies found impairments in RLS, namely impaired in attention (trail-making test), executive function (Weschler Adult Intelligence Test), and mental flexibility (Wisconsin Card Sorting Task; Pearson, Allen, Dean, Gamaldo, Lesage, and Earlet, 2006; Driver-Dunckley, Connor, and Hentz, 2009; Galbiati, Marello, Giora, Zuccomi, Oldani, and Ferini-Strambi, 2015; Lee, Ramsay, Spira, Vachon, Allen, and Munro, 2014; Rist, Elbaz, Dufoiul, Tzourio, and Kurth, 2015); three studies found no difference between RLS and healthy control (Cognitive set shifting [time to completion], Galbaiti et al., 2015; Mini Mental State Examination, Rist et al., 2015; Stroop Task, Porteus Maze, and Trail-making Task; Gamaldo, Benbrook, Allen, Oguntimein, and Earley, 2008), and two studies surprisingly reported better verbal in RLS patients compared to controls (Kim et al., 2014; Lee et al., 2014). However, the results in these two studies appear to be attributable to participants taking dopaminergic medication, which have previously been found to have enhancing effects on cognition; Moon, Song, Lee, Koo, Lee, and Jung, 2014; Allen, Picchietti, and Hening, 2003). The inconsistencies across studies were attributed to differences in sample characteristics (middle-aged vs old age), task choice (battery vs specific), and medication effects (medicated vs unmedicated). Notably, only one study revealed memory deficits (visuospatial, verbal,

and working memory; Galbiati et al., 2015, in Jung, 2015). Despite the clear promise of this sample population, and the strong theoretical prediction that RLS patients may exhibit impaired declarative memory encoding and consolidation (since their sleep is fragmented, e.g., Feld and Born, 2017, review), little research has been carried out on the topic.

Recently, however, Cha and colleagues (2020) informed this question by comparing sleep parameters among RLS and control using PSG and found differences in SO-spindle coupling. Specifically, the authors found reductions in spindle density, spindle power, SO_spindle coupling, dispersed and delayed spindle phase, and an increase in SO duration. These findings are highly relevant because they demonstrate that the key sleep parameters (spindles, SOs, SO_spindle coupling) implicated in systems consolidation are disrupted in RLS. This leads to the hypothesis that declarative memory effects (relating to both encoding and consolidation) may exist in this population, some of which may be independently attributable to the influence of sleep fragmentation.

### *Maternal Postpartum Sleep Fragmentation*

New mothers also commonly experience fragmented sleep (Gay, Lee, and Lee, 2004; Hunter, Rychnovsky, and Yount, 2009) and changes in sleep architecture (Driver, and Shapiro, 1992; Nishihara, Horiuchi, Eto, Uchida, and Honda, 2004), alongside complaints of “foggy” memory, “baby brain”, or “momnesia” (see Brown and Schaffir, 2019, for a recent review). With little studies available which have looked at postpartum sleep fragmentation and encoding or consolidation more specifically, this section will outline the available evidence that cognition is impaired among postpartum parents, that this likely too applies to declarative memory more specifically given SHY and systems consolidation models, and that sleep fragmentation partially accounts for these deficits.

The first thing to note is that sleep fragmentation is much different among new mothers than in, say, RLS. In general, sleep fragmentation tends to be infrequent (fewer awakenings, on average 2.9 per night) but with much longer awakenings (on average 34 minutes per awakening; McBean, and Montgomery-Downs, 2014). There

is also large inter-individual variability in sleep quality during the postpartum period (due to factors including but not confined to breastfeeding; Gay et. al., 2004; resilience; McBean, Kinsey, and Montgomery-Downs, 2016; number of children; Richter, Kramer, Tang, Montgomery-Downs, and Lemola, 2019. Most women experience reduced sleep quality immediately after birth (Baratte-Beebe, and Lee, 1999; Montgomery-Downs, Insana, Clegg-Krayanok, and Mancini, 2010; Siversten, Hysing, Dorheim, and Eberhard-Gran, 2015). Siversten and colleagues (2015) surveyed 1480 women at week 8 and year 2 postpartum and found persistently elevated levels of insomnia across timepoints (60% at week 8 and 41% at year 2). Kramer and colleagues (2019) also provided longitudinal evidence suggesting that new mothers sleep an average of one hour less each night during the early postpartum period compared with pre-pregnancy and this does not fully recover until around 6 years after birth. Whilst total sleep time in new mothers is relatively stable across the postpartum period, sleep fragmentation peaks in the first 8 weeks (Salvatore, Insana, Stacom, Hawley, and Montgomery-Downs, 2011) and gradually improves for at least the next 4 months, with sleep fragmentation and not total sleep time relating to perceived sleep quality (Creti, Rizzo, Fichten, Bailes, Zelkowitz, and Libman, 2013). It is argued that since the postpartum period is rich in natural variation of sleep fragmentation, and that cognition more broadly is clearly impacted by this, that so too likely will be encoding and consolidation.

Beginning with cognitive deficits in the postpartum period more broadly, it has been reported that many new mothers (75%) experience transient (with at least some recovery by the end of the first year) cognitive difficulties including but not limited to memory, concentration, and reading difficulties (Buckwalter, Buckwalter, Bluestein, and Stancyzk, 2001; Logan, Hill, Jones, Holt-Lunstad, and Larson, 2014). “Baby brain” (or “porridge brain”, “maternal amnesia”, “momnesia” to name but a few; Pawluski, Lambert, and Kinsey, 2016), is widely acknowledged in anecdotal reports from new mothers, and has received some support from the scientific literature. There is an abundance of evidence to suggest that, broadly speaking, cognitive function is impaired in the postpartum period (See Brown and Schaffir, 2019, for review). Evidence of hippocampus-dependent memory impairments more specifically have been found in the animal literature (spatial memory deficits in early postpartum rats, Darnaudery, Perez-Martin, Del Favero, Gomez-Roldan, Garcia-Segura, and Maccari,

2007), as well as findings of impairments in visuospatial memory (Piccardi, Verde, Bianchini, Morgagni, Guariglia, Strollo, and Tomao, 2014), prospective memory (Rendell & Henry, 2008), and explicit conceptual memory (Saraulli et al., 2021) in humans.

Narrowing the scope, some literature exists supporting the proposition that sleep fragmentation specifically in the postpartum period can account for some of the cognitive impairment experienced in this period. For example, when comparing maternal and paternal sleep disruption, it has been suggested that new mothers experience more sleep fragmentation (Insana et al., 2013) and that this is associated with more severe cognitive disruption (with both mothers and fathers exhibiting significantly greater psychomotor vigilance deficits than controls; Insana et al., 2013; Richter et al., 2019). As well as this, new mothers also appear more prone to false memories (Bernt et al., 2014). What seems clear from reviewing the available literature on the topic is that: (1) a significant, if not a majority of women experience some degree of cognitive decline in the postpartum period (Buckwalter et al., 2011); (2) SHY and systems consolidation both strongly predict that encoding and consolidation will also be affected, and (3) sleep fragmentation in this period is likely an important contributor (Janes, Casey, and Huntsdale, 1999; Insana, Williams, Hawley, and Montgomery-Downs, 2013; Insana, Stacom, Hawley, and Montgomery-Downs, 2011).

## Caveats of using clinical samples in declarative memory research

Regarding OSAS, one has to be aware of the potentially confounding influence of vascular changes, neural damage and cell death owing to chronic intermittent hypoxia (Xu, Chi, Row, Xu, Ke, Xu, Luo, Kheirandish, Gozal, and Lui, 2004, in Bucks, Olaithe, Rosenzweig, and Morrell, 2017), as well as the numerous comorbidities associated with the disorder (for example, hypertension; Dempsey et al., 2010; heart disease; Bucks et al., 2017; Stroke; Artz, Young, Finn, Skatrud, and Bradley, 2005; kidney disease, cancer, and diabetes; Gildeh, Drakatos, Higgins, Rosenzweig, and Kent, 2016). Similar issues are faced when researching RLS. Any hypothetical influence of sleep fragmentation on encoding and/or consolidation observed in a sample of restless leg syndrome needs to be evaluated in the context of the confounding influence of dopamine dysfunction (Clemens et al., 2006), its apparent

seasonality (Mercuri, Heogl, and Stefani, 2020) and the numerous examples of potentially noise-generating differential diagnoses and mimic disorders that go along with the disorder (for example, peripheral neuropathy, akathisia, attention deficit hyperactivity disorder; O'Regan and Anderson, 2020). These methodological issues necessitate that researchers make careful sampling decisions when studying either condition – a niche sample with a narrow age band along with the recruitment challenges that presents, versus a more representative sample but an inevitably noisier dataset. There are also other caveats associated with taking a group comparison approach (e.g., reliance upon appropriate control groups; not taking individual differences into account; often small underpowered samples).

Confounds also need to be considered when researching sleep fragmentation in new mothers (who aren't strictly speaking a clinical population but who do experience comparable to those with sleep disorders). For example, it has been proposed that the explanation for memory deficits in the early postpartum period may lie with an interaction between endocrine and plasticity shifts in which the new mother prioritizes cognitive abilities which promote infant care at the expense of those which do not (Pawluski et al., 2016). Supporting this, the animal literature cites numerous examples of apparent cognitive enhancement in the postpartum period in the context of offspring-promoting tasks (for example, spatial learning in rats and monkeys; Kinsey, Blair, Karp, Hester, McNamara, Orthmeyer, et al., 2014).

## Sleep fragmentation and hippocampus-dependent memory in the animal literature

Discussed above in detail, sleep pathologies offer a unique lens to better understand the influence of sleep fragmentation on declarative memory but separating the signal from the noise can at times be challenging. An alternative methodological approach is to seek out or create contexts where comorbidities and/or confounding variables are not present. This section reviews highly relevant findings from the animal literature, which make the case for the independent contribution of sleep fragmentation to hippocampus-dependent memory. Findings from this section will used to set up the

argument that sleep fragmentation independently disrupts declarative encoding and consolidation in humans.

Within the animal literature, sleep fragmentation is much closer to what would be observed in e.g., OSAS, and tends to feature frequent interruptions to sleep (e.g., the orbital shaker method, which interrupts sleep every 90 seconds; Sinton et al., 2009). One such finding is that sleep fragmentation in neonatal rabbits results in abnormalities in cognitive development (poor object-location abilities and slower maze escape time), even when normally structured sleep is restored in the first 8 weeks (using the orbital shaker method for 72 hours in postnatal day 3; Bertranda, Zhanga, and Patela, 2020). This study is a useful starting point in the discussion since it is proof of principle that sleep fragmentation can be disruptive to cognitive processes in the absence of an accompanying pathology. Another recent study from the animal literature was interested in the influence of sleep fragmentation on spatial memory specifically (Kim, Chen, Braden, Williams, Jasso, Garcia, Rho, Bimonte-Nelson, and Maganti, 2015). Kim and colleagues induced acute sleep fragmentation (24h a day for 3 days) in adolescent mice using tactile stimulation and found impaired spatial learning (which is within the declarative umbrella in humans) and synaptic plasticity. Importantly, sleep fragmentation did not impact corticosterone levels, suggesting that the deleterious effect on spatial memory was not attributable to stress (a finding reflected using mice in the sleep deprivation literature, Raven, Heckman, Havekes, and Meerlo, 2019).

A recent major finding from the animal literature (Lui, Meng, Wiggin, Zhang, Rosbash, and Griffith, 2019) has identified a neural circuit responsible for regulating the structure of sleep separate from any mechanisms related to sleep duration. Specifically, the authors showed that in Drosophila Melanogaster (fruit flies), stimulating this circuit of serotonergic neurons causes sleep fragmentation without an impact on sleep duration, and that the resultant disruption is associated with cognitive impairments which can be pharmacologically reversed by restoring healthy sleep. These findings are important for several reasons: (1) they reinforce the large body of literature suggesting sleep fragmentation is associated with disruption of hippocampus-dependent memory (here aversive learning); (2) they represent the first study to offer neurobiological support for the hypothesis that sleep fragmentation and

sleep deprivation are distinct phenomena with divergent regulatory processes which impact upon cognition; and (3) more broadly, these results further support the idea that sleep is something that the brain achieves and regulates each night rather than defaults to. It is not the case that each night the brain no longer wishes to maintain the effort of wakefulness, defaults to a lack of wakefulness, and then capitalizes on the peaceful opportunity to actively facilitate memory consolidation. Rather, these results suggest that sleep is a state the brain needs dedicated circuitry to achieve and regulate after which (or if disrupted) it returns to its default state of wakefulness. These observations, however, are only informative to the extent that fruit flies are comparable to humans. With regards to this issue of generalisability, Lui and colleagues argue that fruit flies have been a mainstay of sleep research for almost two decades (Hendricks, Finn, Packeri, Chavkin, Williams, Sehgal, and Pack, 2000; Shaw, Cirelli, Greenspan, and Tononi, 2000), and that both humans and fruit flies have long consolidated bouts at nighttime and more fragmented sleep during the day (Wiggin, Goodwin, Donelson, Liu, Trinh, Sanyal, and Griffin, 2019; Bonnet and Arand, 2003).

### *Summary*

In sum, there have been numerous findings which suggest that sleep fragmentation influences cognition, leading to the strong prediction that it will independently impair the encoding and consolidation of declarative memories. The available evidence is consistent with the idea that sleep fragmentation and sleep deprivation are distinct phenomena with divergent deleterious properties. As well as this, there now exists evidence that sleep fragmentation disrupts key mechanisms implicated in dominant theories of memory consolidation (for example, SO_spindle coupling). This thesis will attempt to extend these promising findings from the animal literature and separate the influence of sleep fragmentation from sleep deprivation in humans by taking advantage of the rich natural variation in sleep fragmentation in the populations discussed above. The remainder of the discussion will focus on setting up the chapters to come.

## Upcoming Chapters

To bring the content above together, the primary goal of this thesis is to extend the findings from the animal literature to humans, namely which suggest that sleep fragmentation independently contributes to the variance in encoding and consolidation (Lui et al., 2019). This is difficult to test in humans however, since it is a challenge to fragment sleep without disturbing total sleep time. This is because animal research has more invasive means of testing at their disposure. To control for the influence of sleep duration, then, I argue that we will have to embrace the noise and variability of naturalistic sleep fragmentation in groups richly varying in it. Under such a naturalistic observational approach, the independent influence of sleep fragmentation on encoding and consolidation could be quantified using hierarchical multilevel modelling. Within such a paradigm, sleep duration could be allowed to naturally vary in e.g., new parents, or those with RLS, and controlled for statistically by entering it into a hierarchical multilevel regression as a separate step. Naturalistic observational studies like the one I am proposing need to be done at scale, however, given the sampling size requirements of large multilevel models. Therefore, I will need to rely on subjective sleep measurements. To do that, better understanding is needed as to the extent to which subjective and more objective sleep measurements are statistically related to and in agreement with one another. This will be necessary to adequately interpret the findings.

The order of the proceeding chapters is as follows therefore: Chapter 2, among other things, will seek to inform task choice and combination for use in polysomnography studies aiming to test for associations between sleep fragmentation and declarative memory consolidation. Experiment 3 and Experiment 4 will address the primary goal of the thesis and take a pre-sleep/post sleep online memory task approach using hierarchical multilevel modelling to quantify the association between sleep fragmentation and encoding and consolidation among new parents (Experiment 3) and those with RLS (Experiment 4). To facilitate interpretation of these results, Chapter 5 will include a series of mixed effects mega-analyses aimed at better understanding the extent to which subjective and more objective measures of sleep are related to and statistically in agreement with one another. Finally, Chapter 6 will bring the findings together, explore contributions to

the key theoretical models underpinning this thesis, and discuss limitations and future directions.

## Conclusions

The current review sought to evaluate the independent influence of sleep fragmentation on declarative memory. In sum, the literature review leads to several conclusions. There is now good evidence for sleep's active role in the encoding and consolidation of declarative memories. There is also considerable evidence that sleep fragmentation impairs cognition, with some tentative evidence that the resultant cognitive impairment extends to declarative memory specifically. What is less certain is the extent to which sleep fragmentation and sleep deprivation uniquely contribute to the variance in encoding and consolidation, although the available evidence suggests that they may be distinctly disruptive. To help better understand the extent to which this is the case in humans, groups naturally rich in variation in sleep fragmentation are well-situated towards better understanding the extent to which this is the case. Among these groups are those with Obstructive Sleep Apnoea Syndrome, those with Restless Legs Syndrome, and new parents, all which experience sleep fragmentation in distinctive and complementary patterns. To make use of the naturally varying sleep disruption within these populations, however, it will have to done at scale, using subjective sleep measures, and the confounding effect of sleep duration will have to be controlled for statistically taking a hierarchical multilevel modelling approach.

# Chapter 2

# Neighbourhood, interference, and MASC-IT

## Introduction

Obstructive Sleep Apnoea Syndrome (OSAS) involves excessive narrowing of the throat during sleep, obstructed breathing (apnoea) and chronic sleep fragmentation (Wallace and Bucks, 2013). The most common symptoms of OSAS are snoring, stopping breathing episodes whilst asleep and excessive daytime sleepiness. OSAS is a highly prevalent sleep disorder, affecting around 1/25 middle-aged men and 1/50 middle-aged women in the UK (Sleep Apnoea Trust, 2017). As discussed in detail in the previous chapter, converging evidence suggests that OSAS has widespread detrimental effects on sleep, including slow wave sleep (SWS; Dempsey, Veasley, and O'Donnell, 2010), which has been implicated in memory consolidation ("a dynamic, generative, transformative, and lingering process that is posited to balance maintenance of useful experience-dependent internal representations of the world with the need to adapt these representations to the changing world"; Dudai, Karni, and Born, 2015, p.20).

Continuous Positive Airway Pressure (CPAP) therapy is widely regarded as the "gold-standard" for treatment (Rotenberg, Vicini, Pang, and Pang, 2016) and is associated with significant improvements in memory, e.g., working memory; Felder, Bruce, Zimmerman, and Sweet, 2007; and verbal episodic memory; Rozenweig, Glasser, Crum, Kempton, Milsosevic, McMillan, Leschziner, Kumari, Goadsby, Simonds, Williams, and Morrell, 2016), broadly speaking, but not consistently (e.g., Apnea Positive Pressure Long-term Efficacy study; Kushida, Nichols, and Holmes, 2012; Joyeux-Faure, Naegele, Pepin, Tamisier, Levy, and Launois. 2016). Investigating the OSAS population might also give crucial insights into typical human memory, since it gives researchers the opportunity to test how sleep fragmentation affects consolidation. The following chapter seeks to outline a proposal for future research (*MASC-IT*: Memory and Sleep Consolidation Intervention Trial) aimed at better understanding the influence of CPAP (and therefore relief from chronic sleep fragmentation) on declarative memory consolidation. We begin by reviewing the CPAP

literature before presenting findings from two preliminary experiments designed to empirically justify task choice for MASC-IT. Finally, we finish by discussing potential findings from the proposed study and their implications for future research.

## Continuous Positive Airway Pressure (CPAP) Therapy

Since its development in the early 1980s, CPAP therapy has had a profound impact on the treatment of OSAS (O'Sullivan, Issa, Berthon-Jones, and Eves, 1981; See Dempsey et al., 2010 for a detailed discussion of the history of OSAS). CPAP therapy involves the patient wearing either a nasal of oronasal mask through which air is continuously pumped to avoid recurrent and often chronic throat-collapsing episodes. (Eckert et al., 2018). This section will very briefly review the general efficacy of CPAP therapy before considering in greater detail the most influential recent investigations into the influence of CPAP on memory consolidation and sleep architecture.

### *General Efficacy of CPAP Therapy*

The key points to note are that CPAP therapy is: (1) highly effective at restoring healthy nocturnal airway flow in OSAS patients; (2) the current "gold standard" treatment therapy for OSAS; and (3) attitudes towards this treatment and subsequent adherence to it varies considerably (Landry, O'Driscoll, Hamilton, and Conduit 2015; Djonlagic, Guo, Carusona, Matteis, Stickgold, and Malhorta, 2015). Among the many reported positive health outcomes associated with CPAP therapy in OSAS include: successfully reducing the apnea-hypopnea index (Isaa and O'Sullivan, 1986); markedly reducing sympathetic activity and blood pressure (Jahaveri et al, 2016); improving the prognosis of stroke patients (Lyons and Ryan, 2015; Parra, Sanchez-Armengol, and Capote, 2015); reducing the rate of premature ventricular beats during sleep in patients with arrhythmias (Craig, Pepperall, and Kohler, 2009; Ryan, Usui, and Floras, 2005); improving pulmonary hypertension (Jahaveri, Jahaveri, and Jahaveri, 2013); increasing brain volume (Kim et al., 2016), and; improving outcomes in patients with comorbidities such as dementia and epilepsy (Rosenweig et al., 2016). It is also worth noting that CPAP has also been shown have significant positive effects in circumstances where nocturnal supplemental oxygen has not (namely in facilitating

the reduction of blood pressure in patients with cardiovascular disease; Gottlieb, Punjabi, Mehra, Patel, Quan, Babinear, Tracy, Rueschman, Blumenthal, Lewis, Bhattt, and Redline, 2014). However, despite these encouraging findings, investigations comparing aspects of cognition before and after CPAP have yielded mixed results (studies showing significant neuropsychological improvement: Canessa, Castronovo, Cappa, Aloia, Marelli, Falini, Alemanno, and Ferini-Strambi, 2011; Kim et al., 2016; Rosenweig et al., 2016; Landry et al., 2015; studies reporting no effect of CPAP on domains of cognitive functioning: Djonlagic et al., 2015; Jackson, McEvoy, Banks, and Barnes, 2018; Joyeux-Faire, Naegele, Pepin, Tamisier, Levy, and Launois, 2018). The remaining discussion therefore attempts to reconcile the apparent disagreement in the literature as to the influence of CPAP on memory.

## The Influence of CPAP on Memory

### *Studies Reporting Significant Improvements in Memory After CPAP*

Several recent investigations into the influence of CPAP on cognition have highlighted significant improvements after treatment, many of which in association with structural and functional brain alterations. Canessa and colleagues (2011) investigated the relationship between brain morphology changes in OSAS and cognition, and the modifications after treatment, using combined neuropsychological testing and voxel-based morphology. In this study, polysomnography, MRI, and a cognitive battery was carried out on untreated OSAS participants and a corresponding control group, with both being re-tested after three months to assess therapy efficacy. The authors reported pretreatment impairment in the OSA group in most cognitive areas, and in mood and sleepiness, which were associated with focal reduction in gray-matter volume in the left hippocampus, left parietal cortex, and right frontal superior gyrus. After treatment, the authors observed improvements involving short- and long- term memory (verbal and visuospatial), attention, and executive function (digit-span backwards, Stroop and Trail-making test) which related to hippocampal and frontal gray-matter volume gains. Canessa and colleagues suggested here that since the hippocampus is sensitive to hypoxia (Gozal, Row, Schurr, and Gozal, 2001), it may be one of the regions most impacted by intermittent hypoxia. Hippocampal damage because of hypoxic episodes would intuitively therefore result in cognitive

deficits and, importantly, ought to be at least partly reversible due to hippocampal plasticity (Kempermann, Gast, and Gast, 2002). These key findings were also supported by similar experimental work conducted by Kim and colleagues (2016), who also found initial hippocampal atrophy in OSAS patients and a correlation between higher working memory impairment prior to treatment and brain volume increase after treatment (hippocampal and prefrontal).

Recent evidence also exists in the literature to suggest that positive effects of CPAP can be found after only one-month of use. Rosenweig, Glasser, and colleagues (2016) tested (neurocognitive, neuroimaging and polysomnography) 55 moderate to severe OSAS patients newly diagnosed before treatment and one month later. They found hypertrophy after one month in the thalamus, significant reductions in daytime somnolence associated with neuroplastic changes in the brainstem, and improvement in delayed memory scores. These findings reported by Canessa et al., (2011) and Kim et al., (2016), demonstrate that these therapeutic effects can occur in a relatively short timeframe. These findings are further complemented by a recent study significant restoration of cognitive function and white matter over a 12-month period with CPAP in OSAS adults (Castronovo, Scifo, and Castellano, 2014), collectively providing convincing evidence that CPAP can have long-term restorative effects on memory. Finally, positive effects of CPAP have also been found for non-declarative memory in OSAS patients (Landry et al., 2015), suggesting that not only that CPAP can restore memory deficits to some degree, but it can also be efficacious across different memory systems.

### *Studies Failing to Find Significant Improvements in Memory After CPAP*

There are also, however, several studies which did not find positive effects of CPAP use in OSAS patients. This section reviews these negative findings and tries to reconcile them with those discussed above by offering theoretical and methodological explanations as to why positive effects of treatment were not found.

The first study worth considering is the APPLES study (Apnea Positive Pressure Long-term Efficacy study; Kushida, Nichols, and Holmes, 2012). The APPLES study compared CPAP and sham at 2 and 6 months among recently

diagnosed OSAS patients. Participants were asked to complete tests of attention, psychomotor function, verbal and working memory and showed no group differences at either timepoint. Another recent study looked at neurobehavioural function in response to three months of CPAP (OSAS group – N=88; Jackson et al., 2018). Jackson and colleagues reported that OSAS patients fared more poorly than healthy controls on several measures including daytime somnolence, mood, attention, and memory, all of which persisted despite treatment.

There have also been a couple of recent studies which departed from the common approach of having participants complete a large neuropsychological battery of cognitive tests and instead focus their investigation on much fewer measures. The first of these is the recent experimental work of Joyeux-Faure et. al. (2016), who compared CPAP and sham over a period of 6 weeks. This paradigm tested verbal, procedural, and working memory before and after 3 months of CPAP and found no improvement in both groups after 6 weeks of treatment. A final study is one carried out by Djonlagic, Guo, and colleagues (2017), who looked at the impact of one night of CPAP on memory consolidation and attention. Their experimental procedure involved tested 15 healthy controls and 29 OSAS patients sustained attention (Psychomotor Vigilance Task) and procedural memory (motor sequence learning task) in the evening and morning after the first night of CPAP. The authors reported an augmentation of attention after one night of CPAP (faster morning PVT scores) but no memory effect. Taken these findings together, and in the context of several demonstrations of long-term restoration of memory after CPAP use, it is argued that the failure to detect effects of treatment in these studies is best explained using suboptimal methodologies rather than a genuine null effect of treatment. The following sections explore some of the potential methodological and theoretical lessons to be learned from the above review of the literature which will form the methodological basis for the MASC-IT study.

## Limitations and Design Considerations

### *Evening/Morning Designs*

This review agrees with a recent review on the influence of OSAS on memory (Ahuja et al., 2018) that it is suboptimal to investigate the influence of sleep

fragmentation on declarative memory consolidation within a paradigm that only tests memory within a sole period of wakefulness (this neglect of much more sensitive evening-morning designs is largely indicative of the literature so far). Indeed, there are several recent studies the results of which are consistent with this suggestion. For example, Landry and colleagues (2015, discussed above) explicitly acknowledged that they would not have observed the significant effect of treatment they did had they not used an evening/morning paradigm. Specifically, they found comparable daytime performance in motor skill acquisition between OSAS patients and controls, and it was not until re-test in the morning that impairment in the OSAS group became apparent. It is worth noting that some of the major failures to find either impairments of cognition and/or positive benefits of CPAP were in studies which opted to test during daytime periods of normal wakefulness (Joyeux-Faure et al., 2016; Kushida et al., 2012).

Another strong reason for adopting evening/morning designs is that they allow for more rigorous investigations aimed at identifying which key elements of the sleep cycle are associated with impairment/treatment benefit. Few studies are available that attempt to measure specific components of the sleep cycle (e.g., slow oscillations, sleep spindles, sharp-wave ripples) and fewer still make conscious attempts to interpret their results in the context of the active systems model of consolidation (discussed in detail above; see Rasch & Born, 2013 for discussion). Current evidence of associations between key elements of the sleep cycle are spare, but include findings that: OSAS patients (intuitively) spent decreased time in SWS and REM (Bucket al., 2017); performance in the maze test (a measure of non-declarative memory) was correlated with slow wave sleep (N3) in adults with moderate/severe OSAS (Medeiros et al., 2012); sleep spindles during stage 2 sleep are associated with motor sequence learning (a measure of non-declarative memory), with CPAP-treated OSAS patients having higher total numbers of sleep spindles and greater spindle density than treatment-naïve patients (Note that the authors did not carry analyses on any other stages; Djonlagic et al., 2012), and; there is a greater likelihood of airways collapse in REM and apneas and hypopneas are more pronounced and are related with more serious oxygen saturation (Alzoubaidi and Mokhlesi, 2017).

### *Task Choice*

Also in agreement with recent reviews on the influence of OSAS on cognition (Ahuja et al., 2018; Bucks et al., 2017), it is argued that sleep fragmentation paradigms should focus on employing reliable and sensitive memory tasks from the sleep and overnight consolidation literature (i.e., '*sure-thing'* tasks; Rudoy, Voss, Westerberg, and Paller, 2009; Davis & Gaskell, 2009; Gaskell & Dumay, 2003; and Tamminen et al., 2010; for consistent successful detection of significant sleep effects using object location and word learning paradigms respectively). This is not to say, that the above studies used "bad" tasks just that they left a lot to chance by not choosing tasks well validated for use in consolidation research, specifically. As well as this, most studies in the field tend to adopt the pragmatic strategy of having participants complete a battery of neuropsychological tests. However, this approach could be harming the sensitivity of each individual task due to factors such as task fatigue and interference (Brown and Robertson, 2007, discussed above). Future investigations might wish to rigorously test the effects of task interference in healthy participants, perhaps by assigning one group of participants to an evening/morning paradigm testing one memory test, having another group follow the same procedure using a different memory task, and finally comparing these to the performance of a third group tasked with completing a counterbalanced combination of the two tasks under the same protocol.

### *Misconceptions Regarding Effective CPAP Use in the Literature*

Finally, there also seems to be a misconception of what constitutes effective use of CPAP in the literature, at least with respect to what constitutes effective use of CPAP in studies aimed at better understanding the effects of treatment on memory in OSAS. Most available research in the CPAP and memory field does not report data on how compliant to treatment their sample was. As well as this, of the studies that comment on treatment compliance, some underestimate what the research suggests ought to constitute adequate CPAP use for OSAS and memory studies. This issue appears to stem from 3 influential articles which discussed CPAP compliance rates in adults in the mid-1990s (Kribbs, Pack, Kline, Smith, Schwartz, and Schubert, 1993; Engleman, Martin, and Douglas, 1994; Reeves-Hoche, Meck, and Zwillich, 1994)

which suggested that on average CPAP was typically used for 4.7 hours on any given night. Importantly, no claim was made by the authors that 4.7 hours constituted healthy or sufficient usage, yet somehow this figure emerged as a threshold for good CPAP compliance in the literature (See Sawyer, Gooneratne, Marcus, Ofer, Richards, and Weaver, 2001, for a recent systematic review on the topic). This is problematic since several studies (for example: Antic, Catcheside, Buchan, Hensley, Naughton, and Rowland, 2011; Weaver, Maislin, Dinges, Bloxham, George, and Greenberg, 2006) suggest that CPAP usage positively correlates with better outcomes. This is an even more salient issue regarding memory research, since memory seems to be a cognitive domain reliant on high doses of CPAP (one study found that likelihood of significant memory restoration was 8 times more likely when comparing 6 and 2 hours of CPAP/night respectively: Zimmerman, Arnedt, Stanchina, Millman, and Aloia, 2006). These findings might explain the null result reported by Jackson and colleagues (2018, discussed in detail above), who found no benefit of CPAP on memory after 3 months in OSAS patients. In this study, the authors reported that 43.1% of their sample used CPAP for at least 4/hours a night for 70% of nights and went on to conclude that significant decrements in memory persisted despite adequate use of “gold standard” treatment. Reported levels of treatment compliance in this study are well at odds with the evidence reviewed by Sawyer and colleagues (2011) and may explain the lack of effect of treatment in this investigation, and indeed many others.

In sum, CPAP therapy is highly efficacious and is associated with a wide range of positive long-term benefits. With regards to memory specifically, results are less consistent. Studies have found a long-term, positive benefit for CPAP in both declarative and non-declarative domains. However, notable failures to find a positive benefit of CPAP on memory in OSAS patients exist too. It is argued that these failures could possibly be explained by methodological issues as opposed to the lack of a measurable effect of treatment and that this should be tested by adapting what we argue are more optimal testing conditions. The field would benefit from increased use of evening/morning designs and more specific task choices, as well as more consistent reporting of treatment compliance and better understanding of what constitutes adequate CPAP use in studies aimed at assessing memory in OSAS patients. Having considered some key design suggestions which might be of use to researchers looking

to investigate the influence of CPAP on memory consolidation in adults with OSAS, the rest of the review turns to some of the key open questions in the field.

## The Current Investigation

The MASC-IT study seeks to add some clarification of the available literature on the influence of CPAP on declarative memory consolidation by applying the above methodological suggestions longitudinally to a sample of treatment-naïve adults with a recent diagnosis of OSAS. Given the scale of the project, and the sensitivity of consolidation and treatment effects to task choice, we carried out two methodological studies aimed at empirically establishing which tasks are suitable for inclusion in the MASC-IT study and whether they will interfere with each other, which will now be discussed. It is worth noting that we live in extraordinary times (owing to the ongoing COVID-19 pandemic), ones in which the MASC-IT study could not possibly be carried out, and as such this chapter builds up the presentation of a plan to collect data for the study later. It is our opinion that this is still a worthwhile exercise, since we are of the belief that the research questions underlying the study have theoretical, economic, and societal value, and it hoped that someone attempts to answer them as soon as possible, even if they are not answered by us.

The following two experiments were conducted to establish which type of stimuli are best suited towards enabling the detection of treatment effects in the MASC-IT study if they exist. However, the study also attempts to answer relevant questions to the psycholinguistics (Experiment 1) and interference (Exp 2) literature, respectively, and they are written up as if for publication. The logic is as follows:(1) OSAS is a disorder characterised by fragmented sleep; (2) Sleep is intimately involved in declarative memory consolidation; (3) If declarative memory consolidation is observed to be impaired in this population, then it is hypothesised that this will be associated with sleep (e.g., time spent in SWS); (4) CPAP is effective at removing the disruption to sleep in OSAS patients; (5) There it is hypothesised that a restoration of continuous sleep will result in a degree of renormalisation of the consolidation of declarative information; and (6) The type of stimuli best suited to test these hypotheses are ones which rely on sleep heavily to learn, and which do not interfere with each other.

# Experiment 1

Following on from the logic flow above, this experiment was conducted under the assumption that the learning of the type of novel word stimuli which benefited the most from a delay period which included a full night of nocturnal sleep (i.e., the one which showed the greatest overnight improvement in recall) would later be used in the MASC-IT study.

## Introduction

Lexical integration is best understood by applying the Complementary Learning Systems (CLS; McClelland, McNaughton, and O'Reilly, 1995) to word learning. According to the CLS model, there are two memory systems – a hippocampal system in which new memories are rapidly learned and stored in the short-term in isolation from the individual's existing knowledge base, and; a neocortical system in which memory representations are gradually integrated with the lexicon (Davis and Gaskell, 2009; see James, Gaskell, Weighall, and Henderson, 2017, for review).

A robust literature of behavioural (e.g., Dumay and Gaskell, 2012; Tamminen, Payne, Stickgold, Walmsley, and Gaskell, 2010), and neuroimaging (e.g., Takashima, Bakker, van Hell, Janzen, and McQueen, 2014, 2017) research implicates sleep in word learning (specifically, by facilitating the integration of new words into the lexicon; Ellenbogen, Hu, Payne, Titone, and Walker, 2007). For example, Tamminen and colleagues (2010) employed a spoken novel word paradigm and found a within-subjects effect of improved recall between test and re-test, but only when the delay period included nocturnal sleep. Overnight spoken novel word learning, therefore, is a consistent, well understood, and powerful paradigm for use in wider sleep and memory research not necessarily focussed on word learning.

An open question within the spoken word learning literature, however, is the degree to which neighbourhood (the degree of similarity between the novel word and previously established entries in the lexicon) matters. Previous research has found that existing knowledge influences the expression of consolidation. As such, adults have greater amounts of existing knowledge than children and this supports fast

consolidation of new information (Groch, McMakin, Guggenbühl, Rasch, Huber, and Wilhelm, 2012; Wilhelm, Diekelmann, and Born, 2008; Wilhelm, Prehn-Kristensen, and Born, 2012; in James et al., 2017).As well as this, information is consolidated more rapidly when the information to be learned is consistent with an already- existing schema (for example, Tse, Langston, Kakeyama, Bethus, Spooner, Wood, and Morris, 2007, found that rats consolidated food location information quicker when consistent with a previously established mental map of their environment; and Lewis and Durrant, 2011, found that human participants consolidated schema-consistent melodies better than schema-inconsistent melodies). Finally, there is some evidence that close competitors are preferentially consolidated in children but not adults (James, Gaskell, and Henderson, 2018).

The existing data can be interpreted in a couple of competing ways. Firstly, it could be the case that when adults are tasked with learning spoken words which are closely related to established entries in the lexicon, they will be encoded easily and rapidly consolidated overnight, and that the brain's preference for information consistent with prior knowledge will lead to a greater degree of overnight improvement (consolidation effect) relative to distantly related words. Alternatively, it could be the case since substantial links already exist between the new and existing knowledge that there will be less of a need for an active role of sleep (see Rasch and Born, 2013 for a comprehensive review). This would result in a general advantage for close neighbours (higher encoding and next-day retrieval) but less of a consolidation effect (overnight improvement in recall) than would be seen for the overnight learning of distant neighbours.

The current investigation seeks to address this question by directly comparing overnight consolidation between groups of adults tasked with learning close neighbours (e.g., *cathedruke-cathedral*) and those tasked with learning distant neighbours (e.g., *yothedruke-cathedruke-cathedral*). The purpose of this exploratory investigation is to establish which theoretical stance is viable and is one which is in the interest of researchers focussed on better understanding lexical integration as well as those looking to establish powerful and fine-tuned paradigms towards better understanding memory consolidation.

## Methods

### *Participants*

A total of 40 undergraduate students from the University of York with English as their first language completed this study and were recruited using the university's online recruitment system (*Task; Close Neighbours,* N=20*; Distant Neighbours,* N=20). Participants were opportunistically assigned (the first participant to sign up was assigned to group the Close Neighbours group, the second to the Distant Neighbours group, and so on) into two groups which were well-matched for age (years; *Close Neighbours*: Range= 18-23, Mean=19.75, SD= 1.21; *Distant Neighbours*: Range=18-23, M=19.4, SD=1.41), and sex (*Close Neighbours*: 17 females [85 %], 3 males [15%]; *Distant Neighbours*: 17 females [85%], 3 males [15%]). Ethical approval was granted by the Research Ethics Committee of the Department of Psychology, University of York. Participants gave informed consent, were reimbursed for their time, and data was collected and stored in accordance with the General Data Protection Regulation (GDPR). Participants were screened such that only non-smokers, not taking any psychoactive medication, who did not have a known sleep disorder and who reported no history of alcohol or drug abuse, were invited to take part. All participants had normal or corrected-to-normal vision and hearing. Participants were also instructed not to consume alcohol in the 24 hours before and to not consume caffeine on the days of the experiment, which were confirmed by self-report on arrival.

### *Materials*

**Alertness Materials.** Alertness at the beginning of each session was measured subjectively using the Stanford Sleepiness Scale (SSS; Hoddes, Zarcone, Smythe, Phillips, and Dement, 1973). This is a one-item measure of sleepiness in which participants indicate how sleepy they feel on a scale of 0 (*Feeling active, vital, alert, or wide awake) to 7(No longer fighting sleep, sleep onset soon; having dream-like thoughts*). Alertness was also measured objectively at the beginning of each session using the Psychomotor Vigilance Task (PVT),in which participants are exposed to a red counter against a black background on screen for each trial (described in further detail below; the task can also be downloaded for free using the

following link:
https://pcpvt.bhsai.org).

**Lexical Materials.** The lexical materials used for this experiment were either directly taken or adapted from previous investigations into the consolidation of newly acquired novel words (Gaskell & Dumay, 2003; Davis & Gaskell 2009; Tamminen et al., 2010). See appendices for the 40 words used (English base words, N=20[not shown to participants]; Close neighbours N=20; Distant neighbours, N=20). The closely related novel words (e.g., *Cathedral-Cathedruke*) were separated by changing the last two phonemes. The less-related novel words were further separate from their English root words by altering the first two phonemes of the closely related novel words (e.g., *cathedral-cathedruke-yothedruke*). Each word was delivered by a recording of a female native British English speaker.

We made use of the psycholinguistic research tool N-Watch (Davis, 2006), which gives various neighbourhood statistics. The program calculates orthographic similarity against a dictionary of 30,605 words. For the purposes of this study, we imported our word list into N-Watch and selected all tests which involved some aspect of neighbourhood. If the output suggested that any of our stimuli were closely related to another word in the default dictionary, then the word was changed or removed from the list. Our final word lists were also perfectly matched for syllable length and letter length since one list (Distant Neighbours) was adapted from the other (Close Neighbours; N=20/Group; Range 7-10 letters/word; Mean=8.35 letters/word; SD=0.88 letters; Mean No. Syllables = 2.9) and care was taken not to expose participants to more than one word with the same word onset.

**Questionnaires.** ***Pittsburgh Sleep Quality Index (PSQI).*** Overall sleep quality was measured using the PSQI, which calculates a global sleep quality score from 7 sections (Total Sleep Time; Sleep Disturbance; Sleep Latency; Daytime Dysfunction; Sleep Efficiency; Overall Sleep Quality; and Sleep Medication Use), and is scored on a scale of 0-21 (greater score = poorer sleep quality; see Zhong, Gelave, Sanchez, and Williams, 2015, for a recent discussion of the psychometric properties of the PSQI).

***Epworth Sleepiness Scale (ESS).*** Daytime sleepiness was measured using the ESS. The ESS is a 8-item measure (each item is a typical daytime scenario (e.g., watching TV, sitting and reading) scored on the likelihood of falling asleep whilst carrying out each activity (*0- would never fall asleep in that situation; 1-slight chance of falling asleep; 2-medium chance of falling asleep; high chance of falling asleep*).

**Apparatus.** All experimental tasks were carried out in individual testing booths using HP EliteDesk 800 G2 TWR computers, with IIYAMA ProLiteX238OHS monitors with a screen resolution of 1920 x 1080. DT234PRO headphones were used to present auditory stimuli and both brightness and volume were kept at the same moderate, consistent level.

**Stimuli presentation.** Stimuli for the Psychomotor Vigilance Task (PVT) were presented using MATLAB (Mathworks, 2012). Stimuli for the two versions of the novel word-learning tasks were presented using Open Sesame (Mathôt, Schreij, and Theeuwes, 2012).

## Design

A 2 (*Neighbourhood:* Close neighbours, Distant neighbours) x2 (*Session:* Session 1, Session 2) mixed factorial ANOVA was used to determine which version of this verbal declarative memory task would be incorporated into planned future investigations into sleep and memory in both university and clinical samples. The between-subjects independent variable *Neighbourhood* was characterised by the type of stimuli participants were exposed to  Participants either completed a novel word-learning task which has been used in several previous studies (Gaskell & Dumay, 2003; Davis & Gaskell 2009; Tamminen et al., 2010) in which participants are instructed to learn novel-words which are closely-related to already existing English words (e.g., *Cathedral-Cathedruke*), or an adapted version of this task with the only key difference being that the distance between the novel word and the existing English word are further separated by two phonemes at the beginning of each work (e.g., *Cathedral-Cathedruke-Yothedruke*). The within-subjects variable *Session* and had two levels (Session 1[Evening]/Session 2[Morning]).

## Procedure

The experiment was based around a 12-hour protocol whereby Session 1 took place in the evening between 8 and 10pm and Session 2 took place the following morning between 8 and 10am. Participants spent the delay period between sessions going about their typical routine in which they expected and confirmed that they had gone home and had a full night's sleep. All testing took place in isolated test booths in the Psychology department at the University of York.

### *Session 1 (Evening)*

Upon arrival participants filled out consent forms, re-confirmed that they met the relevant inclusion criteria and completed a sleep habits questionnaire which captured relevant information such as each participants' typical bed and wake times, subjective ratings of sleep quality and napping habits. Alertness for Session 1 was measured both by self-report (SSS) and then by completion of the PVT. Participants then completed a learning phase and a test phase for the novel word-learning task relevant to that group. All experimental tasks were conducted on computers in isolated individual testing booths. Stimuli were presented using headphones for the spoken stimuli. Responses were recorded by mouse-click for the PVT. For the word learning tasks, responses were recorded by keyboard stroke for the learning phase (phoneme monitoring), and for the test phase (cued recall) participants made verbal responses which were recorded, transcribed by the experimenter, and coded as being correct or incorrect. Screen brightness and speaker volume were kept the same across the sample.

### *Session 2 (Morning)*

Upon arrival for Session 2, participants filled out the SSS, completed the Psychomotor Vigilance Task and finally completed a re-test phase for the novel word-learning task relevant to that group. Participants were then free to go and were given a comprehensive debrief sheet for reference as well as contact details for the tester and the lead supervisor and were invited to ask any questions they had about the study.

***The tasks:***

**Psychomotor Vigilance Task (PVT).** The Psychomotor Vigilance Task is an objective measure of alertness (Reifman, Kumar, Khitrov, Liu, Sridhar, Ramakrishnan, et al., 2018) designed to complement our measure of subjective sleepiness (SSS). This short task (length=5 mins) involves participants responding as quickly and as accurately as possible to a counter on the computer screen which starts counting in milliseconds at the beginning of each trial. The timing between trials was randomised between limits of 2s and 10s to ensure sustained attention.

**Novel Word-Learning Task (Training).** The novel word-learning task used in this study has been used in several studies (Gaskell & Dumay, 2003; Davis & Gaskell 2009; Tamminen et al., 2010). Participants were trained to learn 20 novel words (e.g., *cathedruke*) that have been derived from existing English words (*cathedral*). Participants were trained over a 15-minute period in which they heard the target words 12 times each and answered questions such as "*does that word have a 'c' sound?"* (Phoneme monitoring). Specifically, participants would have a round of exposure in which they heard each of the 20 words one at a time and then a round of phoneme monitoring in which they were asked one question for each of the words they had just heard. Participants responded to phoneme monitoring by saying the word out loud (responses were recorded) and by indicating yes or no via keyboard stroke. See Table 1 (below) for a breakdown of the structure of the blocks of training each participant engaged with.

**Novel Word-Learning Task (Test).** At test (cued recall), the first syllable of each novel word was presented one at a time and participants were given 10s to recall the novel word. Responses were made in the presence of the experimenter and were verbal. Auditory responses were recorded by the computer and transcribed by the experimenter, who also coded them as being correct or incorrect. The only difference between the Close neighbour group and the Distant neighbour group was that the distance between the novel word and the existing English word are further separated by two phonemes at the beginning of each work (e.g., *Cathedral-Cathedruke-Yothedruke*).

## Results

### *Sleep habits*

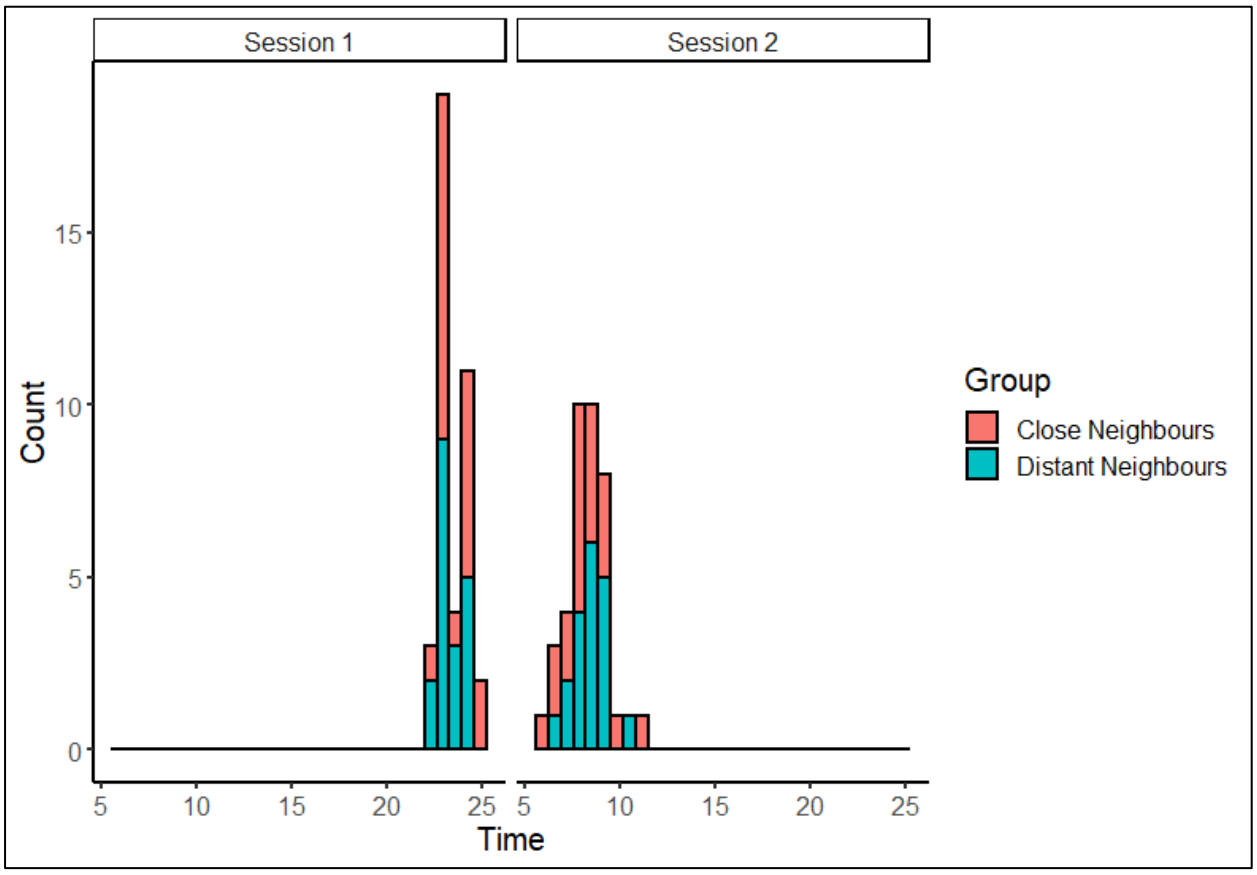



*Figure 1. Stacked histogram showing bedtimes (left) and wake times (right) by Group (Close Neighbours / Distant Neighbours) and Session (Session 1/ Session 2). Time is in hours, e.g., 25 = 1AM.*

Participant self-reports revealed that both groups (*Close Neighbours/ Distant Neighbours*) were well-matched for hours slept (mins), sleep latency (mins), sleep quality (PSQI), daytime somnolence (ESS), sleep onset and wake times. See Figure 1 above for distributions of subjectively estimated sleep onset and wake times for each group. Assumptions were met and independent t-tests revealed no significant differences between the groups for Total Sleep Time, Sleep Onset Latency, sleep quality (PSQI), daytime somnolence (ESS), sleep onset or wake up time (see Tables 1 and 2 below).

*Table 1. Means and standard deviations for self-reported sleep measures by group and at sample level.*

| Variable | Close Neighbours | Distant Neighbours | Total |
|---|---|---|---|
| Total Sleep Time | 487.2 (76.2) | 502.8 (51.0) | 494.9 (64.9) |
| Sleep Onset Latency | 25.3 (18.2) | 27.6 (10.9) | 26.4 (15.0) |
| PSQI Global | 5.1 (1.9) | 5.3 (2.0) | 5.2 (1.9) |
| ESS | 5.8 (3.0) | 5.9 (2.7) | 5.9 (2.8) |
| Sleep Onset | 23:33 (47) | 23:16 (37) | 23:24 (42) |
| Wake | 08:12 (68) | 08:24 (53) | 08:18 (61) |

*Note.* Total Sleep Time (mins); Sleep Onset Latency (mins); PSQI = Pittsburgh Sleep Quality Index; ESS = Epworth Sleepiness Scale. SD for Sleep Onset and Wake (mins).

*Table 2. Results for a series of independent t-tests conducted to test for differences between groups (Close Neighbours / Distant Neighbours) on subjectively measured sleep parameters.*

| Variable | *t* | *p* | *d* |
|---|---|---|---|
| Total Sleep Time | -0.75 | .46 | .24 |
| Sleep Onset Latency | -0.48 | .64 | .15 |
| PSQI Global | -0.40 | .69 | .13 |
| ESS | -0.11 | .91 | .04 |
| Sleep Onset | 1.08 | .29 | .28 |
| Wake | -0.63 | .53 | .20 |

### *Stanford Sleepiness Scale*

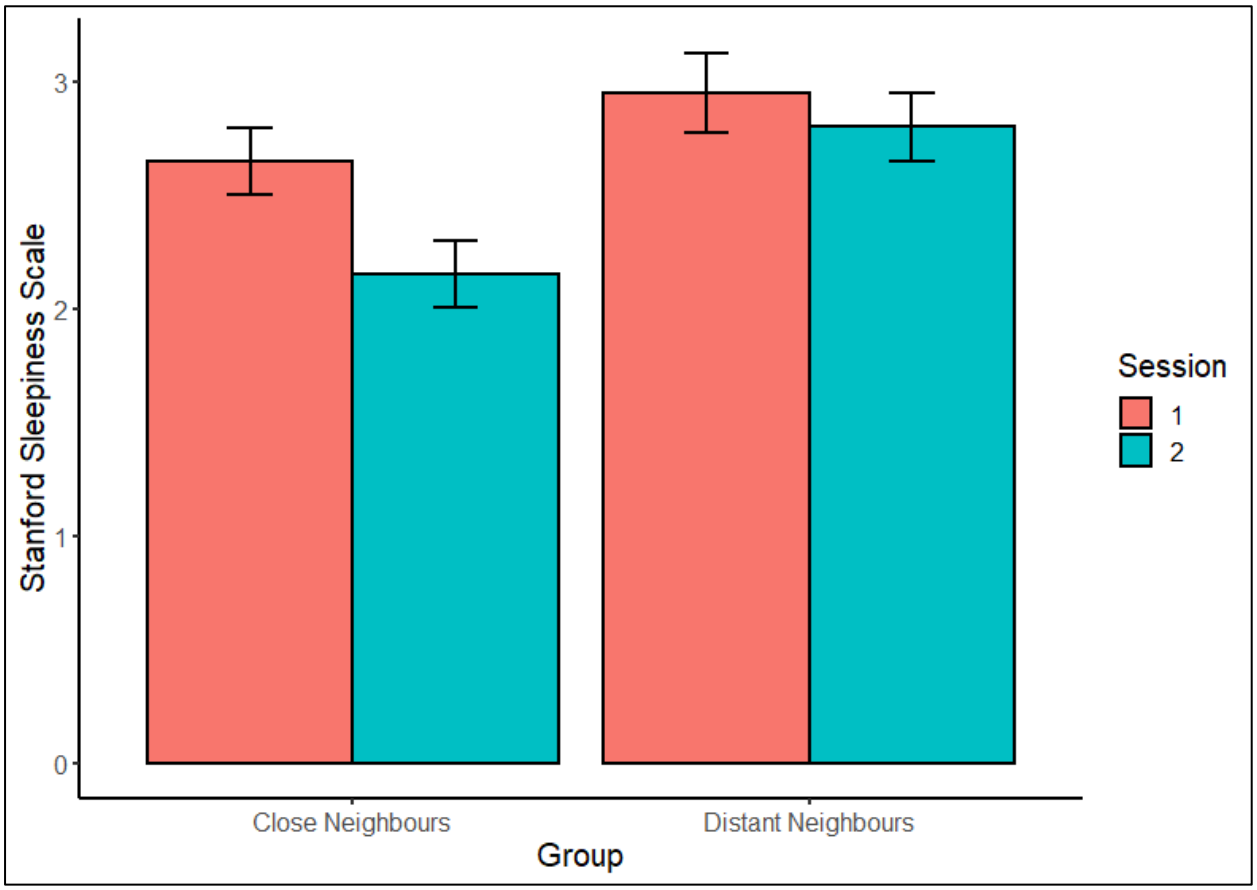



*Figure 2. Grouped bar chart showing Stanford Sleepiness Scale scores by Group (Close Neighbours / Distant Neighbours) and Session (Session 1 / Session 2). Error bars represent standard error. Y axis is a scale from 1 – Feeling active and vital, wide awake to 7 – Almost in reverie, sleep onset soon, lost struggle to stay awake.*

Overall, both groups reported feeling sleepier in Session 2 than Session 1 (see Figure 2). Assumptions were met and a 2x2 mixed factorial ANOVA was conducted on participant SSS ratings across sessions and revealed no main effect of Session (*Session 1/ Session 2*), F (1,38)=3,11, p=.086, *ns*, or Group (*Close Neighbours / Distant Neighbours*), F(1,38)= 3.61, p=.065, *ns*.

### *Psychomotor Vigilance Task*

See Table 3 (below) for a summary of descriptive statistics obtained for the PVT:

*Table 3. Means and standard deviations measures collected as part of the Psychomotor Vigilance Task (PVT) by group and at sample level.*

| | Close Neighbours | Distant Neighbours | Total |
|---|---|---|---|
| **Session 1** | | | |
| Minor lapses | 0.77 (1.36) | 1.11 (1.23) | 0.94 (1.25) |
| Major lapses | 0.00 (0.00) | 0.22 (0.43) | 0.14 (0.36) |
| False Starts | 0.38 (0.65) | 0.83 (1.47) | 0.74 (1.19) |
| RT | 285.59 (40.41) | 283.83 (45.43) | 282.02 (41.40) |
| **Session 2** | | | |
| Minor lapses | 0.85 (0.99) | 1.00 (1.69) | 0.93 (1.41) |
| Major lapses | 0.08 (0.28) | 0.00 (0.00) | 0.03 (0.18) |
| False Starts | 0.31 (0.48) | 0.72 (1.08) | 0.54 (0.89) |
| RT | 282.30 (33.49) | 278.68 (39.60) | 280.20 (36.62) |

*Note.* Minor lapses are a failure to react or any reaction exceeding 355ms; Major lapses are a failure to react or any reaction exceeding 500ms; False starts are premature response times < 100 ms; RT= Mean RT(ms). Minor Lapses, Major Lapses, and False Starts are in counts, RTs are in ms.

Assumptions were met and four 2x2 mixed factorial ANOVAS were conducted on key indicators of participant Psychomotor Vigilance Task performance (minor lapses, major lapses, false starts, mean RT). The within-subjects variable for each was Session (*Session 1/ Session 2*) and the between-subjects independent variable was Group (*Close Neighbours / Distant Neighbours*). Analyses revealed no significant main effects or interactions (reported in Table 4, below), indicating that the groups were well-matched in terms of their psychomotor vigilance and the sample had similar levels of alertness across both sessions.

*Table 4. Results from a series of mixed factorial ANOVAs conducted to test for the effect of Group (Close Neighbours / Distant Neighbours) and Session (Session 1 / Session 2) on psychomotor vigilance.*

| **Variable** | **Main effect** | ***df*** | ***F*** | ***p*** | ***η2*** |
|---|---|---|---|---|---|
| Minor lapses | Session | (1,29) | .003 | .96 | .00 |
| | Group | (1,29) | .417 | .52 | .01 |
| Major lapses | Session | (1,29) | 1.15 | .29 | .04 |
| | Group | (1,29) | 1.15 | .29 | .00 |
| False Starts | Session | (1,29) | .341 | .56 | .01 |
| | Group | (1,29) | 1.55 | .22 | .03 |
| RT | Session | (1,29) | .601 | .44 | .00 |
| | Group | (1,29) | .039 | .85 | .00 |

***Novel word learning task***

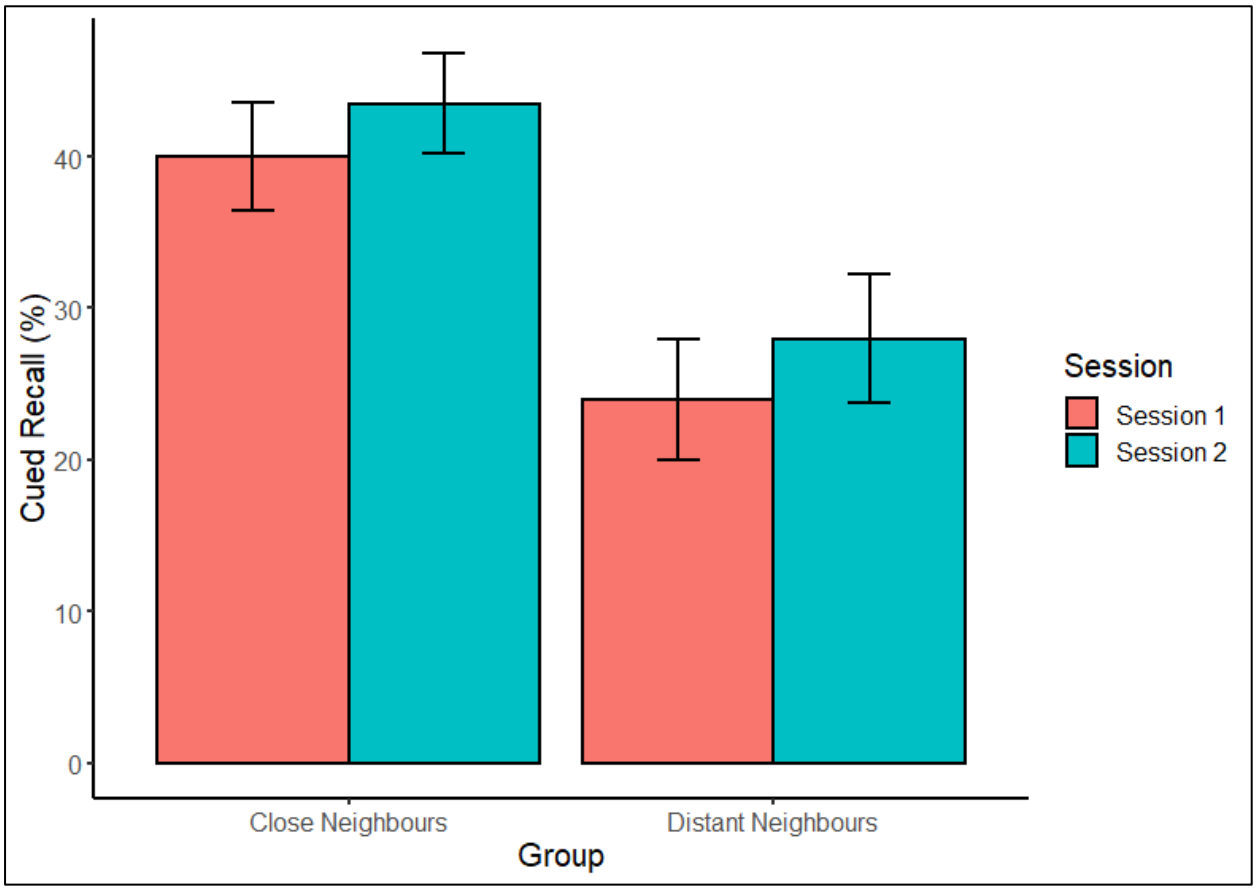



*Figure 3. Grouped bar chart showing cued recall (stem completion; %) by Group (Close Neighbours / Distant Neighbours) and Session (Session 1/ Session 2). Error bars represent standard error.*

Word responses were recorded and later coded by a single member of the research team. The coding criteria adopted was strict, only clear, unambiguous responses of which there was no doubting the accuracy were scored as correct. See Figure 3, above, for descriptive statistics. Assumptions were met and a 2x2 mixed factorial ANOVA was conducted on the cued recall responses for the novel word learning task. The within-subjects independent variable was Session (*Session 1/ Session 2*), and the between-subjects variable was Group (*Close neighbours/ Distant Neighbours*). Analyses revealed a main effect of Session, F (1,38)=6.51, *p*=.015, indicating that the sample as a whole improved overnight (a consolidation effect). There was also a significant main effect of Group, F (1,38)=4.48, *p*=.04, indicating a general advantage for novel words which were more closely related to existing words. There was no significant interaction, F (1,38)=.029, *p*=.86, *ns*.

## Discussion

This study aimed to address an open question in the word learning literature about the extent to which neighbourhood (the degree of similarity between the novel word and previously established entries in the lexicon) matters for overnight memory improvement. This study also aimed to establish which type of stimulus is better suited towards use in sleep and memory consolidation research. A 2x2 mixed factorial ANOVA was carried out on cued recall responses from a novel word learning task which revealed a significant main effect of Session (*Session 1*[Evening], *Session 2*[Morning]) as well as a significant main effect of group (*Close Neighbours/Distant Neighbours*), indicating that both groups improved overnight to a similar degree but also that there was a general advantage of close neighbours over distant neighbours.

These findings are consistent with numerous studies in the literature suggesting that existing knowledge, and therefore, neighbourhood matters (Tse et al., 2007; Lewis and Durrant, 2011; James et al., 2018). The way in which the results suggest that neighbourhood matters were not expected. It was speculated above that either superior overnight improvement would be observed with close neighbours because the brain more rapidly consolidates information which more closely related to existing knowledge or superior overnight improvement would be observed in favour of distant neighbours because sleep would have more of an active process in learning that type of stimuli. What was in fact found was somewhere in between – there was without a doubt an advantage for close neighbours which spanned across both sessions, but it was not an advantage in terms of overnight improvement (consolidation). Instead, the advantage related to encoding – the current evidence suggests that the brain will consolidate information which is successfully encoded during wakefulness (if sufficiently salient, see Stickgold and Walker, 2013, for discussion) at a comparable rate regardless of neighbourhood, but words which relate to already existing lexicon entries will find its way into short-term hippocampal storage significantly easier.

For sleep and memory researchers, the results are clear: (1) overnight spoken novel word learning paradigms are well-situated towards general use in consolidation research owing to their consistency, and; (2) investigations which require a higher baseline of encoding or ones in which it is likely that the population of interest will

underperform (e.g., participants experiencing chronic sleep disruption) should steer towards word stimuli which more closely resemble established entries in the lexicon. Future investigations might seek to adopt a similar paradigm whilst both increasing the number of trials (and therefore the cognitive load, see Born and Feld, 2017, for recent guidelines on how active systems consolidation is affected by task load), increasing the lexical distance in the distant neighbour condition and using polysomnography to test for more subtle differences in the expression of consolidation between close and distant neighbours.

A key limitation is the lack of a wake control group. Without a wake comparison group, the overnight improvement observed in the data cannot be attributed to the role of sleep. For example, the difference in performance between the first and second session could have been due to time elapsed, repeated testing, or some other factor not related to sleep. It is also possible that the results are due to circadian rhythm effects rather than sleep itself (Foster & Kreitzman, 2017).

This study was also limited in that participants from both groups reported feeling sleepier in the evening and the distant neighbour group reported feeling sleepier than the close neighbour group, and both trends closely approached significance. Therefore, it could be possible that increased sleepiness in the distant neighbours group contributed to poorer encoding and blunted consolidation (and therefore that there was a failure to detect an advantage for consolidation of more distant neighbours due to the need for a more active role of sleep). Interestingly, however, there were no significant differences in any PVT measures. If it were the case that sleepiness played an explanatory role in our results, it would be meaningful that subjective fatigue could impact upon encoding and subsequent consolidation despite objectively in-tact vigilance.

Overall, our study set out to investigate the extent to which lexical neighbourhood matters for overnight consolidation and to test which novel word learning paradigm is best situated for use in sleep and memory research. Overnight memory improvement wasobserved regardless of neighbourhood, but a general advantage existed at encoding for word stimuli more closely related to already existing entries in the lexicon. Future research might seek to increase task load and use polysomnography to allow for detection of subtle differences between groups.

# Experiment 2

## Introduction

Previous research has found that declarative memory is maintained with a period of consolidation in sleep (e.g., Plihal & Born, 1997; Gaskell & Dumay, 2003; Davis et al., 2009). It has also been suggested that memory consolidation is supported by interactive systems and that this organisation can result in interference between declarative and procedural memories (but not if the individual is sleeping; Brown & Robertson, 2007; Gagne and Cohen, 2015; Kim, 2020). Interference-based forgetting occurs when "new information acquired either before or after a learning event attenuates memory expression (proactive and retroactive interference, respectively). Multiple learning events often occur in rapid succession, leading to competition between consolidating memories" (Crossley, Lorenzetti, Naskar, O'Shea, Kemenes, Benjamin, and Kemenes, 2019). Interference can be retroactive (where forming a new representation disrupts memory for the preceding one; Alves, Malow, Lange, McEvoWy, Olson, Turkingham, Windisch, Samuels, Stevens, Berry-Kravis, and Weese-Mayer, 2006; Wixted, 2004), or proactive (when the older memory is prioritised and the new one is forgotten; Brown and Robertson, 2007).

What is less clear is whether interference affects the expression of consolidation within a memory system and, more specifically, between verbal and non-verbal measures of declarative memory. Better understanding of interference within the declarative memory system is an imperative both for memory research broadly and for research focussed on better understanding the relationship between sleep and declarative memory consolidation. For example, neuropsychological batteries are commonplace when investigating sleep pathologies (see Ahuja et al., 2018, for discussion), and the available literature on memory dysfunction in common disorders of sleep (e.g., Obstructive Sleep Apnoea Syndrome, see Bucks et al., 2017 for review) is inconsistent. It is plausible that having participants complete many cognitive assessments at once is detrimental and that this approach should be replaced with paradigms favouring a small number of dedicated tasks shown to be consistent in the literature.

However, recent investigations focussed on interference have made several things clear. It seems to be the case that if independent systems are used, there is no interference (Crossley et al., 2019; Gagne and Cohen, 2015). When the newly acquired representation is absorbed during a labile period of consolidation, the newer representation is prioritised over the older one (Crossley et al., 2019). Interference seems to be to some degree regulated by *cognitive control* ("the ability to regulate thoughts and actions in accordance with internally represented behavioural goals"; Brawn, 2013, Nusbaum, and Margoliash, 2013). As such, it has been proposed that cognitive control is regulated by two separate processes: proactive and reactive control (in essence, "early selection" vs "late correction"; Miller and Cohen, 2001); that there exists large intra-individual variation in which strategy is dominant in any given moment, and; these operating modes are sensitive to small tweaks to otherwise largely comparable tasks (the *Dual Model of Control*, Braver et al., 2012).

Taken together, these findings suggest that care should be taken when subjecting participants to more than one learning task on a singular occasion, but that two tasks within the same memory system ought to be able co-exist without directly competing together, if the communication between the two is disrupted (here by sleep). The current investigation seeks to test for interference effects between lexical (verbal declarative) and spatial (non-verbal declarative) information with a period of nocturnal sleep in the delay period between test and re-test. In doing so, we hope to empirically justify use of more than one declarative memory task in consolidation research. It is hypothesised that the two tasks (novel word learning and object location) will not interfere with each other despite the memories occurring within the same memory system (declarative), since communication between the two will be disrupted by sleep.

## Methods

### *Participants*

A total of 80 undergraduate students from the University of York with English as their first language completed this study and were recruited using the university's online recruitment system. Participants were opportunistically assigned into four groups which were well-matched for age (*Word Learning*: Range=18-23, Mean=19.75, SD= 1.21;

*Object Location*: Range=18-24, M=19.85, SD=1.49; *Combined[Word Learning>Object Location]*: Range=18-24, M=19.52, SD=1.88; *Combined[Object Location>Word Learning]*: Range=18-29, M=19.75, SD=2.59), and sex (*Word Learning*: 17 females [85%], 3 males [15%]; *Object Location*: 19 females [95%], 1 male [5%]; *Combined[Word Learning>Object Location]*: 16 females [80%], 4 males [20%]; *Combined[Object Location>Word Learning]*: 15 females [75%], 5 males [25%]). Data from the group who were only tasked with word learning (N=20) was collected as part of Experiment 1 (above, namely the *Close Neighbour* group). Ethical approval was granted by the Research Ethics Committee of the Department of Psychology, University of York. Participants gave informed consent, were reimbursed for their time, and data was collected and stored in accordance with the General Data Protection Regulation (GDPR). Participants were screened such that only non-smokers, not taking any psychoactive medication, who did not have a known sleep disorder and who reported no history of alcohol or drug abuse, were invited to take part. All participants had normal or corrected-to-normal vision and hearing. Participants were also instructed not to consume alcohol in the 24 hours before and to not consume caffeine on the days of the experiment, which were confirmed by self-report on arrival.

### *Materials*

**Alertness materials.** See Experiment 1 (above).

**Lexical materials.** The lexical materials used here were identical to the *Close Neighbour* group in Experiment 1 (above).

**Visual materials.** For the Object-Location Task, 50 unique object images were used. These objects appeared against a blue and orange checkered grid for reference.

**Questionnaires.** See Experiment 1 (above).

**Apparatus.** See Experiment 1 (above).

**Stimuli presentation.** See Experiment 1 (above).

## Design

A 4 (*Task: Word Learning*, *Object Location*, *Word Learning-followed-by-Object Location, Object Location-followed-by-Word Learning)* x2 (*Session:* Session 1[Evening], Session 2[Morning]) mixed factorial ANOVA designed to test whether there are interference effects on overnight consolidation between two different declarative memory tasks: a verbal declarative novel word-learning task, and a non-verbal declarative object-location task. The between-subjects variable *Task* was characterised by the type or combination of tasks participants completed. Participants in the Word Learning group completed only the PVT and a novel word-learning task. Participants in the Object Location group task completed only the PVT and an Object-Location task. Finally, participants in the Combined groups completed both tasks in a counterbalanced order. The within-subjects independent variable *Session* had 2 levels, with Session 1 taking place in the evening and Session 2 taking place the following morning with a delay period including a night of nocturnal sleep.

## Procedure

The same as for Experiment 1 (above)

### *Session 1 (Evening)*

Upon arrival participants filled out consent forms, re-confirmed that they met the relevant inclusion criteria and completed a sleep habits questionnaire which captured relevant information such as each participants' typical bed and wake times, subjective ratings of sleep quality and napping habits Alertness for Session 1 was measured both by self-report (SSS) and then by completion of the Psychomotor Vigilance Task. Participants then completed a learning phase and then a test phase for each task relevant to that group. For the Combined groups, the task order was counterbalanced. All experimental tasks were conducted on computers in isolated individual testing booths. Stimuli were presented using headphones for the spoken stimuli. Responses were recorded by mouse-click for the PVT by clicking and dragging for the Object-Location task. For the word learning tasks, responses were recorded by keyboard stroke for the learning phase (phoneme monitoring), and for the test phase (cued recall) participants made verbal responses which were recorded, transcribed by the experimenter, and

coded as being correct or incorrect. Screen brightness and speaker volume were kept the same across the sample.

***Session 2 (Morning)***

Upon arrival for Session 2, participants filled out the SSS, completed the Psychomotor Vigilance Task and finally completed a re-test phase for the combination of tasks relevant to that group. The order of re-test in Group 3 was the same as it was for that individual in Session 1 the previous evening. Participants were then free to go and were debriefed as in phase 1.

***The tasks:***

**Psychomotor Vigilance Task (PVT).** See Experiment 1 (above).

**Novel Word-Learning Task (Training).** See Experiment 1 (above).

**Novel Word-Learning Task (Test).** See Experiment 1 (above).

**Object-Location Task (Training).** The Object Location task was adapted from one used by Rudoy, Voss, Westerberg, and Paller (2009). The key difference between our version of the task and that of Paller et al., (2009) is that in our task there are no audio stimuli since we are not investigating Targeted Memory Reactivation (TMR). For the Object-Location Task, participants were taught to associate each of 50 unique objects with a location on a computer screen and were tested shortly after the learning phase and then following a night's sleep.

During an initial exposure phase, 50 object stimuli were presented for 3s with a blank window lasting 1s between each trial. Participants were instructed to view the objects from 1m away. The task was designed such that each object was given a random screen location for each participant. Participants were privy to a background grid for reference, but the objects could appear at any coordinate on the screen, i.e., the grid and object's locations were completely independent.

During training, participants were tasked with remembering each object's correct location (in accordance with the preceding exposure phase). Each trial would begin with an object stationary in the centre of the screen. Participants completed each trial by dragging and dropping the object where the participant believed the correct location was. Feedback was provided after completion of each trial by presenting the object in its correct location for 3s. Training was completed in a series of rounds in which the objects were presented in random order. The first two rounds included all 50 objects, and subsequent rounds would skip a particular object if the participant had successfully placed it in its correct location (within 3.5cm) in the previous two rounds. This continued until all objects had been eliminated in this manner.

**Object-Location Task (Test).** Participants were assessed once immediately after training and once the following morning to provide pre- and post-sleep memory results. Testing involved exposing participants to each of the 50 images in random order as in training and having participants click and drag them to where they thought the correct location was. Importantly, no feedback was given at test.

## Results

### *Sleep habits*

Figure 4 below shows distributions of subjectively estimated sleep onset and wake times for each group:

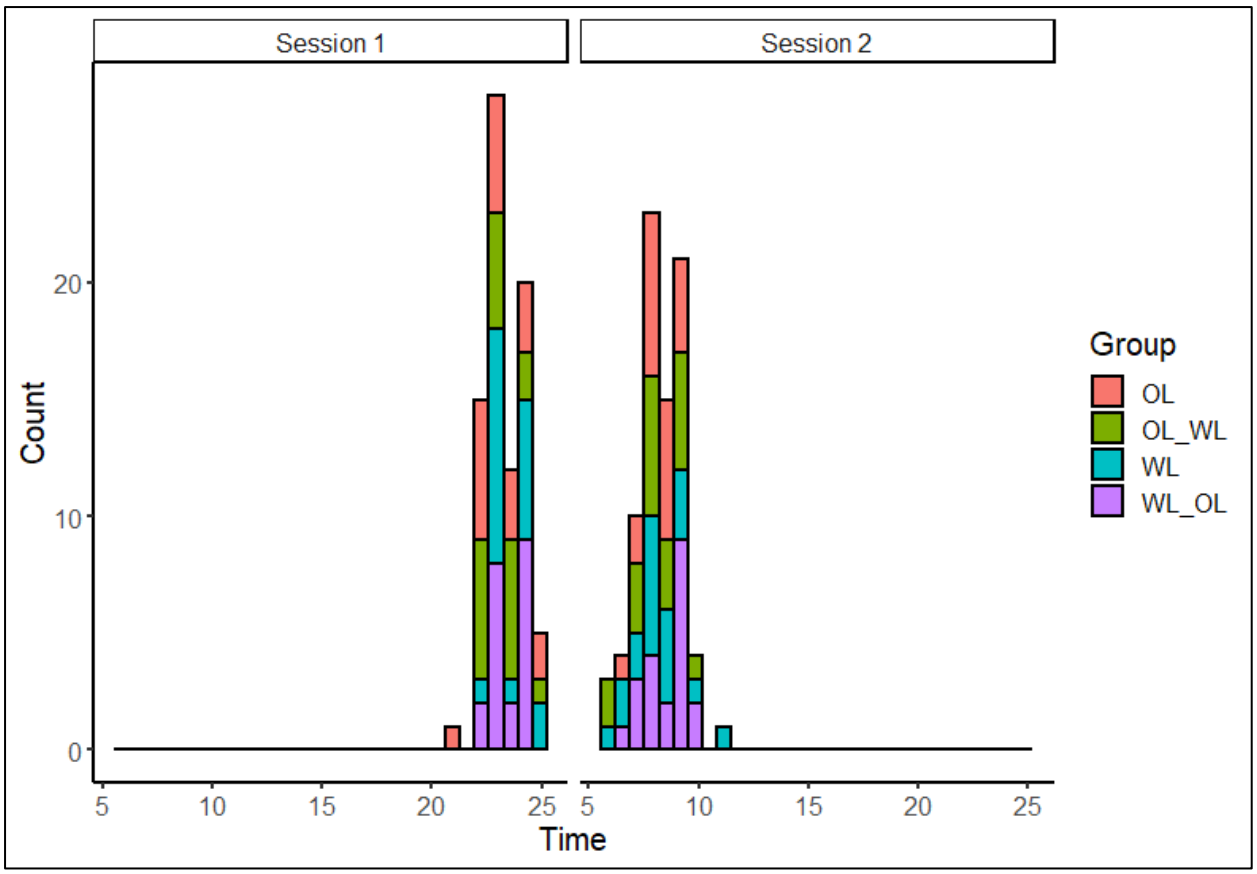



*Figure 4. Stacked histogram showing bedtimes (left) and wake times (right) by Group (OL = Object Location; OL_WL = Object Location followed by Word Learning; WL = Word Learning; WL_OL = Word Learning followed by Object Location) and Session (Session 1/ Session 2 ). Time is in hours, e.g., 25 = 1AM.*

Table 5 below shows descriptive statistics for subjectively estimated sleep parameters by group:

*Table 5. Means and standard deviations for self-reported sleep measures by group and at sample level.*

| Variable | OL | WL | OL > WL | WL > OL | Total |
|---|---|---|---|---|---|
| Total Sleep Time | 528.6 (54.6) | 492.6 (76.2) | 523.2 (76.2) | 520.2 (67.2) | 515.0 (67.2) |
| Sleep Onset Latency | 21.4 (12.07) | 25.3 (18.21) | 23.5 (13.69) | 23.1 (12.19) | 23.3 (14.0) |
| PSQI Global | 4.1 (2.02) | 5.1 (1.90) | 5.4 (2.50) | 4.8 (1.97) | 4.1 (2.3) |
| Sleep Onset | 23:10 (58) | 23:33 (47) | 23:06 (39) | 23:3 (47) | 23:18 (49) |
| Wake | 08:15 (44) | 08:12 (68) | 08:03 (62) | 08:30 (56) | 08:16 (49) |

*Note.* Total Sleep Time (mins); Sleep Onset Latency (mins); PSQI = Pittsburgh Sleep Quality Index; ESS = Epworth Sleepiness Scale. SD for Sleep Onset and Wake (mins). OL=Object Location; WL=Word Learning; OL>WL= Object Location followed by Word Learning; WL>OL= Word Learning followed by Object Location.

*Table 6. Results for a series one-way ANOVAs conducted to test for differences between groups (OL=Object Location; WL=Word Learning; OL>WL= Object Location followed by Word Learning; WL>OL= Word Learning followed by Object Location.) on subjectively measured sleep parameters.*

| Variable | *f* | *p* | *η2* |
|---|---|---|---|
| Total Sleep Time | 1.60 | .196 | .06 |
| Sleep Onset Latency | .26 | .854 | .01 |
| PSQI Global | 1.26 | .293 | .05 |
| Sleep Onset | 1.30 | .281 | .05 |
| Wake | .91 | .442 | .03 |

Participant self-reports revealed that all groups were well-matched for Total Sleep Time (mins), Sleep Onset Latency (mins), sleep quality (PSQI), daytime somnolence (ESS), sleep onset and wake times. Assumptions were met and one-way ANOVAs revealed no significant differences between the groups, as reported in Table 6 above.

***Stanford Sleepiness Scale***

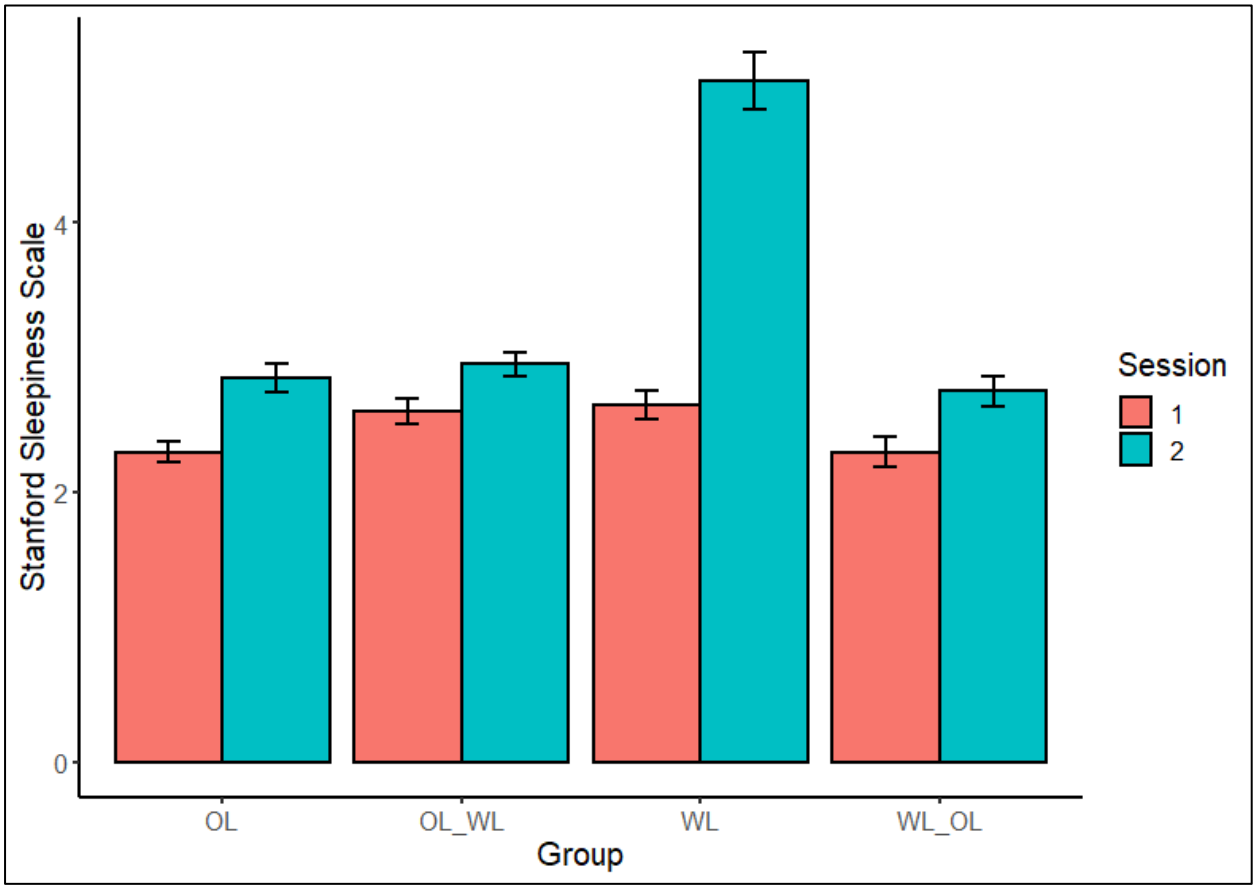



*Figure 5. Grouped bar chart showing Stanford Sleepiness Scale scores by Group (OL = Object Location; OL_WL = Object Location followed by Word Learning; WL = Word Learning; WL_OL = Word Learning followed by Object Location) and Session (Session 1 / Session 2). Error bars represent standard error. Y axis is a scale from 1 – Feeling active and vital, wide awake to 7 – Almost in reverie, sleep onset soon, lost struggle to stay awake.*

See Figure 5 (above) for descriptive statistics for participant SSS ratings. Assumptions were met and a 2x4 mixed factorial ANOVA was conducted on participant SSS ratings across sessions. The within-subjects independent variable was Session *(Session 1/ Session 2*), and the between-subjects independent variable was Group *(Object Location/ Object Location followed by Word Learning/ Word Learning/ Word Learning followed by Object Location).* Analyses revealed a significant main effect of Session, F (1,77)=4.08, *p*=.047, indicating that the sample as a whole felt less alert in the morning session. There was no main effect of Group, F (1,77)=.803, p=.50, *ns*.

### *Psychomotor Vigilance Task*

*Table 7. Means and standard deviations measures collected as part of the Psychomotor Vigilance Task (PVT) by group and at sample level.*

| | OL | WL | OL > WL | WL > OL | Total |
|---|---|---|---|---|---|
| **Session 1** | | | | | |
| Minor lapses | 0.7 (1.2) | 0.8 (1.3) | 0.4 (0.7) | 0.9 (0.9) | 0.7 (1.0) |
| Major lapses | 0.1 (0.3) | 0.00 (0.00) | 0.00 (0.00) | 0.2 (0.4) | 0.1 (0.3) |
| False Starts | 0.8 (1.3) | 0.4 (0.7) | 0.4 (0.8) | 0.4 (0.7) | 0.5 (0.9) |
| RT | 266.3 (33.8) | 281.9 (39.4) | 278.6 (29.2) | 285.0 (25.7) | 275.1 (29.5) |
| **Session 2** | | | | | |
| Minor lapses | 0.7 (1.0) | 0.8 (1.0) | 0.8 (0.9) | 1.2 (2.2) | 0.9 (1.5) |
| Major lapses | 0.02 (0.0) | 0.1 (0.3) | 0.0 (0.0) | 0.1 (0.1) | 0.0 (0.0) |
| False Starts | 0.3 (0.5) | 0.3 (0.5) | 0.3 (0.5) | 0.6 (0.9) | 0.4 (0.7) |
| RT | 279.7 (36.2) | 279.0 (36.2) | 289.4 (28.4) | 285.5 (33.5) | 286.2 (33.4) |

*Note.* Minor lapses are a failure to react or any reaction exceeding 355ms; Major lapses are a failure to react or any reaction exceeding 500ms; False starts are premature response times < 100 ms; RT= Mean RT(ms). Minor Lapses, Major Lapses, and False Starts are in counts, RTs are in ms). OL=Object Location; WL=Word Learning; OL>WL= Object Location followed by Word Learning; WL>OL= Word Learning followed by Object Location.

To complement subjectively measured alertness participants also completed the more objective Psychomotor Vigilance Task (PVT). Table 7 above shows descriptive statistics for 4 objective dependent variables relating to alertness by session and group. Assumptions were met and a series of 4x2 mixed factorial ANOVAs were carried out to test for difference in psychomotor vigilance across group and session, as for the SSS analyses above. These results are displayed in Table 8 (below) and show that the sample made more major lapses in Session 2 than in Session 1. Taken together, the subjective and objective measures complement each other and suggest alertness was lower in the second session than for the first, but that this uniform across the groups.

*Table 8. Results from a series of mixed factorial ANOVAs conducted to test for the effect of Group (OL = Object Location; OL_WL = Object Location followed by Word Learning; WL = Word Learning; WL_OL = Word Learning followed by Object Location) and Session (Session 1 / Session 2) on psychomotor vigilance.*

| Variable | Main effect | *df* | *F* | *p* | *η2* |
|---|---|---|---|---|---|
| Minor lapses | Session | (1,65) | .72 | .398 | .00 |
| | Group | (1,65) | .84 | .473 | .02 |
| Major lapses | Session | (1,65) | 4.62 | **.033*** | .03 |
| | Group | (1,65) | 1.34 | .261 | .03 |
| False Starts | Session | (1,65) | .85 | .359 | .01 |
| | Group | (1,65) | .16 | .926 | .01 |
| RT | Session | (1,65) | 2.64 | .107 | .02 |
| | Group | (1,65) | 1.38 | .252 | .02 |

*Note. * = p < .05.*

### *Novel word learning task*

Assumptions were met and a 2x3 mixed factorial ANOVA was conducted on the cued recall responses for the novel word learning task to test for task interference. The within-subjects independent variable was Session (*Session 1/ Session 2*), and the between-subjects variable was Group (*Object Location/ Object Location followed by*

*Word Learning/ Word Learning followed by Object Location*). Word responses were recorded and later coded by a single member of the research team. The coding criteria adopted was strict, only clear, unambiguous responses of which there was no doubting the accuracy were scored as correct. See Figure 6 (below) for summary statistics.

Analyses revealed no main effect of Session, F (1,57) =3.80, *p*=.056, *ns*, indicating that the sample maintained what they learned the evening before. There was also no significant main effect of Group, F (1,57)=41.45, *p*=.24, indicating that there was no evidence of interference. There was finally no significant interaction, F (1,57)=.026, *p*=.97, *ns*.

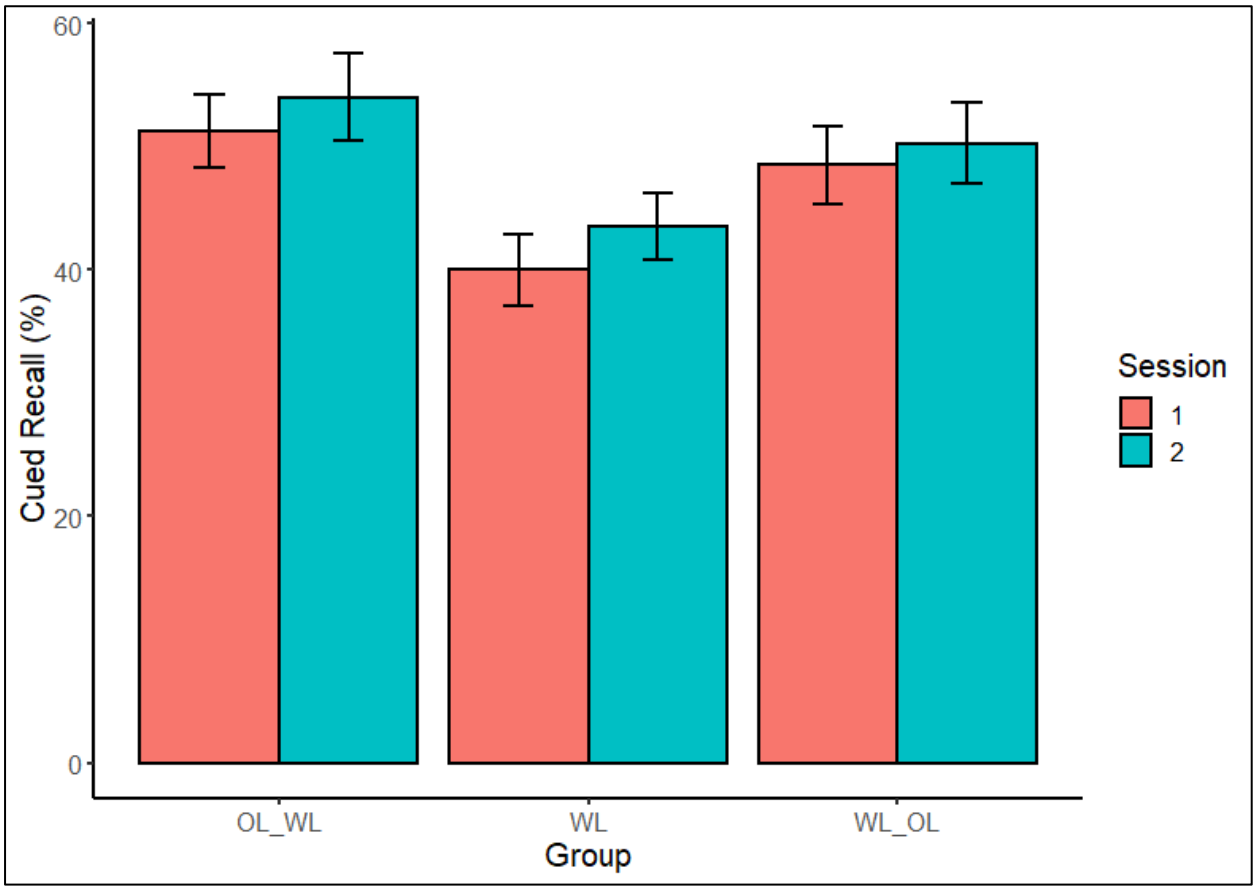



*Figure 6. Grouped bar chart showing cued recall (stem completion; %) by Group (OL = Object Location; OL_WL = Object Location followed by Word Learning; WL = Word Learning; WL_OL = Word Learning followed by Object Location) and Session (Session 1/ Session 2). Error bars represent standard error.*

Finally, a Bayesian mixed factorial ANOVA was conducted to compare the evidence for and against the null hypothesis (no interference) relative to the alternative hypothesis (interference) for the effects of Group (*Word Learning/ Word Learning followed by Object Location; Object Location followed by Word Learning)* and Session

(*Session 1/ Session 2*). As Table 9 (below) shows, there was only anecdotal evidence for no task interference.

*Table 9. Bayes Factors and Confidence Intervals for Bayesian mixed factorial ANOVA.*

| Effect | BF | CI |
|---|---|---|
| Group | .49 | ±0.01% |
| Session | .23 | ±0.03% |
| Group + Session | .11 | ±1.45% |
| Group + Session + Group: Session | .02 | ±2.74% |

### *Decay Index*

Following previous investigations exactly (see Cairney, Lindsay, Sobczak, Paller, and Gaskell, 2016, for more detailed discussion), a memory decay index was created firstly by measuring how accurate (distance in cm) each participant was in each trial and then secondly by subtracting the combined Session 1 score from the combined Session 2 score. A positive score reflects decay (i.e., forgetting), and a negative score here reflects retention (i.e., consolidation/overnight improvement). See Figure 7 below for distributions of mean decay index by group:

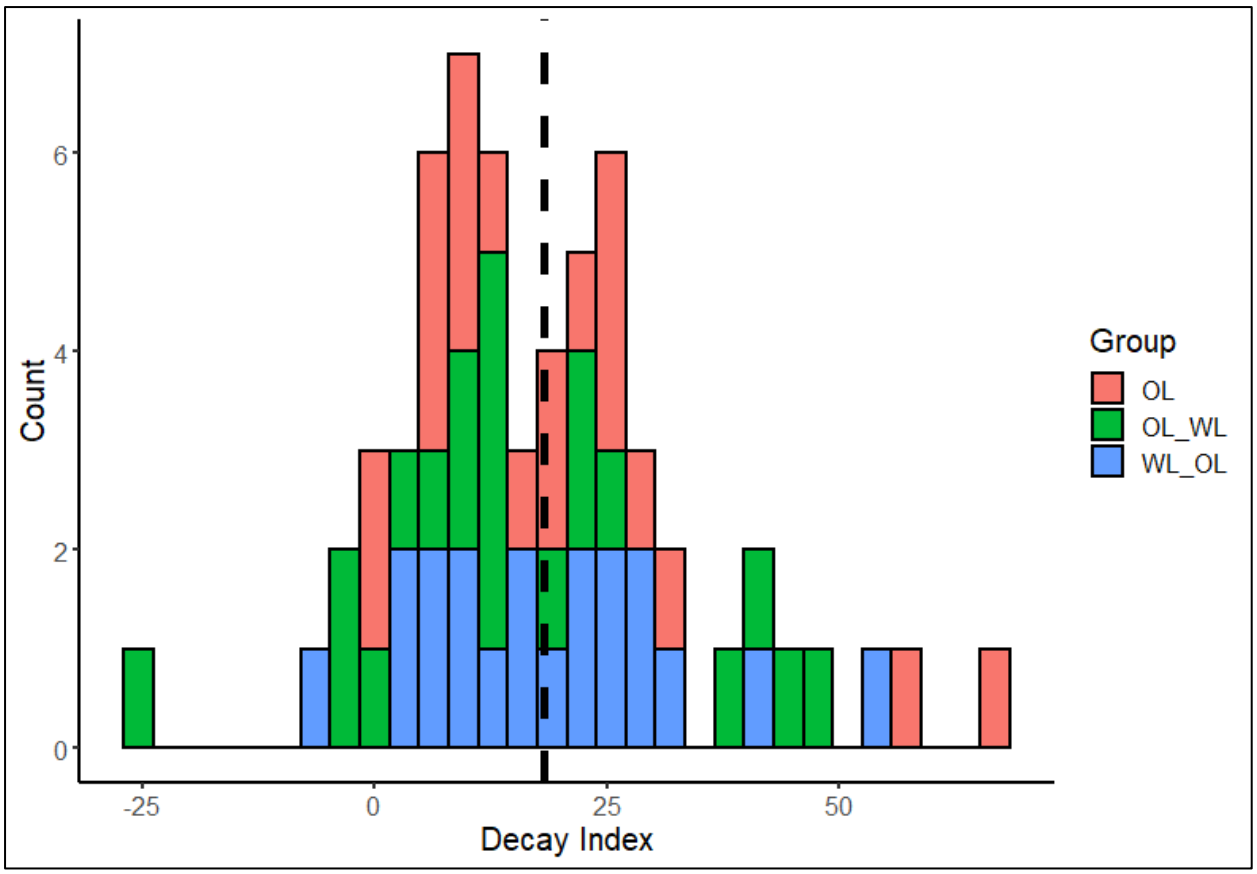



*Figure 7.Stacked histogram showing decay index scores by Group (OL = Object Location; OL_WL = Object Location followed by Word Learning; WL_OL = Word Learning followed by Object Location) for those who complete the Object Location task. Black dashed line is sample mean.*

See Figure 8 (below) for a bar chart displaying the mean decay index by group. Assumptions were met and a one-way ANOVA was carried out on the decay index scores and no significant effect of Group (*Object Location*, *Object Location followed by Word Learning*, *Word Learning followed by Object Location*) was found, $F(2,57)=.382$, $p=.68$, *ns,* suggesting that the groups did not differ in overnight decay and therefore that there was no interference.

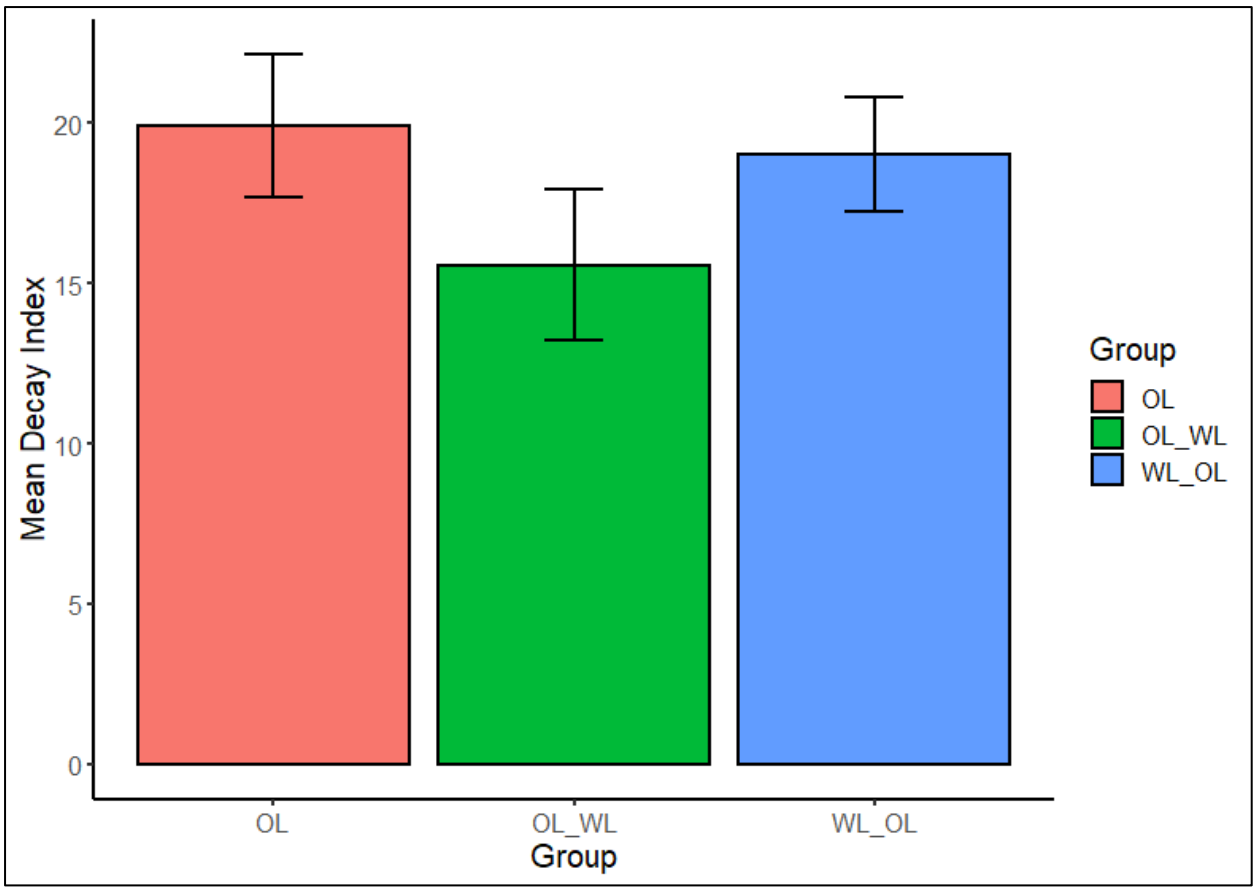



*Figure 8. Bar chart showing mean Decay Index by Group (OL = Object Location; OL_WL = Object Location followed by Word Learning; WL = Word Learning; WL_OL = Word Learning followed by Object Location). Error bars represent standard error.*

Finally, a Bayesian one-way ANOVA was carried out to compare the probability null (no interference) and the alternative hypothesis (interference) for mean decay index scores by group (*Object Location*, *Object Location followed by Word Learning*, *Word Learning followed by Object Location*). The results of this analysis are strongly in support of the null hypothesis, and therefore that no interference between tasks took place ($B_{10}$ = .18, ±0.01%).

## Discussion

This study aimed to address a gap in the interference literature by testing for interference effects on the overnight improvement when participants are tasked with learning lexical (verbal declarative) and spatial (non-verbal declarative) information. It was found that the sample significantly improved overnight . Importantly, no significant main effect of Group (*Object Location*, *Object Location-followed-by-Word Learning*, *Word Learning-followed-by-Object Location*) was found, indicating that there was no evidence of interference between the two tasks. This finding was complemented by a further null result when comparing overnight decay in the object location task.

These results support our initial hypothesis that there would be no interference between the two declarative tasks arguably because communication between these memories was disrupted by sleep (Cohen and Robertson, 2017), and because cognitive control was encouraged (Braver et al., 2012). The lack of an interference effect here also has implications for sleep and memory research more generally, since there now exists empirical support for using more than one declarative memory task within the same paradigm. Whilst these results clearly do not justify the use of multiple tasks as part of a neuropsychological battery (since only two tasks were tested), they do support use of more than one task within the declarative memory system, given they are sufficiently distinct, and learning is protected by sleep.

However, the results are limited in that we found evidence that the sample were both subjectively sleepier and objectively less vigilant in the morning session than the preceding evening session. On the other hand, there were no group differences in subjective sleepiness or psychomotor vigilance which suggests both groups of participants had a similar experience.

As was the case in Experiment 1, another key limitation is the lack of a wake control group. Without a wake comparison group, the overnight improvement observed in the data cannot be attributed definitively to the role of sleep. For example, the difference in performance between the first and second session could have been due to time elapsed, repeated testing, or some other factor not related to sleep. It is also possible that the results are due to circadian rhythm effects rather than sleep itself (Foster &

Kreitzman, 2017). Whatever is driving the overnight improvement effect observed here, however, does not seem to have been affected by interference.

Future research might seek to further explore the Dual Mechanisms of Control framework and attempt to characterise the cognitive control strategy of the sample using neuroimaging techniques (the DMC framework makes clear predictions what sort of functional activity ought to be expected when comparing proactive and reactive control, Braver et al., 2012). As such, proactive control is hypothesised to be related to sustained and/or anticipatory lateral prefrontal cortex activation in the interests of goal maintenance, and reactive control ought to be characterised by transient lateral PFC activity (Botnovick, Braver, Barch, Carter, and Cohen, 2001). Such investigations might seek to better understand the relationship between subtle changes in protocol, intra-individual variation in cognitive control, and interference by trying to induce proactive/reactive control scenarios through experimenter prompts. Systematic attempts to reduce intra-individual variability in cognitive control behaviour gives the researcher greater control over potential interference and limits the likelihood of noisy data. Finally, future research might seek to adapt the experimental conditions of the current paradigm and further test whether interference can occur between verbal and non-declarative tasks in the right circumstances. Potential ways to adapt the current paradigm would be to interweave word learning and object location trials so as to increase the likelihood of competition/disruption, or to bring the two tasks closer together by introducing a common semantic or conceptual theme to both tasks (e.g., have participants learn a list of animal names whilst learning the locations for a series of animal pictures). Increased understanding of the boundaries of interference is both inherently valuable and pragmatic.

Overall, our study sought to investigate interference effects within the declarative memory system. No interference effects were observed. This finding is encouraging for researchers looking to use more than one declarative memory task in sleep and memory research, although more research is needed to fully understand the rules of interference. Future research might seek to further explore the influence of intra-individual variability in cognitive control, as well as increasing the likelihood of competition between tasks to see where the boundaries lie.

## The MASC-IT Study

As discussed above, the following study aims to facilitate better understanding of the influence of CPAP therapy on declarative memory consolidation in adults with OSAS. It was our intention to run this project, but unfortunately it was not possible due to the current circumstances.

The two preliminary studies discussed above allowed for data-driven stimuli and task choice. Experiment 1 made it clear that close lexical neighbours are more appropriate for use in studies in which participants may be sleep deprived, since the level of encoding is considerably higher for more distant neighbours. Experiment 2 also justified the use of two declarative memory tasks in the same investigation, if communication between these memories, which theoretically would otherwise be in direct competition, is disrupted by a delay period containing sleep.

### Methods

#### *Participants*

A total of 50 NHS patients aged between 18-45 years and recently diagnosed with moderate Obstructive Sleep Apnoea Syndrome (AHI 15-30) with English as their first language will complete this study (*Treatment,* N=25; *Comparison,* N=25). Participants will be randomly assigned into two groups which will be well-matched for age, BMI and cognitive ability.

**Recruitment and consenting.** Participants will give informed consent and will be reimbursed for their time. Specifically, two testers with NHS Research Passports and Letters of Access will attend New Patient Clinics and ask individuals who might potentially satisfy the inclusion and exclusion criteria (outlined below, see sections on *Inclusion Criteria* and *Exclusion Criteria*) for consent to be contacted about taking part in a future study should they meet the criteria. Individuals who are approached will be given information sheets which make it clear that they will be contacted by telephone and e-mail with further information about the study and to check that they are still interested

should they meet the relevant criteria. Individuals who meet the relevant criteria will be contacted in the manner discussed and those individuals who agree to participate in the study will re-consented on each experimental night they take part in.

**Ethical approval.** Ethical approval has been granted by the Research Ethics Committee of the Department of Psychology, University of York. Ethical approval has also granted by the Health Research Authority, NHS.

**Inclusion criteria.** Participants will be included in this study if they are between the age of 18-45, have been recently diagnosed with moderate Obstructive Sleep Apnoea Syndrome (OSAS), have normal or corrected-to-normal vision and hearing, have English as their first language and do not meet any of the exclusion criteria. Smokers will be included in this study on the basis that whether a participant smokes or not and how often they smoke in a typical day will be recorded. Our decision to include smokers was since previous studies have estimated a smoking prevalence of 35% in OSA populations as compared to 18% in the general population (Kashyap, Hock, and Bowman, 2001). As well as this, smoking has been reported as being the third most important predictor for sleep-disordered breathing after age and sleepiness (Tzischinsky, Cohen, and Doveh, 2012).

**Exclusion criteria.** Participants will be screened such that an individual will not be included if they are taking psychoactive medication, have another known sleep disorder, psychiatric disorder, or comorbidity. Examples of comorbidities common to OSAS populations include cardiovascular conditions such as hypertension, ischaemic heart disease, and stroke (note that obesity, highly comorbid with OSAS, is not included as part of the exclusion criteria).

**Confidentiality.** Data will be collected and stored in accordance with the General Data Protection Regulation (GDPR). All data will be stored pseudonymously, with each participant only being referred to using a unique participant identifier. Participant ID's will be linked to personal data, but this will only be accessible to members of the project team. Only the minimum amount of data necessary for the project will be collected and will be stored on university devices provided by the Department of Psychology. It will be included within the participant information sheets for the study that participants have a general

right to access their data, a right to rectification, erasure, restriction, and objection. It will also be made clear that they have the right to withdraw at any time. Participants will be advised that if they have a concern about any aspect of the study, they can speak with a member of the research team and that we will do our best to answer any questions they have. Finally, participants will be advised that if they remain unhappy and wish to complain formally, they can do this through the complaints procedure at the University of York.

### *Materials*

**Questionnaires. *Alertness.*** See Experiment 1 (above).

***ESS.*** See Experiment 1.

***PSQI***. See Experiment 1.

***Mood*.** Anxiety during the past month will be captured by having participants fill out the Beck Anxiety Inventory (Steer & Beck, 1997). Depression in participants will be measured using the Beck Depression Inventory (Beck, Ward, Mendelson, Mock, and Erbaugh, 1961) and the Patient Health Questionnaire (PHQ-9; Kroenke, Spitzer, and Williams, 2001), and the Zung Self-Rating Depression Scale (Zung, 1965). Positive and negative affect will be measured using the Positive and Negative Affect Schedule (PANAS; Watson, Clark, and Tellegen, 1988).

**Alertness materials.** See Experiment 1.

**Word Stimuli.** See Experiment 1.

**Visual materials.** See Experiment 2.

**Apparatus.** See Experiments 1 and 2.

**Stimuli presentation.** See Experiment 1.

**Polysomnography.** An Embla N7000 polysomnography system with a 16-bit resolution and 200 Hz sampling rate will be used to conduct polysomnographic monitoring for all participants. Scalp electrodes will be attached according to the 10-20 system at 8 standard locations: F3, F4, C3, C4, P3, P4, O1, 02, and all referenced to the combined mean of left and right mastoid. Left and right electro-oculogram, 3 electromyograms, and a forehead ground electrode will also be attached. Impedance <5kΩ will be verified at each electrode, and the digital sampling rate will be 200 Hz throughout the experiment. Sleep data will be scored using RemLogic 3.0 in accordance with the AASM Manual (American Academy of Sleep Medicine, Westchester, IL).

### *Design*

A 2 (*Intervention:* Treatment, Comparison) x 2 (*Timepoint:* Baseline, CPAP) mixed factorial ANOVA has been designed to test the influence of Continuous Positive Airways Pressure (CPAP) therapy on memory consolidation and sleep architecture in adults with Obstructive Sleep Apnoea Syndrome (OSAS). The between subjects variable *Intervention* is characterised by the timing of CPAP onset. More specifically, we want to collect behavioural and Polysomnography data for OSAS individuals on their first night of CPAP use *(Treatment*) and compare this to no-treatment baseline *(Comparison).* Participants in the *Treatment* group will be tested before receiving treatment (*Baseline*), and on the first night of the CPAP use (*CPAP*). Participants in the comparison group will be tested once at Baseline and once again within 7 days before they were scheduled to begin their CPAP therapy, at which point their involvement in the study will cease. Therefore, the within-subjects *Timepoint* has two levels for the *Treatment* group (*Baseline, CPAP*) and two levels for the *Comparison* group (*Baseline, CPAP*).

In addition to this, the *Treatment* group will have a further follow-up one month after their first night of treatment (*CPAP)* to allow us to look at the influence of CPAP on memory consolidation and sleep architecture over time (See Analysis Plan, below). Carrying out the study in this way allows the following to be tested whilst controlling for habituation effects with regards to repeated declarative memory testing: (1) whether there are declarative memory deficits in OSAS individuals; (2) whether any observed deficits can be restored at all by CPAP; (3) to what degree they can restored by CPAP; (3)

whether any significant or meaningful degree of restoration can be achieved after just one; (4) which sleep parameters are reliably implicated in these processes?

### *Procedure*

The experiment will be based around a 12-hour overnight protocol whereby at each timepoint Session 1 will take place in the evening between 8 and 10pm and Session 2 will take place the following morning between 8 and 10am (see Figure 9, below). Participants will spend the delay period between sessions sleeping in the sleep lab at the University of York where polysomnography will be recorded. All testing will take place in isolated booths in the Psychology department at the University of York.

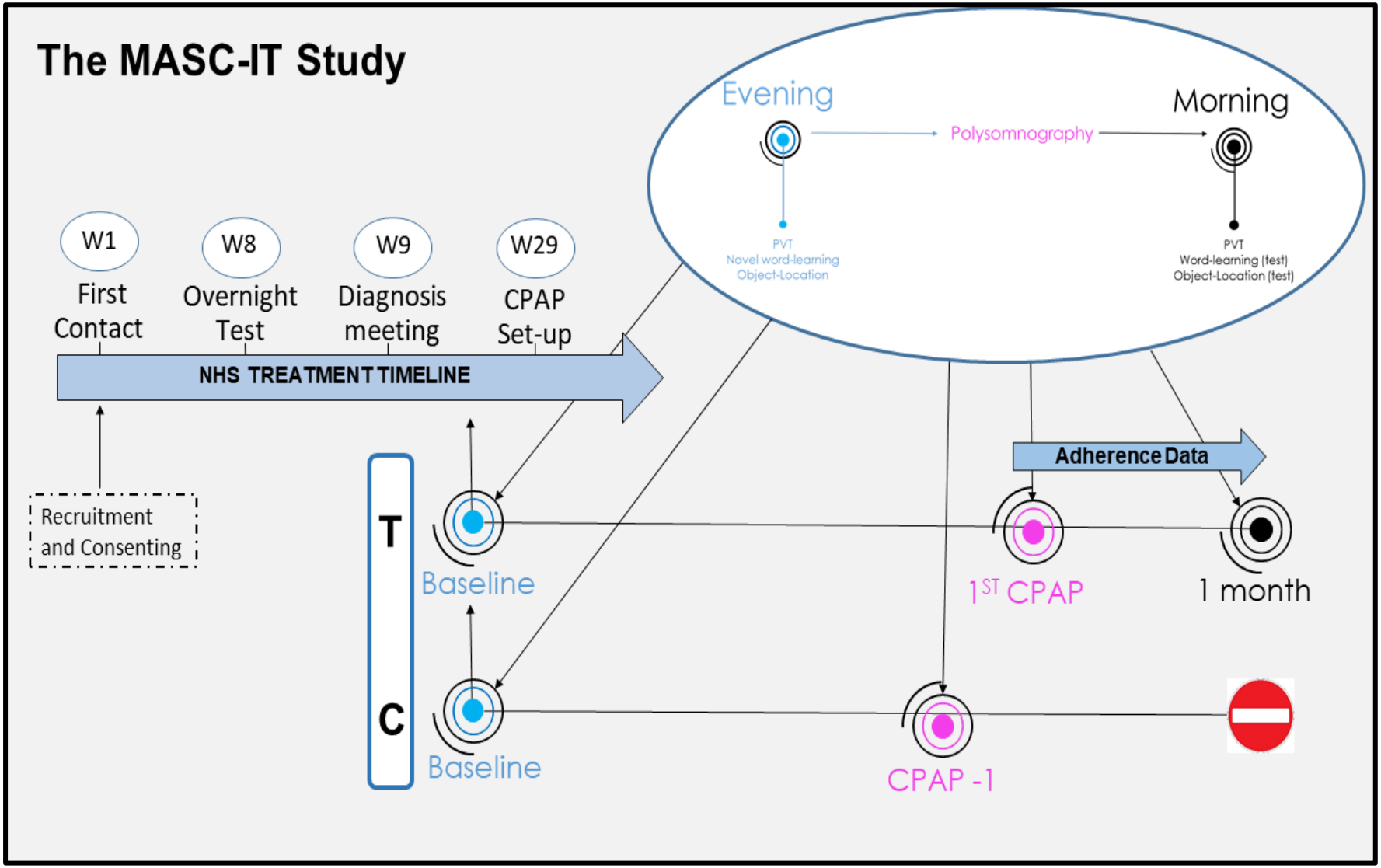



*Figure 9.* Procedure and timeline for the MASC-IT study.

**Session 1 (Evening).** Upon arrival participants will fill out consent forms and complete a battery of questionnaires. Blood saturation will then be measured using pulse oximetry. Participants will then complete our objective measure of alertness; the PVT. Participants will then complete a learning phase and then a test phase one at a time for each task. The order of the tasks will be counterbalanced. Stimuli will be presented using headphones for the spoken stimuli. Responses will be recorded by keyboard stroke for the PVT and by clicking and dragging objects from one screen location to another for the Object-Location task. For the word learning tasks, responses will be recorded by keyboard stroke for the learning phase (phoneme monitoring), and for the test phase (cued recall) participants will make verbal responses which will be recorded, transcribed by the experimenter, and coded as being correct or incorrect. Screen brightness and speaker volume will be kept the same across the sample.

**Delay period.** Participants will be settled into the sleep lab after session 1 and will be instructed to go to sleep when it feels natural. An agreed wake-up time the next morning will be established, at which point participants will be awakened if they are not in deep sleep (N3) or REM. If they are in deep sleep or REM, they will be awakened when they transition into a different sleep stage (i.e., N1, N2). All participants will be woken up in the same way, namely by the tester knocking the door 3 times and saying "*Good morning – Is it okay to come in?"*. If there was no response, the experimenter will pause for around 10 seconds and repeat. Participants will be given a 20-minute period to minimise the influence of sleep inertia before beginning the second session.

Session 2 (Morning)

In the morning session, participants will complete some questionnaires before completing the PVT and will then complete test phases for both the declarative memory tasks in counterbalanced fashion. For the *Treatment* group we will also take pulse oximetry in the morning session because it will be of interest for us to see if there are any acute overnight changes as a result of one night's use of CPAP.

***The tasks*:**

The Object location task and novel word learning task are the same as those discussed elsewhere in this chapter.

### Analysis Template

See Table 10 (below) for a brief description of the key analyses.

*Table 10.* Research questions and corresponding intended analyses for MASC-IT.

| Research Question | Statistical Analysis |
|---|---|
| Is SWS related to declarative memory? | Correlational analyses testing the relationship between SWS and overnight improvement. |
| Is adherence to treatment related to [memory, mood, attention]? | Correlational analyses testing the relationship between adherence to treatment and overnight improvement. |
| Will there be a difference in [memory, mood, attention] after 1 night of CPAP? | 2x2 Mixed Factorial ANOVA: Group (CPAP, Control) x Session (1st night S1, 1st Night S2). |
| Will there be any improvement at all? | 3x2 Mixed Factorial ANOVA: Group (CPAP, Control) x Overnight Change (Baseline, 1 week, 1 month). |

## General Discussion

### *How reversible are the memory deficits observed in OSAS?*

The MASC-IT study has the potential to inform the open question as to how reversible the cognitive impairment in OSAS is, and therefore the extent to which recovery from chronic sleep fragmentation is possible. Severe neurobehavioural impairments typically do not reverse (Dempsey et al., 2010). Despite this, most research to date has taken the position that the observed cognitive deficits in OSAS can be reversed with adequate CPAP usage (Ahuja et al., 2018). However, if treatment compliant OSAS patients continued to show memory deficits after considerable doses of treatment, this would suggest that some or all of these deficits are irreversible. Given that we have already reviewed evidence of positive long-term benefits of CPAP on memory

in OSAS patients at one month (Rosenweig et al., 2016; Kim et a., 2016), 3 months (Cannessa et al,, 2011), 12 months (Castronovo et al., 2010), and even protective effects after 10 years (Crawford-Achour, Daphinet, and Martin, 2015, in Joyeux-Faire et al., 2016), it seems reasonable to suggest that the salient issue is not whether memory deficits in OSAS are reversible, but rather how reversible are they?

To our knowledge there are no studies claiming to have demonstrated a full cognitive recovery in adults with OSAS as most investigations have only been interested in whether there is measurable improvement within a pragmatic timespan (typically less than 6 months). This suggests that rigorous longitudinal research is well-situated to answer this question. From a theoretical perspective, since the most likely explanation for memory deficits for at least a significant portion of the OSAS population is that they are being driven by the sensitivity of the hippocampus to hypoxic episodes (Gozal et al., 2001), then it stands to reason that these deficits are reversible to the extent that the hippocampus can repair itself. Given that hippocampal plasticity has been well established (e.g., Kempermann et al, 2002), there are no obvious reasons as to why complete recovery is impossible (assuming that the patient's sleep-disordered breathing is adequately managed using CPAP). Future research might seek to compare the magnitude of recovery by age (with the hypothesis that since hippocampal plasticity is less prevalent in elderly populations (Lister and Barnes, 2009) impairments will be significantly more persistent over time).

Another potentially informative approach might be to compare recovery over time with adequate CPAP use compared to a combined therapy of CPAP in conjunction with another complementary intervention suspected to boost hippocampal regeneration (candidates include mindfulness and/or antidepressants; Yang, Barros-Loscertales, Pinazo, Ventura-Campos, Borchardt, Bustamante, Rodriguez-Pujadas, Fuentes-Claramonte, Balaguer, Avila, and Walter, 2015). Regarding the latter suggestion for future research, it is worth noting that combining CPAP with other interventions (weight loss) has been shown to lead to positive long-term cardiovascular benefits in circumstances where either CPAP or weight loss in isolation showed no such benefits (Chirinos, Gurubhagavatula, Teff, Rader, Wadden, Townsend, Foster, Maislin, Saif, Broderick, Chittams, Hanlon, and Pack, 2014). As well as this, despite it being likely that a significant portion of the OSAS population have been prescribed antidepressants at

some point in their adult life (owing to the fact that OSAS has been consistently associated with depression; McCall, Harding, and O'Donovan, 2006; Baran, and Richert, 2003), most sleep and memory consolidation studies exclude individuals on antidepressant medication due to the potentially confounding effects they have on sleep (Wichniak, Wierzbicka, Walecka, and Jernajcyk, 2017). Such an investigation, therefore, is yet to our knowledge to have been carried out and has the potential to be informative.

***The exact contributions of intermittent hypoxia vs sleep fragmentation to memory deficits***

The MASC-IT study is well-situated to help answer this question because if no recovery is observed despite prolonged relief, it would suggest that memory deficits are a result of the damage caused by intermittent hypoxia. As discussed at length above, OSAS is "a chronic condition in which repetitive upper airway collapse results in intermittent hypoxia and sleep fragmentation, making it the perfect lens through which to explore the interaction of a clinical sleep disorder with memory" (Ahuja et al., 2018). It has to date, however, been difficult to tease the two apart, leaving it unclear which mechanism is responsible for which impairment, and whether intermittent hypoxia and sleep fragmentation are additive and/or synergistic.

In individuals with OSAS, impairments in memory thought to be caused in part by intermittent hypoxia have been attributed to reduced cell neurogenesis and density of the hippocampus (Bartlett, Ray, and Thompson, 2004), the frontal cortex (Hatipoglu and Rubinstein, 2004) and gray matter (Macey, Henderson, and Macey, 2002, in Bucks et al., 2013). The only attempts to isolate the specific contribution of intermittent hypoxia to memory deficits in OSAS come from the animal literature. In one study (Gozal, Daniel, and Dohanich, 2001, in Dempsey et al., 2010), researchers exposed rodents to either 2 weeks of air patterns designed to model OSAS, hypoxia in the absence of OSAS, or control. The authors found cognitive impairments in only the hypoxia condition. In another similar study (Decker, Hue, Caudle, Miller, Keating, and Rye, 2003), hypoxia exposure was associated with spatial memory impairment in newborn rats.

The specific contribution of sleep fragmentation towards memory deficits in OSAP seem to be trickier to isolate. Bucks and colleagues (2013) reviewed the available

evidence on the topic and suggested that sleep fragmentation “reduces the efficacy of restorative processes in the prefrontal cortex leading to cellular and biochemical stress” (Horne, 1993; Maquet, 1995). These stresses, in turn, “disrupt functional homeostasis, altering glial and neuronal viability, contributing to memory dysfunction” (Daurat, Foret, and Bret-Dibat, 2008) via the slowing of cognitive processing (Verstraeten, 2007; Torun-Yazihan, Aydin, and Karakas, 2007). The necessity of at least some degree of sleep continuity was demonstrated by Djonlagic, Sabiosky, and colleagues (2017). The authors tested procedural learning in OSAS patients against a carefully matched control group (only reported differences were AHI and arousal index and oxygen nadir) and found a significant poorer performance in the OSAS group. Importantly, the observed difference between groups was predicted by the arousal index rather than oxygen saturation, providing compelling evidence for the influence of sleep fragmentation in humans which is independent from the influence of intermittent hypoxia.

A sensible approach towards isolating the contribution of sleep fragmentation to memory deficits in OSAS might be to compare performance on memory tests in an evening/morning paradigm between an OSAS group and a carefully matched control group who experience frequent arousals in the absence of intermittent hypoxia. A candidate group here might individuals with Periodic limb movements (PLMs, discussed above). Not only could comparing OSAS patients and individuals with PLMs potentially inform the discussion on the separate contributions of intermittent hypoxia and sleep fragmentation on memory consolidation, longitudinally reassessing these patient groups could potentially provide evidence as to whether CPAP restores impairments specifically caused by fragmented sleep, hypoxia, or both. However, great care would be necessary to ensure comparable arousal indexes across both groups.

***Which model of cognitive harm in OSAS is most viable?***

By longitudinally measuring memory consolidation across treatment timepoints , the MASC-IT study could inform which model of cognitive harm in OSAS is most viable. Evidence of quick memory renormalization support the former, a lack of recovery supports the latter. On the one hand, cognitive impairments in OSAS could theoretically be the consequence of sleep disruption and the resulting daytime somnolence/attentional deficits. This would suggest that memory and executive deficits in OSAS are attributable

to these daytime sleepiness/attentional deficits and are therefore secondary to them. If this model is to be considered a dominant explanation for cognitive deficits in OSAS, then CPAP therapy ought to benefit daytime somnolence, attention, memory, and executive functions. This model would also imply that the degree of relief from daytime somnolence should be associated with any restoration of cognitive function (Buck et al., 2017).

The second class of theoretical model, again outlined by Bucks and colleagues (2017), takes the position observed deficits in OSAS are attributable to brain damage (e.g., vascular changes, neural damage, and cell death). Candidate mechanisms for this explanation include chronic intermittent hypoxia (Xu et al., 2004, in Bucks et al., 2017) or blood-gas abnormalities and cerebral homeostatic changes as a result of fragmented sleep (e.g., Beebe, 2005, in Bucks et al., 2017). This account would be supported by evidence of damage to brain regions associated with the aspects of cognition patients demonstrate deficits in.

### *What do we stand to learn about sleep and memory more generally?*

The MASC-IT study finally stands to further inform our understanding about sleep and memory more generally. Whilst it wouldn't necessarily be ground-breaking to observe further evidence of impaired sleep being associated with memory deficits, or even restored sleep being associated with a degree of normalisation, it would be of interest to better understand: (1) the degree of memory renormalisation following acute sleep restoration; (2) whether one night of sleep restoration is enough to significantly impact upon memory consolidation; (3) how much continuous sleep is necessary to constitute a restoration of sleep (i.e., what level of adherence is necessary to observe a therapeutic effect), and (5) which sleep parameters are most plausibly implicated in these processes and what this means for dominant theories of sleep and declarative memory consolidation.

## Conclusions

In sum, CPAP therapy is highly efficacious and is associated with a wide range of positive long-term benefits, and a handful of well-designed, convincing studies demonstrate a long-term, positive benefit for CPAP in both declarative and non-

declarative domains. Notable failures to find a positive benefit of CPAP on memory in OSAS patients are best explained by methodological issues as opposed to the lack of a measurable effect of treatment. The field would be benefitted from increased use of evening/morning designs and more specific task choices, as well as more consistent reporting of treatment compliance and better understanding of what constitutes adequate CPAP use in studies aimed at assessing memory in OSAS patients.

The MASC-IT study seeks to settle the debate in the literature as to the influence of CPAP on declarative memory consolidation by applying the above methodological suggestions longitudinally to a sample of treatment-naïve adults with a recent diagnosis of OSAS. To carry out the MASC-IT study, two methodological studies were carried out aimed at empirically justifying the number type of task best suited for the study. It was found that novel word learning (verbal declarative) and object location (non-verbal declarative) are ideally situated for use in sleep and memory studies. As well as this, it was found that words more closely related established entries in the lexicon have a general advantage in terms of encoding over more distal novel words, as well as consistently producing significant consolidation effects. Finally, no evidence was found that applying two different tasks within the declarative memory system would interfere with the expression of consolidation. The proposed project will offer a valuable insight into the influence of chronic sleep fragmentation, the extent to which recovery is possible from chronic sleep fragmentation, how much relief is necessary, and which model of cognitive harm in OSAS is most viable

# Chapter 3

# Assessing the unique contributions of sleep duration and number of awakenings to measures of affect and declarative memory.

## Abstract

The duration and continuity of sleep are both thought to be important for the governance of declarative memory and affect. Little is known about the unique contribution of sleep continuity in humans since controlling for sleep duration is challenging. We conducted an observational pre-sleep/post-sleep memory study online among new parents and students (N=519) to investigate whether the continuity of sleep predicts the encoding and overnight forgetting of declarative information as well as the next-day expression of positive and negative affect whilst controlling for the influence of sleep duration. Participants learned spatial locations of familiar objects (e.g., *paw, cupcake*) and their memory was tested immediately (i.e., pre-sleep) and the following day (post-sleep) to measure encoding and overnight forgetting respectively. They estimated their total sleep duration and number of nocturnal awakenings for each night preceding completion of an experimental session and for a typical night across the previous month and completed the Positive and Negative Affect Scale (PANAS) in the post-sleep session. When controlling for sleep duration, the number of nocturnal awakenings predicted encoding and positive and negative affect (albeit with small effects). Sleep duration did not predict encoding or affect and neither sleep duration nor the number of awakenings predicted overnight forgetting. However, when encoding error was added into the multilevel model testing the relationship between awakenings and overnight forgetting as a control variable, a small significant positive relationship was found. When we replaced awakenings with a variable capturing categories of Parent status, we strikingly found that all categories of parenthood were associated with increased encoding error relative to not having a child. However, Parent status did not contribute to the explained variance in Overnight forgetting, Positive affect, or Negative affect. These results are consistent with the possibility that continuity of sleep is associated with the encoding and consolidation of declarative memories and the expression of affect independently of sleep duration in human adults. These results also

suggest that encoding deficits in new parents persist many years after the birth of the youngest child and may never renormalise. However, effect sizes were small and further research is required.

## Introduction

It is well-documented that sufficient sleep is essential to sustain various aspects of human performance and quality of life. Insufficient sleep is also associated with many severe negative outcomes including cardiac problems (Gradner et al., 2016), obesity (Spiegler et al., 2015), dementia (Sabia et al., 2021), psychiatric disturbance (Krystal, 2012), and millions of lost working hours for each of the world's leading economies (Hafner et al., 2016). Even in less extreme circumstances, poor sleep is associated with daytime somnolence (Brown et al., 2020), poor affect (Goldstein and Walker, 2014), low alertness (Kilgore, 2010), and impaired memory (Rasch and Born, 2013). Yet, individuals are increasingly attempting to keep up with the demands of modern life by cutting down on sleep, with the recent 'Need for Sleep' study (Milling et al., 2022) surveying over 4000 UK adults and finding that 71 percent slept less than the recommended 7-9 hours per night (Consensus Conference Panel, 2015). There is also evidence that sleep in the modern era is becoming increasingly fragmented, with numerous studies reporting shallow sleep, cardiovascular reactivity, and cognitive deficits among individuals living near noise nuisances such as busy roads, railway stations, airports and wind turbines (Tassi et al., 2010; Morsing et al., 2018). Some groups such as shift workers (Schwartz and Roth, 2006) are frequently at risk of mild chronic exposure to disrupted sleep whilst others such as new parents (McBean et al., 2016; Pawluski et al., 2016) experience a relatively sudden and acute dose of sleep disruption during the postpartum period. Given the clear societal imperative to better understand the deleterious impact of poor sleep, this paper is focussed on the relationship between sleep, declarative memory, and affect. To do so, we are capitalising on natural variation in sleep restriction in the parent population and considering the unique contributions of total sleep time (as a proxy for sleep deprivation) and number of nocturnal awakenings (as a proxy for sleep fragmentation). We are also exploring potential associations between parenthood and deficits in declarative memory and affect.

### *Sleep and declarative memory*

Human sleep consists of three core sleep stages: NREM stage 2 sleep; NREM stage 3 sleep (slow-wave sleep; SWS) and rapid eye-movement (REM) sleep, which alternate in a cyclic manner. SWS and REM predominate during the early and later parts of the sleep period, respectively (Rasch and Born, 2013). Awareness of the relationship between sleep and memory extends at least as far back as Ebbinghaus (1885) who observed less forgetting when participants slept between observations and since then there have been numerous studies robustly showing that sleep promotes the formation and maintenance of memory across multiple domains, including (but not restricted to) declarative memory (e.g., Plihal and Born 1997; 1999). Declarative memories are defined as ones that are explicit (as opposed to implicit) and are broadly categorised as being episodic (memories of what, where and when in relation to an event or episode) or semantic (memories of facts and figures) (Camina & Guell, 2017).

There are two contemporary models that account for a relationship between sleep and declarative memory, suggesting that sleep actively supports both the encoding and the longer-term consolidation (strengthening of a newly formed or otherwise vulnerable memory representation) of new declarative information: the synaptic homeostasis hypothesis (Tononi and Cirelli, 2014), and the active systems consolidation hypothesis (Born and Wilhelm, 2012). The synaptic homeostasis hypothesis has become a highly influential model implicating sleep in memory formation (See Tononi and Cirelli, 2014, for review). The synaptic homeostasis hypothesis suggests that sleep is “the price we pay for waking plasticity” (Tononi and Cirelli, 2020) and is essential towards regulating synaptic strength. The key premise is that during wake we are constantly learning and that this comes at a cost (namely the costs associated with a net increase in synaptic strength, e.g., energy). Towards the end of the day, these respective neuronal networks approach saturation to the point that sleep is required to reset the balance via a process called synaptic down-selection. Without sleep and synaptic down-selection, the encoding of new material would be impaired. In keeping with this model, rodent research using electron microscopy has shown that over wake there is a net increase in synaptic strength and over sleep there is a reduction (e.g., Vivo et al., 2017).

On the other hand, the active systems model states that when newly encoded memory representations are tagged as being of future use (e.g., emotionally salient stimuli, Hu et al., 2006; items of high future relevance, Wilhelm et al., 2011; items intended to be remembered, Rauchs et al., 2011), sleep plays an active role in transforming those representations from being labile and vulnerable-to-forgetting to being strong and enduring. This is proposed to occur via repeated reactivations of new memory traces, a product of a delicate synchrony of oscillations which take place in NREM sleep and serve the purpose of relieving the hippocampus of its duty to temporarily store the representation and trigger plasticity and ultimately facilitating long-term storage and availability.

It must also be noted that alternative accounts also exist purporting to explain the clear association between sleep and declarative memory consolidation. Yonelidas et al., (2019) proposed a Contextual Binding account of consolidation in which the hippocampus is continually involved in shaping and reshaping memories over time in accordance with new contextual information. This account differs from SHY and Active Systems in that for Contextual Binding, forgetting is interference, and SWS is an excellent protector from interference. Sleep fragmentation may be a useful measure in testing some of these core assumptions of Contextual Binding. For example, if SWS serves the function of protecting vulnerable memory traces to contextual interference, then it is not intuitive that extremely brief, intermittent awakenings in which the individual is highly unlikely to remember the episode (see Winser et al., 2013) could be regarded as sufficient towards promoting contextual interference.

Spatial memory tasks are thought to be sensitive to declarative memory deficits (e.g., Wallace and Bucks, 2012 meta-analysis on cognitive impairments among obstructive sleep apnoea syndrome patients). Rudoy, Voss, Westerberg, and Paller (2009) taught participants object locations before sleep and paired each object with a sound (e.g., dog/bark). During slow-wave sleep, the sounds for half of the objects were played and memories for these object locations upon waking were found to be stronger relative to non-cued objects in 83% of participants (N=12). Variants of this object location task have been used in several studies since (Cairney et al., 2016; Guttesen et al., 2022; Lewis et al., 2021), offering support for both the active systems model generally and this

task is a robust and sensitive measure of declarative memory storage in neocortical areas (see Rasch and Born, 2013).

Collectively, the above models of sleep and memory suggest that both deprived and fragmented sleep should have implications for learning and consolidation. However, studies have rarely considered the role of sleep fragmentation in isolation of total sleep time. Considering sleep deprivation, Newbury, Crowley, Rastle, and Tamminen (2021) recently carried out meta-analyses on the effects of total sleep deprivation (i.e., no opportunity for sleep) before and after learning. The researchers considered 45 papers in which the effect of sleep deprivation before learning and 31 papers after learning, with each study having at least one full night of sleep deprivation in healthy adults. They found that sleep deprivation significantly impairs learning of declarative information both at the point of initial encoding and when tested after a period of consolidation. Notably however, the effect size for sleep deprivation on learning was considerably larger than the effect size for sleep deprivation on consolidation. Corroborating this, Guttesen and colleagues (2022) compared overnight forgetting for items learned in a visuospatial memory task (N=30 young adults) following sleep and following one night of total sleep deprivation and found significantly reduced performance following sleep deprivation. The authors also found a significant effect of sleep deprivation on encoding which was larger than the consolidation effect they detected. Drummond et al., (2000) also found that in humans, one night of sleep deprivation impaired post-sleep word learning in healthy adult volunteers and that impairment was associated with reduced temporal lobe activity, consistent with the active systems model.

However, there is an important distinction to be made between studies of sleep deprivation (i.e., no sleep) and those that use more naturalistic partial sleep deprivation regimes (around 4-5 hours of time in bed, e.g., Cousins et al., 2018). The literature is less clear on the effects of partial sleep deprivation, and this is important because this is the kind of sleep deprivation the general population is likely to encounter in day-to-day life. In their review, Cousins, and Fernandez (2019) argued that declarative learning and consolidation are often unaffected by several nights of partial sleep deprivation. This is not to say that partial sleep deprivation cannot affect learning and consolidation. For example, Cousins et al., (2018) compared learning of items in a picture encoding task when adults had received 9 hours in bed versus 5 hours in bed and found impaired

encoding for the latter relative to rested controls. Therefore, more research is required to better understand the effects of more naturalistic forms of sleep deprivation.

Another key gap in the literature is that there is little evidence on the influence of sleep fragmentation on encoding and consolidation. In one exception, Ficca and colleagues (2000) found that one night of sleep fragmentation (waking participants up 40 minutes after each NREM sequence began) significantly impaired verbal recall of word-pairs learned prior to sleep in healthy young adults. Consistent evidence is also available from the animal research literature. For example, Tartar and colleagues (2006) found that waking rodents every 2 minutes and 30s across a 24-hour sleep interruption protocol significantly impacted the encoding of spatial memory as demonstrated using a water maze task known to be hippocampus-dependent. Similarly, Wallace and colleagues (2015) found that interrupting sleep in rodents every 2 minutes for 72 hours impacted the encoding of spatial memory. Importantly, the authors also ruled out the potentially confounding effect of stress hormones by measuring levels throughout and finding no increase during acute sleep fragmentation. In humans, Hrubos-Strom, and colleagues (2012) found mild impairments in the encoding of verbal memory in 290 adults at elevated risk of obstructive sleep apnoea (a sleep disorder characterised by several awakenings each hour in which the individual wakes in response to airway collapse). The participants were considered at high-risk owing to their scores on the Berlin Questionnaire, which assesses complaints of excessive daytime sleepiness and snoring. These participants are likely to have been experiencing fragmented sleep since their symptoms suggest that they are experiencing frequent upper-airway collapse during sleep.

A particular challenge in researching the influence of sleep fragmentation is separating its effects from the potentially confounding influence of sleep duration (often referred to as total sleep time). Whilst there is little research in this area conducted in humans and in the domain of declarative memory formation and consolidation, there is again some informative work in the animal literature. One example comes from Lee et al (2016). They used a forced desynchrony protocol to fragment sleep whilst ensuring that sleep duration remained intact. Specifically, they delivered 11-hour light/11-hour dark phases for 12 days to artificially adapt the circadian rhythms of mice. The result of this protocol is that sleep duration, NREM, and REM time remained intact, but NREM and REM sleep became fragmented (i.e., shorter bouts with more stage transitions).The

authors reported impairment in a hippocampus-dependent fear conditioning task in mice with altered 22-hour circadian rhythms compared to controls. Another study from Lui and colleagues (2019) directly addressed the question of whether sleep fragmentation makes a unique contribution to learning deficits compared to sleep duration and found convincing evidence for the existence of a serotonin-modulated sleep fragmentation circuit in fruit flies (who share a similar diurnal sleep pattern to mammals). By stimulating what they refer to as the sleep fragmentation circuit, the flies started exhibiting fragmented sleep without changes in total sleep time until the manipulation stopped. The authors also established a causal link between activation of the sleep fragmentation circuit and impairments in aversive associative memory: By increasing the activation of this circuit to trigger the learning impairment and then restoring sleep architecture pharmacologically they observed a renormalisation of memory function.

This investigation relies on subjective measures and therefore is not intended to directly affirm or refute any of the dominant theories of sleep on its own. It is intended to be a starting point towards entrenching sleep fragmentation firmly within the thinking of those seeking to extend models of sleep and declarative memory consolidation by showing that sleep fragmentation may also unique be able to influence memory consolidation.

Taken together, these findings offer support for the following: (1) sleep and declarative memory formation and consolidation are strongly associated, with some evidence available to suggest a causal link; (2) both sleep fragmentation and sleep deprivation can have deleterious effects on the encoding and consolidation of declarative memory if the disruption is adequate; (3) it is difficult to tease apart fragmentation and deprivation, but some promising work has been carried out in the animal literature, and; (4) a gap exists in the literature with regards to investigating the unique contributions of each form of sleep disruption in humans.

### *Sleep and affect*

A large body of literature also supports a relationship between sleep and affect, with both low affect and poor sleep transdiagnostic features of psychiatric disturbance (Walker and Goldstein, 2014, Benca et al., 1997). However, as for declarative memory,

we are again lacking understanding of the unique contributions of sleep deprivation and fragmentation. Addressing this is important to be able to identify risk markers and therapeutic targets. Thus, a secondary aim of this study is to assess the unique contributions of sleep deprivation and sleep fragmentation to measures of affect.

The term affect describes what we feel and is reflected in our mood and emotions and can extend to our preferences (Rosenberg, 1998). Affect is conceptualised and measured in the two broad domains of positive and negative (Bradburn, 1969; Watson & Tellegen, 1985). Within each domain exists several emotions (e.g., enthusiastic, proud, and inspired are regarded as examples of positive affect and irritability, fear, and distress would be regarded as examples of negative affect). A rich literature exists supporting the broad proposition that sleep disruption alters and impairs affect, and an emerging literature is beginning to better understand how different properties of the sleep cycle influence each domain of affect respectively.

At its most general, the literature suggests that loss of sleep increases self-reported negative affect and decreases self-reported positive affect (Cote et al., 2019; Khan et al., 2013). Tomasso and colleagues (2020) carried out three meta-analyses looking at the influence of sleep on mood, emotion, and emotion regulation. They pooled 241 effect sizes from 64 studies and found this same pattern. They also reported that the association between positive affect and sleep loss is a large effect and the association between sleep loss and negative affect is moderate. Importantly, they also found that the difference in effect size between sleep loss and positive affect vs negative affect only existed under conditions of total sleep deprivation. Current models of sleep and positive affect implicate sleep in regulating dopamine and its role in reward processing (Krause et al., 2017). Notably, sleep deprivation is argued to reduce the activity of dopamine receptors (Brower et al., 2001; Wong et al., 2004; Michaelides et al., 2012). With regards to negative affect, Walker, and Goldstein (2014) put forward the REM sleep emotion calibration model. According to this model, there exists an emotional salience network including the Amygdala, Locus Coeruleus and Prefrontal Cortex which relies on noradrenaline to function. The key argument is that REM sleep restores noradrenaline levels for next-day functioning of the emotional salience network, and sleep disruption in REM impairs the activity of this network.

Offering broad support for these models, Shen et al (2018) measured sleep quality and duration via questionnaire and mood via the Positive and Negative Affect Scale (PANAS) in 4582 adolescents finding that different types of sleep disruption have different consequences for positive and negative affect. Specifically, they observed that levels of sleep duration similar to those in partial sleep deprivation studies were associated with the expression of less positive affect, and poorer self-reported sleep quality (known to be related to sleep fragmentation, e.g., Conte, 2021) was associated with greater expression of negative affect.

Considering the influence of sleep continuity on affect more directly, Finan and colleagues (2015) compared sleep restriction, forced awakenings, and rest on subjective ratings of positive and negative affect. The manipulations took place across 3 nights over an 8-hour nocturnal phase each night. In the forced awakenings group, no sleep was allowed for 1 of the 8 hours determined at random. For the remaining 7 chunks of 1 hour, they were awoken and kept awake for a randomly determined 20 minutes of each hour. In this condition no-one slept for more than 280 minutes. In the sleep restriction group, pairwise matching was used such that participants were given the maximum total sleep time of their allocated match in the forced awakenings groups, except with no interruption. A control group slept for 7-8 hours each night without disruption. The authors reported that forced awakenings (i.e., sleep fragmentation) had a greater detrimental effect on positive emotion than sleep restriction (i.e., partial sleep deprivation). The authors also reported that the limited expression of positive affect in the forced awakenings group was statistically accounted for by a greater reduction of SWS in this group relative to the restriction group. The groups were equivalent in their negative mood. Finan et al (2017) went on to show that this effect can emerge after just one night of sleep continuity disruption. The authors suggested that the effect of disrupted SWS on positive affect is independent of negative affect, which seems to be regulated by REM sleep, and that the disruptive effect of SWS loss may be gradual following partial sleep deprivation and that fragmenting sleep on top of that can expedite the process. It should be noted that whilst the authors carefully matched for total sleep time, their design did not permit a true comparison of the unique contributions of sleep deprivation and sleep fragmentation to the positive and negative affect since the forced awakenings group also underwent partial sleep deprivation (sleeping a maximum of 280 mins each evening, which is 140-260 minutes short of recommended sleep duration).

In sum, there exists some evidence that sleep disruption impairs positive (and to a lesser extent negative) affect with emerging evidence on underlying mechanisms, but that a necessary step to advance the literature is to isolate the contributions of fragmentation versus sleep duration. Isolating the exact contributions of sleep fragmentation and sleep duration is challenging and further research is required. Specifically, fresh approaches to controlling for the effect of sleep duration on affect in humans are needed to complement the existing experimental literature. One promising candidate is to control for sleep duration statistically using multilevel models.

### *Sleep disruption in the postpartum period*

A rich literature exists demonstrating that new parents, particularly new mothers, experience disrupted sleep in the postpartum period, which is generally shorter, lighter, and more fragmented (Bernt et al., 2014. Also see McBean et al., 2016; Kenny et al., 2020; McBean and Montgomery-Downs, 2013; Quin et al., 2022; Tikotzky et al., 2022, and Insana et al., 2011). This is due to a variety of reasons, most notably the need to feed and care for the infant at night (Astbury et al., 2022). For example, despite considerable variability, McBean, Kinsey, and Montgomery-Downs modelled postpartum sleep disturbance in the first 6 months following giving birth and found that breastfed newborns wake for feeding approximately 2.9 times each night and are tended to for an average of 33.9 minutes each awakening. Given that 50-73% of new parents report both types of sleep disruption in the postpartum period (Quin et al., 2022), the trajectory of sleep restoration is gradual (Insana et al., 2013), and can last as long as 6 years (Kenny et al. 2020), this population is ideal for our investigation. Speaking to that variability, Bruni and colleagues (2014) reported that nocturnal awakenings vary both within each cross-section and longitudinally across infancy. In addition to our analyses on the associations between sleep and memory/affect more generally, we are also interested to see whether new parenthood (with a youngest child age of 6 or less) is associated with poorer declarative memory and affect relative to healthy controls.

### *The current investigation*

The current investigation aims to investigate and thus complement and extend the sleep and declarative memory and sleep and affect literature in two key ways. Firstly, the available evidence for the unique influence of sleep fragmentation on the encoding and consolidation of declarative information needs to be complemented by research with human participants. To our knowledge, our investigation is the first to isolate the unique contributions of sleep duration and sleep fragmentation in this area using adult human participants. We aim to achieve this by utilising a pre-sleep/post-sleep spatial declarative memory task online amongst a large enough sample to satisfy the power requirements of large multilevel models. By doing so, we can statistically quantify the unique contribution of sleep duration to our dependent variables of interest (namely encoding, consolidation, positive affect, and negative affect, respectively) and then do the same for sleep fragmentation *whilst controlling for sleep duration.* Our task choice (object location) has the benefit of being a well-established measure of declarative memory that is sensitive to the effects of sleep and has a sizable existing animal literature for comparative purposes. As well as this, our approach potentially complements and adds robustness to a field dominated by highly controlled objective measurements amongst small sample sizes. Finally, our approach offers a complementary approach to that of Finan and colleagues (2015) by examining natural variation in sleep deprivation/fragmentation as opposed to manipulating it. Arguably then, the present results may have greater ecological validity. Taking an individual differences approach requires a sample that has rich natural variation in sleep duration, sleep fragmentation, and affect. To achieve both an adequate sample size and sufficient natural variation in sleep disturbance, we draw both from the undergraduate pool and from new parents. In doing so, we also hope to extend the literature on sleep, memory, and affect in the postpartum period by investigating whether distinct categories of parenthood determined by youngest child age are associated with deficits in declarative memory and affect. Doing so will empirically address widely held anecdotally driven and intuitive beliefs that new parents suffer deficits in these areas.

## Methods

### *Hypotheses*

The aims expressed above can be broken down into the following hypotheses:

**H1.** Increased number of self-reported nocturnal awakenings will predict increased Encoding error, increased overnight forgetting, less expression of positive affect, and more expression of negative affect whilst controlling for sleep duration.

**H2**. Based on the reasoning that the key differences between new parents and healthy controls are lower sleep duration and increased number of nocturnal awakenings, we expect to see increased Encoding error, increased Overnight forgetting, less expression of positive affect and more expression of negative affect in new parents whilst controlling for sleep duration since we argue that the effects of sleep fragmentation are at least to some degree independent of sleep duration.

Several analyses were left as exploratory. Namely, from the available literature it was unclear whether sleep duration would predict encoding, overnight decay, positive affect, or negative affect. It was also unclear which measurements of sleep duration and number of awakenings would make for the best predictors in our models. Hence, we measured sleep duration for the night before the completion of each session and for a typical night across the previous month. Similarly, we measured the total number of nocturnal awakenings for the night before the completion of each session and for a typical night across the previous month as well as the number of awakenings for each of these that lasted 5 minutes or longer.

### *Participants*

**Sample.** Participants were adults aged 18 or over who reported being free from a diagnosed sleep disorder. As a guide, an a priori power calculation was carried out using G*Power 3.1 for a hierarchical multiple regression with 6 predictors with $\alpha=0.05$ at 80% power. This power calculation was carried out to detect small effects ($r^2=0.02$) and suggested that 688 participants would be required. We used this figure as a target and

recruited as close to this number of participants as possible within our recruitment window and with the intention of stopping recruitment when this target had been met. Inclusion in the final analyses required completion of both sessions in the appropriate time frame (with Session 2 being completed within the calendar day following completion of Session 1) and with only one recorded attempt at Session 1. A total of 1085 participants completed the first questionnaire, with participant retention and exclusion criteria resulting in a total sample of 705 participants who completed the first session and a final sample of 519 eligible participants completed both sessions in the appropriate timeframe and were included in the remainder of the analyses. Table 11 (below), describes the characteristics of this sample:

*Table 11: Means, standard deviations, and proportions of sample demographics for parents, non-parents, and the full sample.*

| Characteristic | Parents | Non-Parents | All |
|---|---|---|---|
| Age (years) | M = 35.97 (SD = 6.66) | M = 22.41 (SD = 5.46) | M = 29.50 (SD = 9.13) |
| **Sex assigned at birth** | | | |
| Male | N = 26 (9.49%) | N = 56 (22.49%) | N = 82 (15.68%) |
| Female | N = 248 (90.51%) | N = 193 (77.51%) | N = 441 (84.32%) |
| **Employment** | | | |
| Part-time | N = 90 (32.85%) | N = 11 (4.40%) | N = 101 (19.27%) |
| Full-time | N = 82 (29.93%) | N = 44 (17.60%) | N = 126 (24.04%) |
| Unemployed | N = 35 (12.77%) | N = 13 (5.20%) | N = 48 (9.16%) |
| Student | N = 10 (3.65%) | N = 182 (72.80%) | N = 192 (36.64%) |
| Maternity/Paternity | N = 39 (14.23%) | N = 0 (0.00%) | N = 39 (7.44%) |
| Homemaker | N = 18 (6.57%) | N = 0 (0.00%) | N = 18 (3.44%) |
| **Education** | | | |
| High School | N = 73 (26.64%) | N = 166 (66.40%) | N = 239 (45.61%) |
| University/college | N = 97 (35.41%) | N = 61 (24.40%) | N = 158 (30.15%) |
| Postgraduate | N = 83 (30.29%) | N = 22 (8.80%) | N = 105 (20.04%) |
| Specialised vocational training | N = 21 (7.66%) | N = 1 (0.40%). | N = 22 (4.20%) |
| **Partnership** | | | |
| Yes | N = 242 (88.32%) | N = 45 (18.00%) | N = 287 (54.77%) |
| No | N = 45 (11.68%) | N = 205 (82.00%) | N = 237 45.22%) |
| **Youngest Child Age** | | | |

| | |
|---|---|
| 0-2 | N = 101 (36.86%) |
| two-six | N = 145 (52.92%) |
| 6+ | N = 27 (9.85%) |

*Table 12: Means, standard deviations, and proportions of sample demographics for each category of parents.*

| Characteristic | YCA 0-2 | YCA 2-6 | YCA 6+ |
|---|---|---|---|
| Age (years) | M = 33.33 (4.41) | M = 36.12 (5.11) | M = 44.59 (11.69) |
| **Sex assigned at birth** | | | |
| Male | N = 7 (6.93%) | N = 12 (8.28%) | N = 7 (25.92%) |
| Female | N = 94 (93.07%) | N = 133 (91.72%) | N = 20 (74.07%) |
| **Employment** | | | |
| Part-time | N = 27 (26.73%) | N = 60 (41.38%) | N = 3 (11.11%) |
| Full-time | N = 20 (19.80%) | N = 48 (33.10%) | N = 14 (51.85%) |
| Unemployed | N = 10 (9.90%) | N = 16 (11.03%) | N = 8 (29.63%) |
| Student | N = 1 (0.99%) | N = 9 (6.21%) | N = 0 (0.00%) |
| Maternity/Paternity | N = 39 (38.61%) | N = 0 (0.00%) | N = 0 (0.00%) |
| Homemaker | N = 4 (3.96%) | N = 12 (8.28%) | N = 2 (7.41%) |
| **Education** | | | |
| High School | N = 19 (18.81%) | N = 43 (29.65%) | N = 11 (40.74%) |
| University/college | N = 37 (36.63 %) | N = 52 (35.86%) | N = 8 (29.62%) |
| Postgraduate | N = 37 (36.63%) | N = 42 (28.97%) | N = 4 (14.81%) |
| Specialised vocational training | N = 8 (7.92%) | N = 8 (5.52%) | N = 4 (14.81%) |
| **Partnership** | | | |
| Yes | N = 93 (92.08%) | N = 126 (86.90%) | N = 23 (85.19%) |
| No | N = 8 (7.92%) | N = 9 (13.10%) | N = 4 (14.81%) |

Data collection took place online between April 2020 and April 2021. New parents were predominantly recruited by contacting primary schools and asking them to advertise the study using their notice boards and/or mailing list. New parents were also recruited by requesting permission to advertise the study in closed parent groups, by creating Instagram and twitter posts with relevant hashtags (e.g., "#Nightfeeds), and by requesting permission from moderators of relevant subreddits (e.g., r/daddit) to engage with their communities. Finally, new parents were recruited by searching hashtags on twitter (e.g., #mum) and advertising the study to individuals whose accounts allowed message requests. Our non-parent participants were mainly recruited from within the University of York's participation credit scheme. It was important to safeguard against attrition so participants were asked to volunteer their mobile phone number so that they could be reminded to complete the second session. As well as this, participants were informed that this would enable them to be entered into a prize draw for a £50 Amazon voucher, since they would otherwise be anonymous. Participants who elected to receive a text message reminder received a text message in the morning of the second session with a direct link to the study. Text messages were sent in bulk and scheduled so as to avoid waking up participants from outside of the UK. 69.7% of participants elected to receive a reminder text message. For payment, participants were entered into a prize draw for a £50 Amazon voucher or received course credit. All participants provided informed consent.

See Figure 10 (below) for a visualisation of the study timeline:

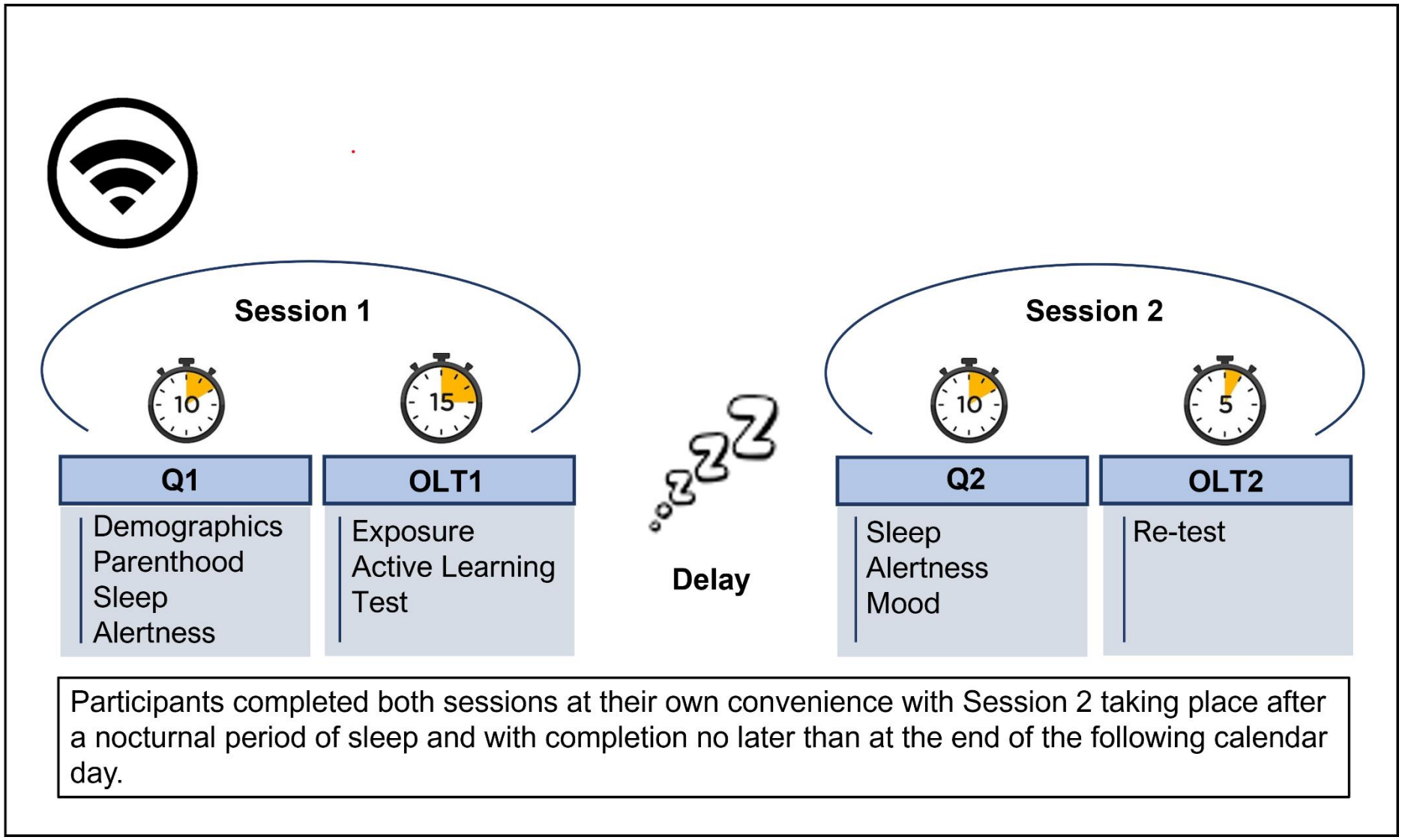



*Figure 10. Experimental procedure.*

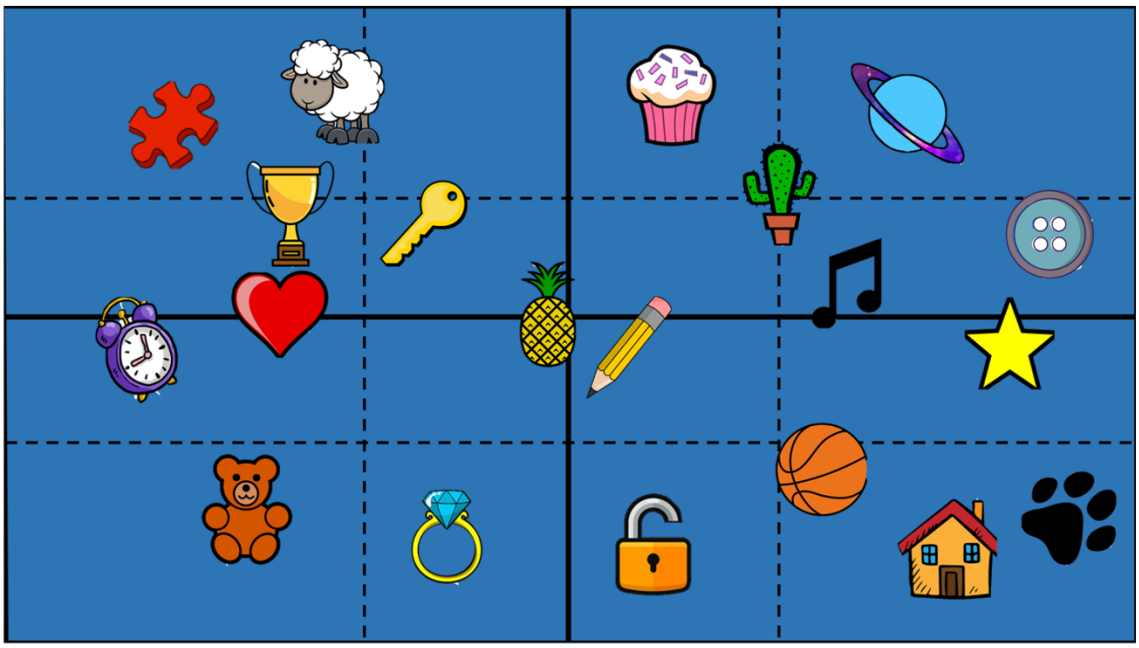
### *Stimulus Presentation*

Participants completed the study on a device of their choosing. The experimental task was programmed using Psychopy (Pierce et al., 2019), was hosted online using Pavlovia (Bridges, Pitiot, MacAskill, and Peirce, 2020), and was compatible with smartphones, laptops, and desktop computers. The task was coded in normative screen units to ensure compatibility across devices and numerous test runs were carried out among a wide variety of devices to detect bugs of which none were found. All screens are 2 units wide and 2 units high. For the encoding dataset, 54.3% completed the experiment using a smartphone, 42.0% of participants completed the experiment using a laptop, and the remaining 3.6% completed the experiment using a tablet.

Participants were exposed to 20 objects within the object location task, each presented one at a time against a reference grid (see Figures 11 and 12, below). Participants viewed the objects from varying distances based on their own preference. The objects were all simple, bold, animated and with glossy bold colours and thick black outlines to enhance presentation quality across potential devices. Each object was contained within a 0.2355 x 02355 area in normative units. Each object had its own unique coordinates in normative screen units. These coordinates were originally generated randomly using excel within boundaries that ensured that the full object would appear on the screen and then slightly tweaked to avoid clustering. Finally, care was taken to avoid objects sharing the same obvious semantic category, e.g., farm animals.

**Visual stimuli.** The 20 objects from the object location task, in order of presentation at exposure. Each object had one correct location and objects were presented in random order during active learning, test, and re-test.

*Figure 11. Visual stimuli for the Object Location task.*



**Screen locations.** Approximate screen locations for each of the 20 objects in the object location task. Objects were presented one at a time.

*Figure 12. Screen locations for Object Location task.*

***Procedure***

This experiment was conducted online and consisted of two sessions (test/re-test) separated by at least a period of nocturnal sleep and with the second session occurring no later than the end of the calendar day following the first session (see Figure 10, above). Each session began with participants completing an online survey via Google Forms and clicking the link in the confirmation message of the form submission which took them to an object location task (~25 mins total for session 1 and ~15 mins total for session 2) hosted by Pavlovia (Bridges et al., 2020). The questionnaires allowed a measure of several independent variables (e.g., number of nocturnal awakenings, sleep duration) and two dependent variables (positive affect and negative affect) whilst the object location task allowed measures of declarative memory encoding and overnight forgetting, our primary dependent variables. At each stage of the experiment participants disclosed either their full mobile telephone number or at least the last 5 digits of it, allowing us to

match their questionnaire responses to their object location data. Session 2 (below) shows start times for each session for all participants who completed the study:

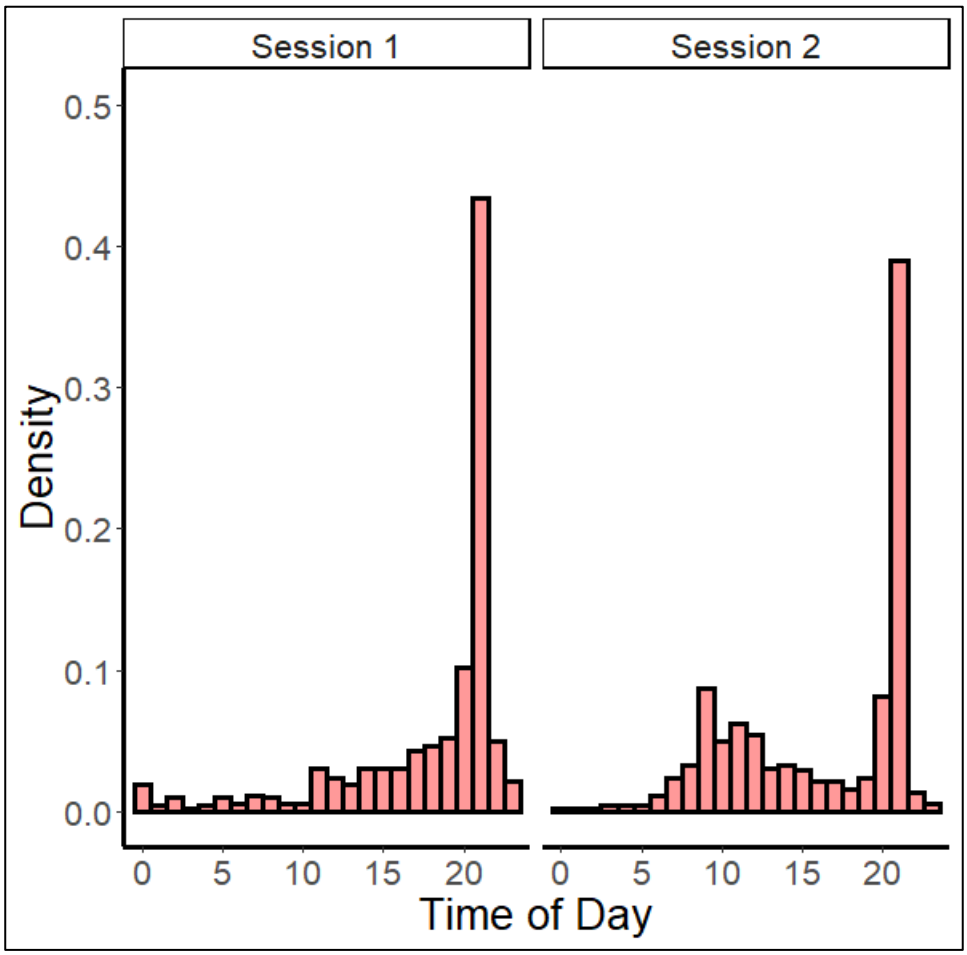



*Figure 13. Density plots for session start times. X axis = hour of day (0-24), Y axis = Density. N=519.*

**Session 1.** Participants were told that it would be ideal if they completed the first session in the evening. However, Session 1 began at a time of day of the participants convenience (see Figure 13 above). Participants would begin by accessing a dedicated website created to walk participants through the study independently. The website contained a brief introduction to the study and contact information for the research team. Session 1 began by clicking the hyperlink to Session 1 at which point the participant would be brought to the first questionnaire hosted by Google Forms. Consent, unique ID, demographic, parenthood, sleep, and alertness information was collected at this stage (for details see section on questionnaires below) and typically lasted around 10 minutes. Participants then completed an exposure phase in which they passively viewed each object presented one at a time against the reference grid. This was followed by an active

learning phase, in which participants would try to correctly identify the correct location of each object (and if they did so twice in a row this object would be eliminated from further training). Finally, in a final test phase each object was presented once, one at a time, and the participant clicked or touched the screen where they thought the object's location was. The duration of the object location task varied as a consequence of participants' proficiency in the active learning phase, with the average completion time being around 15 minutes (range: ~7-20 mins).

**Delay Period.** Participants went about a typical night of nocturnal sleep in the delay period. Participants who disclosed their full mobile phone number were sent text message reminders at 9am in their local time-zone containing a link for the second session of the experiment.

**Session 2.** Participants were told that it would be ideal if they completed the second session in the morning. However, Session 2 began at a time of day of the participants convenience (see Figure 4 above).Session 2 began with an online survey which again collected information on sleep, alertness and mood (as detailed below), and was followed by the object location task retest (taking roughly 10 minutes to complete). The retest involved participants being exposed to the object once, one at a time in the centre of the screen and they had to click the mouse or touch the screen where they believed the object's true location was. This phase lasted no more than 5 minutes.

**Session 1 Questionnaire.** The first questionnaire was focussed on collecting information on background demographics, including gender, age, parenthood status, as well as collecting data on sleep and alertness.

***Demographics.*** Participants were asked if they had been diagnosed with Restless Legs Syndrome or a sleep disorder, e.g., Sleep Apnoea. Declaring a diagnosis of either of these would result in the resulting data not being included in the final analyses. Participants then disclosed: their age in years; the first half of their post or zip code; whether or not they lived with a partner (*Yes/No)*; what their employment status was (*Full-time employment, Part-time employment, Maternity/Paternity leave, Student, Temporary employment, Full-time volunteer, Part-time volunteer, Unemployed, Other*); what the highest level of education was that they had achieved (*No education,*

*Primary/Elementary, High school/Secondary, Specialised vocational training, University/college graduate, Postgraduate, Other*), and; their sex assigned at birth (*Female, Male, Don't know, Prefer not to say*). Parenthood data was captured by firstly asking whether or not they were a parent (*Biological mother, Biological father, Step/adoptive/foster mother, Step/adoptive/foster father, not a parent, other).* If the participant indicated that they were a parent, they then disclosed: whether they lived with their child full-time (*Yes, Part-time, No*); How many children they had, and; what the age of their youngest child was (*Years and months)*.

***Sleep.*** Self-reported data was collected to measure sleep duration and number of nocturnal awakenings. Participants were asked: how much sleep they would get in a typical 24 hour period in hours and minutes (to the nearest 15 minutes) based on their sleep patterns over the last month; How long in hours and minutes (to the nearest 15 minutes) they sleep for the previous night; how long they had napped for that day in hours and minutes (to the nearest 15 minutes); how many times they are likely to wake up in a typical night of sleep based on sleep patterns across the previous month; how many of these awakenings would last 5 minutes or more; how many times they woke up the previous night, and; how many of these awakenings lasted 5 minutes or longer (the approximate requisite amount of time for an awakening to be remembered the following day, Winser et al., 2013).

***Alertness****.* Alertness was measured using the Stanford Sleepiness Scale (Hoddes, Dement, and Zarcone, 1972), and as such participants described how they felt currently (1=*Feeling active, vital, alert, or wide awake; 2=Functioning at high levels, but not at peak; able to concentrate; 3=Awake, but relaxed, responsive but not fully alert; 4=Somewhat foggy, let down; 5=Foggy, losing interest in remaining awake; 6=Sleepy, woozy, fighting sleep, prefer to lie down; 7=No longer fighting sleep, sleep onset soon, having dream-like thoughts)*.

**Session 2 Questionnaire. *Sleep****.*** The second set of sleep questions was again designed to capture information about sleep duration and number of nocturnal awakenings and was focussed on sleep in the delay period. Participants were asked: how much night-time sleep they had had since the first session in hours and minutes (to the nearest 15 minutes); how much nap time they had had since the first session in hours

and minutes (to the nearest 15 minutes); how many times they woke up the previous night, and; how many of these awakenings had lasted 5 minutes or longer.

***Alertness.*** Alertness was again captured using the Stanford Sleepiness Scale, as in the first session.

***Affect.*** Affect was measured subjectively using the Positive and Negative Affect Schedule Short Form (PANAS-SF; Watson, Clark, and Tellegen, 1988). This measure splits mood into positive and negative affect and asks participants for a list of 20 items to indicate the extent to which they had felt that way over the last week (*Interest; Distressed; Excited; Upset; Strong; Guilty; Scared; Hostile; Enthusiastic; Proud; Irritable; Alert; Ashamed; Inspired; Nervous; Determined; Attentive; Jittery; Active; Afraid).* The value for Cronbach's alpha was reported as α=0.86-0.90 for the positive affect scale and α=0.84-0.87 for the negative affect scale by Watson and Tellegen in 1988.

**Object location task.** The object location task was an adaptation of one used by Rudoy, Voss, Westerberg, and Paller (2009). The number of items was reduced to 20 to maximise participant retention whilst maintaining statistical power. The items were animated as opposed to photographic to maximise picture quality across devices. The task was also described to the participant as a game, for the purposes of retention and enhancing the participant's experience. As such, a scoreboard was included for the active learning phase to reinforce the idea that the goal was to eliminate all the objects. Finally, several minor changes were made including changing the colour scheme of the reference grid to be more visually appealing across devices, outlining the shape in red or green to provide intuitive feedback and to block off the active learning phase into rounds with the option of breaks.

The object location task began in Session 1 with a short round of exposure (~2 minutes). In this phase, participants were instructed to passively view the presentation of each of the 20 objects, one at a time, in their correct locations against the reference grid. The objects were presented in the order indicated by Figure 11. The participant triggered the exposure phase by clicking the mouse or touching the screen when ready, the first object was presented after 0.5s, each object was presented for 2s with 0.5s between each presentation.

Participants were instructed that they were about to play the “Object memory game” and their goal was to eliminate each object. Each object would be presented one at a time, in the centre of the screen, and the participant had to click or touch the area of the screen that they thought the object belonged to. Active learning began 0.5s after the participant clicked their mouse or touched the screen, after which each object was presented in the screen for up to 4s, at which point the trial would be deemed incorrect. After each trial, participants would receive feedback in the form of the object just presented being shown in its correct location against the reference grid for 1.5s. The object would be thickly highlighted in either green (correct) or red (incorrect). Unknown to the participant, during each trial an invisible rectangle surrounded each object (0.4167 x 0.4167 normative screen units, in line with Rudoy et. al., (2009). This error zone occupied 4.34% of the screen). This area represented the correct zone, a zone within which if the participant clicked or touched the screen, they would be treated as having correctly identified the location of the object. An object would there be “eliminated” if the participant correctly responded within this zone for a particular object twice in a row, at which point the scoreboard would reflect that the object had been eliminated and that the participant was one point closer to completion. Participants were also instructed that they had 10 rounds with which to eliminate all of the objects and complete the game. Each round contained each object, minus any that had been eliminated, and presented in random order. At the end of each round, the participant had to click or touch the screen again to signify that they were ready to continue, giving the opportunity for short breaks if needed. This phase lasted approximately 15 minutes, but completion varied considerably, ranging from approximately 4-20 minutes.

The first session concluded with a short test (~5 minutes) in which the participant would again be presented once, one at a time, with each of the 20 objects. After clicking the mouse or touching the screen 0.5s would elapse, at which point each object would be presented in the centre of the screen until the participant clicked the mouse or touched the screen where they thought the correct object location was. There was a pause of 0.5s between each trial and no feedback was given. The second test phase was identical to test phase 1.

## Statistical Analysis

### *Models*

The current investigation used subjective sleep data measured across two days along with scores from an object location task carried out across two sessions and two consecutive calendar days. As well as this, demographic data and data relating to the family status of the participant (e.g., parent status, whether in a partnership, youngest child age) were used. This data was used to predict several different outcomes measures across several similar models. The data collected was used to predict encoding (declarative), overnight decay (declarative), positive affect, negative affect. Supplementary analyses were also carried out using the data collected to predict sleep duration and number of nocturnal awakenings. Since it was sought to better understand the unique contributions of each independent variable (most notable the unique contributions of sleep duration and number of nocturnal awakenings), a hierarchical multiple regression approach was used. For analyses which use the object location data, random effects of participant and object were controlled for, making these hierarchical multilevel models. The section below describes the variables used across these respective models.

### *Dependent variables*

The following were response variables in separate models:

**Encoding error**. A mathematical index of distance was used to measure retention of object locations after the first test in session 1. This method has been used in several previous studies (e.g., Rudoy et al., 2009; Cairney, Lindsay, Sobczak, Paller, and Gaskell, 2016). Firstly, the absolute value of the subtraction of the test x-coordinate from the correct x-coordinate is calculated for each object (*S1(x)Distance = ABS(S1X-CorrectX*)) in normative screen units (ranging from -1 to 1 in screen height and -1 to 1 in screen width; a 1x1 normative unit object would occupy 25% of screen area and these values are generated within Psychopy). The same is then carried out for the y-coordinates (*S1(y)Distance = ABS(S1Y-CorrectY*)) for each object before calculating the square root of (*S1XDistance*S1XDistance)* + (*S1YDistance*S1YDistance)* . The resulting

value (*S1Distance*) is one index of encoding and is in normative screen units, with higher values representing a lower amount of success in the participant's attempt to remember the object location of all the objects. Mean indices of encoding error per participant in our study ranged from 0.07 (almost no error) to 0.89 (almost no accurately encoded information).

**Overnight forgetting.** The same index as for encoding error was carried out for Session 2 test scores and an index of overnight forgetting was calculated by subtracting *S2Distance - S1Distance.* A positive index of overnight forgetting is indicative of memory loss and a negative index of overnight forgetting is indicative of memory improvement (consolidation). Indices of overnight forgetting in our study ranged from mean values per participant of -0.38 (substantial overnight improvement) to 0.63 (substantial overnight forgetting).

**Positive affect.** A positive affect score was calculated by adding the scores on items 1,3,5,9,10,12,14,16,17, and 19 from the PANAS-SF. Scores can range from 10-50 and a higher score corresponds to a higher degree of expressed positive affect.
Negative affect. A negative affect score was calculated by adding the scores on items 2,4,6,7,8,11,13,15,18, and 20 from the PANAS-SF. Scores can range from 10-50 and a higher score corresponds to a higher degree of expressed negative affect.

### *Control variables*

The following group of variables were control variables in separate models:

**Age.** Age was measured in years.
Sex assigned at birth. Participants selected from: *Female; Male; Don't know;* and *prefer not to say*. Dummy coding was used with "Female" as the intercept.
Level of Education. Participants selected from: *No education; Primary/Elementary; High school/Secondary; A levels/Higher/Advanced placement; Specialised vocational training; University/College graduate; postgraduate;* and *Other.* Dummy coding was used with "High School" as the intercept.

**Employment Status.** Participants selected from: *Full-time employment; Part-time employment; Maternity/paternity leave; Student; Temporary employment; Full-time volunteer; Part-time volunteer; Unemployed; and Other.* Dummy coding was used with "Part-time" as the intercept.

**Alertness.** Alertness was measured using the Stanford Sleepiness Scale, an ordinal variable ranging from *1 - Feeling active, vital, alert, or wide awake* to *7 - No longer fighting sleep, sleep onset soon; having dream-like thoughts*.

**Random effects.** Participant ID and stimulus (object name) were included as random effects.

***Independent variables***

The following were predictor variables in separate models:

**Sleep duration.** Subjectively estimated sleep duration was measured in several ways: *sleep duration for the night preceding Session 1 in minutes; sleep duration for the night preceding Session 2 in minutes; sleep duration for a typical night based on the past month in minutes; and each of the first two variables plus any nap time in the preceding 24 hours*. These variables were included together as one hierarchical step except for measures inclusive of nap time, which were removed to address multicollinearity concerns (see below). The variables were included together for practical reasons because we had no factual basis to determine which would be most likely to reveal an effect if it existed.

**Awakenings.** Subjectively estimated number of nocturnal awakenings was measured in several ways: *number of awakenings for the previous night of sleep; number of awakenings for the previous night of sleep which lasted 5 minutes or more; typical number of awakenings for night based on the previous month; typical number of awakenings for night based on the previous month with last 5 minutes or more.* In general, these variables can be categorised as (1) estimates derived from memory (for awakenings lasting 5 minutes or more, Winser et al., 2013) or (2) estimates derived from gist feeling (e.g., Insana et. al., 2013). Our general strategy was exploratory and was to

include all these variables as one hierarchical step, and then re-run the model with only 5-minute awakenings included.

**Parent status.** We collected data on the youngest child age in months and recoded it into the following categories: “No child”, “2 or younger”, “2-6”, and “6+” based on literature suggesting that these were meaningful milestones within parenthood (e.g., significant reduction in insomnia symptoms when youngest child is aged 2, Siversten et al., 2015; complete restoration of pre-pregnancy sleep, Richter et al., 2019). This variable was dummy coded with “No child” as the intercept.

## Results

### *Data preparation*

Unique IDs for questionnaires and object location scores were converted into participant IDs. To ensure that the resultant participant IDs were accurately indexed across tasks and sessions, 100 IDs were randomly generated and closely double-checked. No exclusions were made because of this process. The decision was taken to not remove any outliers to reduce the probability of type I error (Gress, Denver, and Shapiro, 2018).

**Collinearity.** Two variables (‘TST_Total’ - self-reported sleep duration in 24 hours preceding Session 1 inclusive of naps, and ‘A_month’ - self-reported total number of awakenings for a typical night across previous month) were removed due to having VIF statistics greater than 5 (VIF[TST_Total] = 18.2; VIF[A_month] = 6.31).

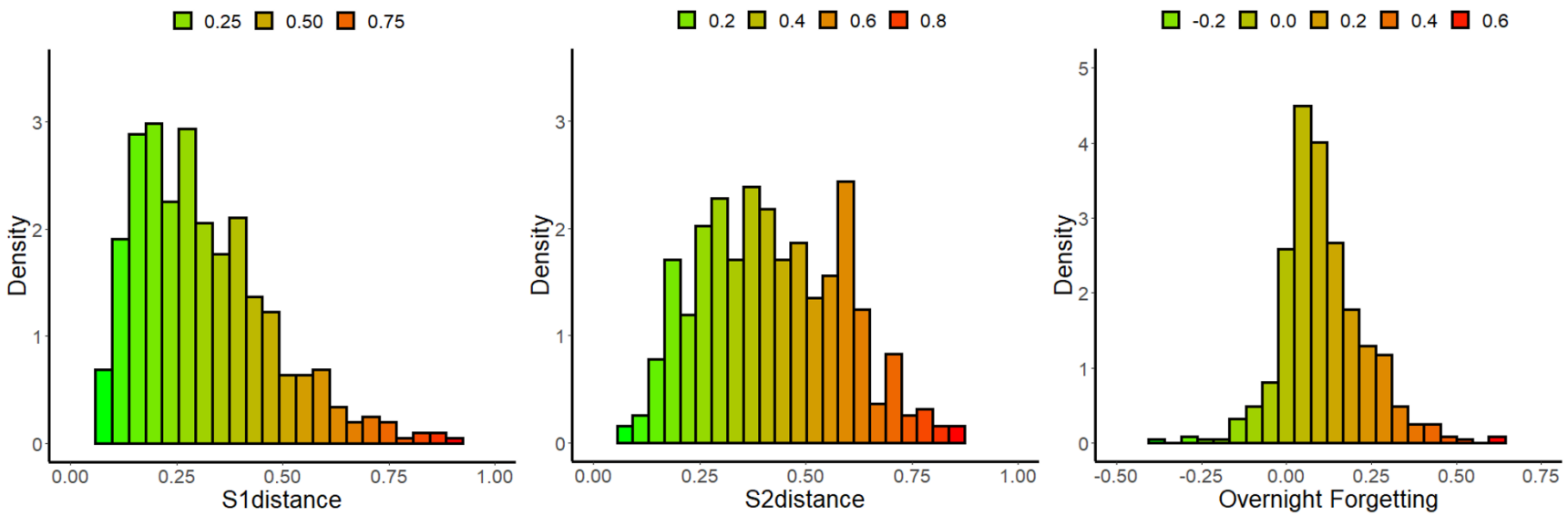
### *Descriptive statistics*

Table 12 below shows means and standard deviations for sleep, alertness, memory, and mood variables, and Figure 14 shows how Encoding error and Overnight forgetting are distributed:

*Table 13. Descriptive statistics.*

| Variable | New Parents | Controls | Total |
|---|---|---|---|
| Overnight forgetting | 0.14 (0.44) | 0.09 (0.41) | 0.11 (0.43) |
| Encoding error | 0.32 (0.32) | 0.26 (0.29) | 0.28 (0.30) |
| Positive affect | 26.90 (7.34) | 28.40 (7.70) | 27.66 (7.56) |
| Negative affect | 20.91 (7.90) | 19.82 (7.13) | 20.29 (7.52) |
| TST (S1) | 395.6 (74.1) | 461.8 (92.9) | 429.7 (90.6) |
| TST (S2) | 402.9 (83.8) | 466.2 (88.6) | 435.8 (91.9) |
| TST (Month) | 401.7 (62.2) | 470.3 (76.2) | 437.3 (77.8) |
| Awakenings (S1) | 2.83 (2.62) | 1.25 (1.63) | 2.00 (2.29) |
| Awakenings (S2) | 2.90 (2.64) | 1.33 (1.65) | 2.07 (2.31) |
| Awakenings+5 (S1) | 1.57 (1.42) | 0.57 (0.92) | 1.04 (1.28) |
| Awakenings+5 (S2) | 1.62 (1.72) | 0.57 (0.96) | 1.07 (1.46) |
| Awakenings+5 (Month) | 1.93 (1.52) | 0.79 (1.09) | 1.33 (1.43) |
| SSS1 | 3.42 (1.24) | 3.10 (1.23) | 3.25 (1.24) |
| SSS2 | 2.90 (1.25) | 2.78 (1.19) | 2.84 (1.22) |

*Note.* Means and Standard deviations objective variables across models. Overnight forgetting = index of decay of object location memory across sessions; Encoding error = index of retention based on S1 test scores; TST(S1) = Subjectively estimated sleep duration for night previous to completion of Session 1 in minutes; TST(S2) = Subjectively estimated sleep duration for night previous to completion of Session 2 in minutes; TST(Month) = Subjectively estimated typical sleep duration based on previous month in minutes; Awakenings(S1) = subjectively estimated number of nocturnal awakenings for night previous to S1; Awakenings(S1) = subjectively estimated number of nocturnal awakenings for night previous to S1; Awakenings(S2) = subjectively estimated number of nocturnal awakenings for night previous to S2; +5 = an awakening lasting 5 minutes or more. SSS1= Stanford Sleepiness Scale scores for Session 1; SSS2= Stanford Sleepiness Scale scores for Session 2.




Encoding error and Overnight retention. *Note.* Mean Encoding error per participant ranges from 0 (no error) to around 1 (very poor memory). Increasing value corresponds with poorer memory. Mean Overnight retention per participant ranges from approximately -0.5 (substantial overnight improvement) to 0.5 (substantial overnight forgetting). Increasing value corresponds with poorer memory.

*Figure 14.* Density plots for encoding error and overnight forgetting.

Figure 15 below shows correlation matrices for objective measures:

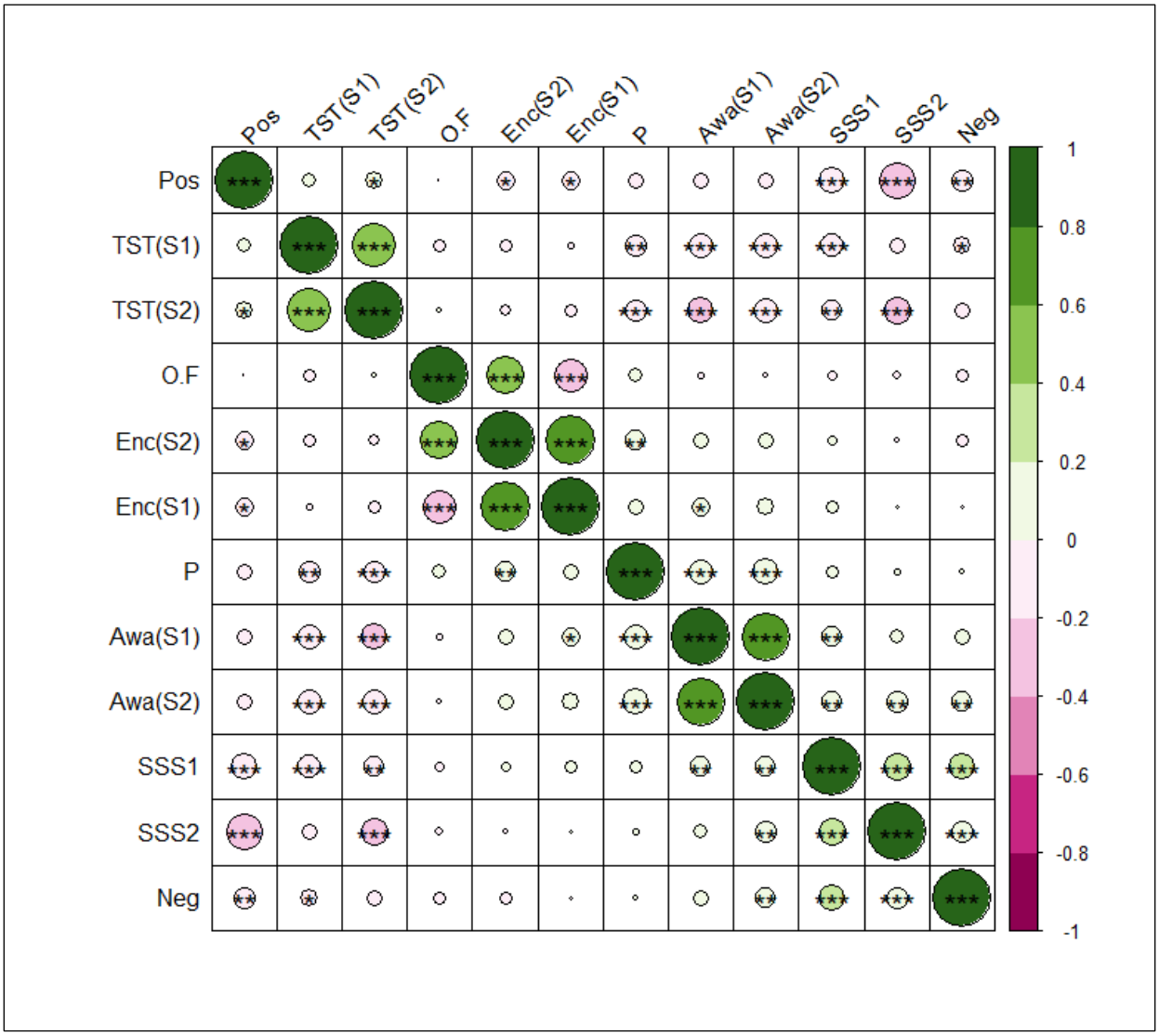



*Figure 15. Correlation matrix for objective measures. The above matrix represents correlations between mean objective measures for the data across both sessions. 'O.F' = Overnight forgetting; 'Enc' = Encoding error; 'TST' = Sleep duration in minutes; SSS' = Stanford Sleepiness Scale; 'Awa' = Total number of awakenings. ' Pos' = Positive affect; 'Neg' = Negative affect. P-values are as follows: * = <0.5; ** = <0.01; *** = <0.001.*

**Hypothesis 1**

Increased number of self-reported number of nocturnal awakenings will predict increased Encoding error, increased overnight forgetting, less expression of positive affect, and more expression of negative affect whilst controlling for sleep duration.

***Does the number of awakenings predict Encoding error whilst controlling for sleep duration?***

This section outlines relevant results for a hierarchical multilevel model which sought to evaluate the relationship between number of awakenings and encoding whilst controlling for sleep duration. For this and all subsequent analyses where the dependent variable is encoding error (i.e., S1Distance), the dependent variable was log-transformed to meet the assumption of normality. All other assumptions were met. Table 13 shows the evolution of the coefficients of key predictors as other variables are added to the model and Table 14 shows model comparison statistics:

*Table 14. Coefficients for hierarchical multilevel model testing predictors against Encoding Error.*

| Predictor | | *b* | SE | *CI* | *p* |
|---|---|---|---|---|---|
| Model 1 | | | | | |
| | TST_month | -0.00 | 0.00 | -0.00 - 0.00 | .492 |
| | TST_S1 | 0.00 | 0.00 | -0.00 - 0.00 | .104 |
| Model 2 | | | | | |
| | TST_month | -0.00 | 0.00 | -0.00 - 0.00 | .937 |
| | TST_S1 | 0.00 | 0.00 | -0.00 - 0.00 | .078 |
| | Awa_month_5 | 0.02 | 0.01 | -0.00 - 0.03 | .052 |
| | Awa_5_S1 | -0.00 | 0.01 | -0.02 - 0.02 | .939 |
| | AWA_S1 | -0.00 | 0.01 | -0.01 - 0.01 | .563 |
| | Awa_month_5 | 0.01 | 0.01 | -0.00 - 0.03 | **.01**** |

*Note.* N=519. Model comparison for hierarchical multilevel model testing predictors against the log transformation of the S1Distance whilst controlling for age, sex assigned at birth, highest level of education, employment status, alertness, participant (random effect) and object (random effect). Sleep duration is a step with 2 variables: TST_ S1 = subjectively estimated sleep duration for the night before Session 1; TST_month = subjectively estimated sleep duration for a typical night of sleep based on the previous month. Awakenings is a step containing three variables: A_5_S1 = subjectively estimated number of nocturnal awakenings lasting 5 minutes or more the previous night; and Awa_month_5 = subjectively estimated number of nocturnal awakenings lasting 5 minutes or more for a typical night based on the previous month; AWA_S1 = total number of awakenings for the night preceding completion of Session 1. The bottom row represents Awa_month_5 being entered into the model separately without the other awakenings variables.

*Table 15. Model comparison, DV =Encoding Error.*

| Model | npar | Test | deviance | Chisq | df | Pr(>Chisq) |
|---|---|---|---|---|---|---|
| Null | 14 | | 5466.1 | | | |
| Sleep duration | 16 | 1x2 | 5463.4 | 2.76 | 2 | .251 |
| Awakenings | 19 | 2x3a | 5456.3 | 7.08 | 3 | .069 |
| Awa_month_5 | 17 | 2x3b | 5456.7 | 6.68 | 1 | **<.01**** |

The data suggests that sleep duration did not significantly predict encoding error. Also, the number of awakenings lasting 5 minutes or longer for a typical night across the previous month seems to predict encoding error whilst controlling for sleep duration. However, the three variables entered as a step did not uniquely contribute to the variance. When only the unique contribution of Awa_month_5 was considered, it significantly predicted encoding error whilst considering sleep duration, with increased awakenings being associated with increased encoding error (see tables 13 and 14 above). The model's total explanatory power is moderate ($R^2 = 0.23$), and the part related to the fixed effects alone (marginal $R^2$) is of 0.03.

***Does the number of awakenings predict Encoding error whilst controlling for sleep duration and affect?***

As outlined above, we ascertained that awakenings predict Encoding error whilst controlling for sleep duration. In an exploratory fashion we then went on to test whether the number of awakenings still predicts encoding whilst controlling for affect. It is important to take this step as the sleep literature often considers the influence of sleep on memory and affect in isolation. Since the variable Awa_month_5 (self-reported number of nocturnal awakenings lasting 5 minutes or more for typical night across previous month) drove the effect shown in Table 13 above, it was entered into the model separately. All assumptions were met. Table 15 shows coefficients at each step and Table 16 shows model comparison statistics:

*Table 16. Coefficients table for multilevel model. DV = Encoding Error (whilst controlling for affect).*

| Predictor | | *b* | SE | *CI* | *p* |
|---|---|---|---|---|---|
| Model 1 | | | | | |
| | Positive | -0.00 | 0.00 | -0.00 - 0.00 | **,03*** |
| | Negative | -0.00 | 0.00 | -0.00 - 0.00 | .470 |
| | TST_month | -0.00 | 0.00 | -0.00 - 0.00 | .617 |
| | TST_S1 | 0.00 | 0.00 | -0.00 - 0.00 | .122 |
| Model 2 | | | | | |
| | Positive | -0.00 | 0.00 | -0.00 - 0.00 | .053 |
| | Negative | 0.00 | 0.00 | -0.00 - 0.00 | .343 |
| | TST_month | -0.00 | 0.00 | -0.00 - 0.00 | .999 |
| | TST_S1 | 0.00 | 0.00 | -0.00 - 0.00 | .086 |
| | Awa_month_5 | 0.01 | 0.01 | -0.00 - 0.03 | **.014**** |

*Note*. Model comparison for hierarchical multilevel model testing predictors against log transformed S1Distance whilst controlling for age, sex assigned at birth, highest level of education, employment status, alertness, positive affect, negative affect, participant (random) and object (random). *p < .05, **p < .01, ***p < .001

*Table 17. Model comparison for analyses reported in Table 15. DV = Encoding Error*

| Model | npar | Test | deviance | Chisq | df | Pr(>Chisq) |
|---|---|---|---|---|---|---|
| Null | 15 | | 5461.1 | | | |
| Sleep duration | 18 | 1x2 | 5458.4 | 2.67 | 2 | .262 |
| Awakenings | 19 | 2x3 | 5452.3 | 6.06 | 1 | **.013*** |

This data further supports the idea that the number of awakenings predicts encoding whilst controlling for sleep duration, even when affect is also controlled for. It is important to note here that the effect of awakenings on encoding disappears when multiple measurements of awakenings are included in the same step. The model's total explanatory power is moderate (conditional $R^2$ = 0.23), and the part related to the fixed effects alone (marginal $R^2$) is of 0.02.

***Does the number of awakenings predict Overnight forgetting whilst controlling for sleep duration?***

This section outlines relevant results for a hierarchical multilevel model which tested the relationship between number of awakenings and overnight decay whilst controlling for sleep duration. All assumptions were met, and no outliers were removed. No effects were significant as shown in Tables 17 and 18:

*Table 18. Coefficients table for multilevel model. DV = Overnight Forgetting.*

| Predictor | | *b* | SE | *CI* | *p* |
|---|---|---|---|---|---|
| Model 1 | | | | | |
| | TST_month | -0.00 | 0.00 | -0.00 - 0.00 | .716 |
| | TST_S2 | 0.00 | 0.00 | -0.00 - 0.00 | .916 |
| Model 2 | | | | | |
| | TST_month | -0.00 | 0.00 | -0.00 - 0.00 | .681 |
| | TST_S2 | 0.00 | 0.00 | -0.00 - 0.00 | .873 |
| | Awa_month_5 | 0.02 | 0.01 | -0.01 - 0.01 | .641 |
| | Awa_5_S2 | 0.00 | 0.01 | -0.01 - 0.02 | .553 |
| | Awa_S2 | -0.00 | 0.00 | -0.01 - 0.01 | .721 |

*Note*. Model comparison for hierarchical multilevel model testing predictors against Overnight Forgetting whilst controlling for age, sex assigned at birth, highest level of education, employment status, alertness, participant (random) and object (random).

*Table 19. Model comparison for analyses reported in Table 17. DV = Overnight Forgetting*

| Model | npar | Test | deviance | Chisq | df | Pr(>Chisq) |
|---|---|---|---|---|---|---|
| Null | 15 | | 13261 | | | |
| Sleep duration | 17 | 1x2 | 13261 | 0.12 | 2 | .931 |
| Awakenings | 20 | 2x3 | 13261 | 0.42 | 3 | .936 |

One possible reason for not finding a relationship between awakenings and overnight forgetting is that in our study, encoding level was allowed to naturally vary. In other comparable lab studies (e.g., Rudoy et al., 2009; Cairney et al., 2016) encoding level was controlled for with a 100% criterion. Therefore, we carried out the same multilinear model as before to test the relationship between awakenings and overnight forgetting (*overnight forgetting ~ age + sex + education + employment + alertness + sleep duration + awakenings + (1 | participant) + (1| object)*) and included encoding error as a control variable. Since the variable Awa_month_5 (self-reported number of nocturnal awakenings lasting 5 minutes or more for typical night across previous month) drove the effect shown in Table 3 above, it was entered into the model separately. The results are displayed in Tables 19 and 20 below:

*Table 20. Coefficients table for multilevel model. DV = Overnight Forgetting (controlling for Encoding Error).*

| Predictor | | *b* | SE | *CI* | *p* |
|---|---|---|---|---|---|
| Model 1 | | | | | |
| | TST_month | 0.00 | 0.00 | -0.00 – 0.00 | .788 |
| | TST_q2 | 0.00 | 0.00 | -0.00 – 0.00 | .623 |
| Model 2 | | | | | |
| | TST_month | 0.00 | 0.00 | -0.00 – 0.00 | .591 |
| | TST_q2 | 0.00 | 0.00 | -0.00 – 0.00 | .513 |
| | Awa_month_5 | 0.01 | 0.01 | 0.00 – 0.02 | **.027*** |

*Note*. Model comparison for hierarchical multilevel model testing predictors against Overnight Forgetting whilst controlling for age, sex assigned at birth, highest level of education, employment status, alertness, encoding error, participant (random) and object (random). *p < .05, **p < .01, ***p < .001

*Table 21. Model comparison for analyses reported in Table 19. DV = Overnight Forgetting (controlling for Encoding Error).*

| Model | npar | Test | deviance | Chisq | df | Pr(>Chisq) |
|---|---|---|---|---|---|---|
| Null | 16 | | 7838.7 | | | |
| Sleep duration | 18 | 1x2 | 7838.5 | 0.24 | 2 | .886 |
| Awakenings | 19 | 2x3 | 7833.6 | 4.88 | 1 | **.027*** |

*Note*. *p < .05, **p < .01, ***p < .001

The data suggests that awakenings predict overnight forgetting whilst controlling for encoding and sleep duration and the effect of awakenings on overnight forgetting is significant regardless of whether multiple measurements of awakenings are included in one step or not. Increasing number of nocturnal awakenings is associated with greater forgetting. The model's total explanatory power is substantial (conditional $R^2$ = 0.62) and the part related to the fixed effects alone (marginal $R^2$) is of 0.55. It should be noted however that this relationship was significant only when including self-reported number of awakenings lasting 5 minutes or more for a typical night across the previous month.

### *Does the number of awakenings predict Positive affect whilst controlling for sleep duration?*

This section outlines relevant results for a hierarchical multiple regression model which aimed to evaluate the relationship between number of awakenings and positive affect whilst controlling for sleep duration. All assumptions were met, and no outliers were removed. Table 21 shows coefficients at each step and table 22 shows model comparison statistics:

*Table 22.* Coefficients table for hierarchical multiple regression. DV = Positive Affect.

| Predictor | | *b* | SE | *t* | *p* |
|---|---|---|---|---|---|
| Model 1 | | | | | |
| | TST_month | 0.01 | 0.00 | 2.04 | **.042*** |
| | TST_S2 | -0.01 | 0.00 | -1.35 | .126 |
| Model 2 | | | | | |
| | TST_month | 0.01 | 0.01 | 1.34 | .182 |
| | TST_S2 | -0.01 | 0.00 | -1.51 | .131 |
| | Awa_month_5 | -0.79 | 0.30 | -2.63 | **<.01**** |
| | Awa_5_S2 | 0.07 | 0.34 | 0.21 | .837 |
| | Awa_S2 | 0.19 | 0.18 | 1.02 | .310 |

*Note*. Model comparison for hierarchical multiple regression testing predictors against Positive Affect whilst controlling for age, sex assigned at birth, highest level of education, employment status, and alertness. *p < .05, **p < .01, ***p < .001

*Table 23. Model comparison for analyses reported in Table 21. DV = Positive Affect.*

| Model | Res.df | df | Test | F | p | $\Delta R^2$ |
|---|---|---|---|---|---|---|
| Null | 507 | | | | | 0.17 |
| Sleep duration | 505 | 2 | 1x2 | 2.14 | .118 | 0.17 |
| Awakenings | 502 | 3 | 2x3 | 2.73 | **.04*** | 0.18 |

*Note*. *p < .05, **p < .01, ***p < .001

The data suggests that sleep duration does not predict positive affect. However, the number of awakenings significantly predicted positive affect whilst taking into account sleep duration, with decreased awakenings being associated with increased expression of positive affect. The model explains a statistically significant and moderate proportion of variance ($R^2$ = 0.23, f(25,493) = 5.93, p < .001, adj. $R^2$ = 0.19).

***Does the number of awakenings predict negative affect whilst controlling for sleep duration?***

This section outlines relevant results for a hierarchical multiple regression model which investigated the relationship between number of awakenings and negative affect whilst controlling for sleep duration. All assumptions were met, and no outliers were removed. Table 23 shows coefficients at each step and table 24 shows model comparison statistics:

*Table 24. Coefficients table for hierarchical multiple regression. DV = Negative Affect.*

| Predictor | | *b* | SE | *t* | *p* |
|---|---|---|---|---|---|
| Model 1 | | | | | |
| | TST_month | -0.01 | 0.01 | -2.26 | **.027*** |
| | TST_S2 | 0.01 | 0.00 | 0.77 | .445 |
| Model 2 | | | | | |
| | TST_month | -0.01 | 0.01 | -1.62 | .106 |
| | TST_S2 | 0.01 | 0.00 | 1.75 | .240 |
| | Awa_month_5 | 0.28 | 0.32 | 0.87 | .386 |
| | Awa_5_S2 | 0.69 | 0.36 | 1.91 | .056 |
| | Awa_S2 | -0.03 | 0.20 | -0.17 | .869 |

*Note*. Model comparison for hierarchical multiple regression testing predictors against Negative Affect whilst controlling for age, sex assigned at birth, highest level of education, employment status, and alertness. *p < .05, **p < .01, ***p < .001

*Table 25. Model comparison for analyses reported in Table 23. DV = Negative Affect.*

| Model | Res.df | df | Test | F | p | $\Delta R^2$ |
|---|---|---|---|---|---|---|
| Null | 507 | | | | | 0.04 |
| Sleep duration | 505 | 2 | 1x2 | 2.56 | .078 | 0.05 |
| Awakenings | 502 | 3 | 2x3 | 3.79 | **.010**** | 0.06 |

*Note*. *p < .05, **p < .01, ***p < .001

Sleep duration did not significantly predict negative affect whilst taking into account awakenings. Number of awakenings did significantly predict negative affect whilst taking into account sleep duration, with increased awakenings being associated with increased expression of negative affect. The model explains a statistically significant and small proportion of variance ($R^2$=0.09, F(502) = 3.14, p < .001, adj. $R^2$ = 0.06). Adding awakenings to the model did significantly contribute to the variance. Since a nonsignificant main effect and a significant f-test for that step suggests multicollinearity, each variable was entered into the model separately as shown in Tables 25 and 26:

*Table 26.* Coefficients table comparing the separate insertion of each subjective measure of awakenings into the hierarchical multiple regression described in Table 23 above. DV = Negative Affect.

| Predictor | *b* | SE | *t* | *p* |
|---|---|---|---|---|
| Awa_month_5 | 0.67 | 0.26 | 2.56 | **.011*** |
| Awa_5_S2 | 0.81 | 0.25 | 3.26 | **<.01**** |
| Awa_S2 | 0.30 | 0.15 | 1.97 | **.048*** |

*Note*. *p < .05, **p < .01, ***p < .001

*Table 27.* Model comparison statistics for the separate insertion each measure of awakenings. DV = Negative Affect.

| Model | Res.df | df | Test | F | p | $\Delta R^2$ |
|---|---|---|---|---|---|---|
| Null | 507 | | | | | 0.04 |
| Sleep duration | 505 | 2 | 1x2 | 2.56 | .078 | 0.05 |
| Awa_month_5 | 504 | 1 | 2x3 | 6.58 | **.011*** | 0.06 |
| Awa_5_s2 | 504 | 1 | 2x4 | 10.63 | **<.01**** | 0.06 |
| Awa_s2 | 504 | 1 | 2x5 | 3.90 | **.048*** | 0.06 |

*Note*. *p < .05, **p < .01, ***p < .001

The data in Tables 25 and 26 suggest that, when entered separately into the model, awakenings are significantly associated with negative affect whilst controlling for sleep duration and significantly contribute to the explained variance. As the number of nocturnal awakenings increase so too do expressions of negative affect.

***Does affect predict Encoding error whilst controlling for sleep duration and the number of awakenings?***

This section outlines relevant results for a hierarchical multilevel model testing whether positive and/or negative affect predicts encoding error whilst controlling for sleep, alertness, and demographics. This is important because there is some evidence to suggest that both positive affect and encoding is governed by SWS (Tononi and Cirelli, 2014; Krause et al., 2017), whereas negative affect is largely governed by REM (Walker and Goldstein, 2014), suggesting that a much stronger relationship might be observed between positive affect and encoding. All assumptions were met, and no outliers were removed.

*Table 28. Coefficients table for hierarchical multilevel model. DV = Encoding Error.*

| Predictor | | *b* | SE | *CI* | *p* |
|---|---|---|---|---|---|
| Model 1 | | | | | |
| | TST_q2 | 0.00 | 0.00 | -0.00 - 0.00 | .610 |
| Model 2 | | | | | |
| | TST_q2 | 0.00 | 0.00 | -0.00 - 0.00 | .486 |
| | Awa_month_5 | 0.01 | 0.01 | 0.00 - 0.02 | **.012*** |
| Model 3 | | | | | |
| | TST_q2 | 0.00 | 0.00 | -0.00 - 0.00 | .576 |
| | Awa_month_5 | 0.01 | 0.01 | 0.00 - 0.02 | **.017*** |
| | Positive | -0.01 | 0.00 | -0.01 – 0.00 | **.023*** |
| | Negative | -0.01 | 0.00 | -0.01 - -0.00 | .082 |
| | Pos*Neg | 0.00 | 0.00 | -0.00 – 0.00 | .120 |

*Note*. Model comparison for hierarchical multilevel model testing predictors against Encoding Error whilst controlling for age, sex assigned at birth, highest level of education, employment status, and alertness, and number of awakenings. *p < .05, **p < .01, ***p < .001

*Table 29. Model comparison for analyses reported in Table 27. DV = Encoding Error.*

| Model | npar | Test | deviance | Chisq | df | Pr(>Chisq) |
|---|---|---|---|---|---|---|
| Null | 15 | | 5466.1 | | | |
| Sleep duration | 17 | 1x2 | 5465.7 | 0.39 | 2 | .821 |
| Awakenings | 18 | 2x3 | 5459.4 | 6.35 | 1 | **.011*** |
| Affect | 19 | 3x4 | 5452.1 | 14.03 | 6 | **.029*** |

*Note*. *p < .05, **p < .01, ***p < .001

The data contained in tables 27 and 28 support the above hypotheses. The relationship between positive affect and encoding error was statistically significant and negative, suggesting that lower encoding error levels were associated with increased positive affect. The relationship between negative affect and encoding error was statistically non-significant. We also entered positive and negative affect as an interaction term in an exploratory fashion. The interaction effect of negative affect on positive affect was not statistically significant. These findings suggest expressions of positive affect are related to encoding error in a way that expressions of negative affect are not. The models total explanatory power is moderate (conditional $R^2$ = 0.18), and the part related to the fixed effects alone is of 0.02.

***Summary (H1)***

Table 29 below summarises the associations reported in this section and shows standardised measures of effect size:

*Table 30. Summary table of standardised coefficients.*

| Predictor | Outcome | Control | β | SE_beta | *p* | Size |
|---|---|---|---|---|---|---|
| Awakenings | Encoding error | Sleep duration | 0.07 | 0.03 | **.010**** | VS - S |
| Awakenings | Encoding error | Sleep duration, Affect | 0.06 | 0.02 | **.014*** | VS |
| Awakenings | Overnight forgetting | Sleep duration | -0.00 | 0.01 | *ns* | *na* |
| Awakenings | Overnight forgetting | Encoding, Sleep duration | 0.04 | 0.02 | **.027*** | VS |
| Awakenings | Positive affect | Sleep duration | -0.12 | 0.05 | **<.01**** | S |
| Awakenings | Negative Affect | Sleep duration | 0.13 | 0.05 | **<.01**** | S |
| Positive | Encoding error | Sleep duration, Awakenings | -0.04 | 0.02 | **.023*** | VS |
| Negative | Encoding error | Sleep duration, Awakenings | -0.01 | 0.02 | *ns* | *na* |

*Note*. *p < .05, **p < .01, ***p < .001. VS – Very Small; S = Small.

**Hypothesis 2**

Parent status will be associated with increased Encoding error, increased Overnight forgetting, less expression of positive affect and more expression of negative affect in new parents whilst controlling for sleep duration.

The following section takes the strategy of repeating the analyses above with the exception that instead of a subjective measure of awakenings we have included a variable which breaks parent status into the following categories: “2 or under”, “2-6”, “6+”, and “No child” as the intercept. This allows for a general evaluation of whether parenthood is associated with differences in declarative memory and affect, but also whether the pattern changes across the time periods the literature suggests being meaningful milestones within the postpartum period and beyond for parents. As before, multilevel models will be carried out for measures of declarative memory and hierarchical multiple regressions will be carried out for measures of affect. Age, sex, level of education, employment status, and alertness will be controlled for throughout. When Overnight forgetting is the dependent variable, level of encoding error will also be controlled for.

***Are there group differences in total sleep time and number of awakenings among categories of parenthood?***

Before carrying out the main analyses in this section, we wanted to test our assumption that new parents would on average sleep less and wake up during the night more often than healthy controls. As such, we carried out hierarchical multiple regressions assessing whether Parent status was associated with minutes slept for a typical night across the past month and the number of awakenings for a typical night in the previous month which lasted 5 minutes or longer. Tables 30 and 31 below show data for the model with sleep duration as dependent variable:

*Table 31. Coefficient table for hierarchical multiple regression. DV = Sleep Duration.*

| Predictor | | *b* | SE | t | *p* |
|---|---|---|---|---|---|
| Model 1 | | | | | |
| | Age | -1.49 | 0.54 | -2.80 | **.005**** |
| | Sex | 13.40 | 9.10 | 1.47 | .141 |
| | Education | -6.86 | 8.68 | -0.79 | .430 |
| | Employment | 22.06 | 13.83 | -0.38 | .111 |
| | Alertness | -10.13 | 2.61 | -3.87 | **<.001***** |
| Model 2 | | | | | |
| | Age | -0.48 | 0.63 | -0.71 | .453 |
| | Sex | 5.26 | 8.83 | 0.60 | .552 |
| | Education | -10.29 | 8.47 | -1.21 | .225 |
| | Employment | -2.69 | 13.75 | -2.02 | .845 |
| | Alertness | -8.94 | 2.53 | -3.53 | **<.001***** |
| | YCA (2 or under) | -80.92 | 12.40 | 2.28 | **<.001***** |
| | YCA (2-6) | -56.56 | 11.20 | 3.02 | **<.001***** |
| | YCA (6+) | -26.74 | 18.77 | 6.53 | .155 |

*Note.* Model comparison for hierarchical multilevel model testing Parent status against sleep duration whilst controlling for age, sex assigned at birth, highest level of education, employment, alertness, and sleep duration. Sleep duration is the self-reported nocturnal sleep duration for a typical night in the previous month in minutes. Alertness is captured by the Stanford Sleepiness Scale. YCA breaks down parent status into categories with "No child" as the intercept. $R^2$ = 0.174, R2 adjusted = 0.156, p <.001*** for Model 1 (null); $R^2$ = 0.247, R2 adjusted = 0.226, p<.001*** for Model 2 (+ Parent Status).

*Table 32. Model comparison statistics for the hierarchical multiple regression described in Table 30.*

| Model | *Res.df* | *df* | *Test* | *F* | *p* | *ΔR²* |
|---|---|---|---|---|---|---|
| Model 1 (null) | 506 | | | | | 0.16 |
| + Parent Status | 503 | 3 | 1X2 | 16.30 | **<.001***** | 0.23 |

The data broadly suggest that our choice of categories for Parent status was justified. Parents with very young children (0-2) slept on average 80.92 minutes less on a typical night across the past month compared to non-parents, parents with children aged 2-6 slept on average 56.56 minutes less, and there was no difference between parents with a youngest child aged 6 or over and non-parents. Tables 32 and 33 below show the data with number of awakenings as dependent variable:

*Table 33. Coefficient table for hierarchical multiple regression. DV = Number of Awakenings.*

| Predictor | | *b* | SE | *t* | *p* |
|---|---|---|---|---|---|
| Model 1 | | | | | |
| | Age | 0.02 | 0.01 | 1.63 | .105 |
| | Sex | -0.60 | 0.17 | -3.66 | **<.001*** ** |
| | Education | -0.20 | 0.16 | -1.29 | .198 |
| | Employment | -0.94 | 0.25 | -0.04 | **<.001*** ** |
| | Alertness | 0.10 | 0.05 | 2.11 | **.035*** |
| Model 2 | | | | | |
| | Age | 0.01 | 0.01 | 1.27 | .205 |
| | Sex | -0.53 | 0.16 | -3.39 | **.001*** ** |
| | Education | -0.21 | 0.15 | -1.34 | .178 |
| | Employment | -0.55 | 0.25 | 2.09 | **.028*** |
| | Alertness | 0.07 | 0.05 | 1.57 | .117 |
| | YCA (2 or under) | 1.42 | 0.23 | 6.29 | **<.001*** ** |
| | YCA (2-6) | 0.44 | 0.20 | 2.17 | **.031*** |
| | YCA (6+) | -0.01 | 0.34 | -0.04 | .970 |

*Note.* $R^2$ = 0.174, R2 adjusted = 0.156, p <.001*** for Model 1 (null); $R^2$ = 0.247, R2 adjusted = 0.226, p<.001*** for Model 2 (+ Parent Status).

*Table 34. Model comparison statistics for the hierarchical multiple regression described in Table 32.*

| Model | *Res.df* | *df* | *Test* | *F* | *p* | $\Delta R^2$ |
|---|---|---|---|---|---|---|
| Model 1 (null) | 506 | | | | | 0.18 |
| + Parent Status | 503 | 3 | 1X2 | 16.01 | **<.001*** ** | 0.24 |

The awakenings data presented here again nicely supports our categorisation of Parent status, with increased awakenings being predicted in the 0-2 and 2-6 category but not the 6+ category. Parents with children aged 0-2 have 1.44 extra awakenings lasting 5 minutes or longer on a typical evening, 0.44 extra for parents with a youngest child aged 2-6, and no significant difference for parents with a youngest child 6 or older. However, the difference in sleep fragmentation is arguably very small between parents and non-parents unless the youngest child is aged 2 or under. Overall, our aim to recruit a sample rich variation in sleep fragmentation was not convincingly achieved since only 19.27% of the total sample and only 38.31% of our recruited parents experience the upper band of observed sleep fragmentation in this study.

***Does parent status predict level of encoding error?***

Tables 34 and 35 below show the results for a multilevel model investigation the association between Parent Status and Encoding error:

*Table 35. Coefficient table for hierarchical multilevel model. DV = Encoding Error.*

| Predictor | | *b* | SE | *CI* | *p* |
|---|---|---|---|---|---|
| Model 1 | | | | | |
| | Age | 0.01 | 0.00 | 0.00 – 0.02 | **.003*** |
| | Sex | -0.05 | 0.06 | -0.17 – 0.07 | .442 |
| | Education | -0.03 | 0.06 | -0.15 – 0.08 | .552 |
| | Employment | -0.06 | 0.07 | -0.19 – 0.07 | .397 |
| | Alertness | 0.02 | 0.02 | -0.01 – 0.05 | .256 |
| | Sleep duration | 0.00 | 0.00 | -0.00 – 0.00 | .276 |
| Model 2 | | | | | |
| | Age | 0.00 | 0.19 | -0.00 – 0.01 | .269 |
| | Sex | -0.02 | 0.00 | -0.12 – 0.11 | .874 |
| | Education | -0.01 | 0.06 | -0.14 – 0.09 | .624 |
| | Employment | -0.02 | 0.07 | -0.15 – 0.12 | .791 |
| | Alertness | 0.02 | 0.02 | -0.02 – 0.05 | .303 |
| | Sleep duration | 0.00 | 0.00 | -0.00 – 0.00 | .118 |
| | YCA (2 or under) | 0.25 | 0.06 | 0.08 – 0.42 | **.004**** |
| | YCA (2-6) | 0.19 | 0.08 | 0.04 – 0.35 | **.013*** |
| | YCA (6+) | 0.26 | 0.13 | 0.00 – 0.51 | **.048*** |

*Table 36. Model comparison statistics for the hierarchical multilevel model described in Table 34.*

| Model | npar | Test | deviance | Chisq | df | Pr(>Chisq) |
|---|---|---|---|---|---|---|
| Model 1 (null) | 15 | | 29340 | | | |
| + Parent status | 18 | 1x2 | 29330 | 9.33 | 1 | **.025*** |

The data suggests being a parent is associated with increased Encoding error regardless of category relative to not having a child. Standardised coefficients are small for parents with younger children (β = 0.08, SE_β = 0.03 for both parents with children 2 and under and between 2 and 6) and very small for parents with children older than 6 (β = 0.05, SE_β = 0.03). The finding of encoding deficits among parents with older children was not expected since it has been suggested in the literature that parents' sleep is completely restored by around the time the youngest child is aged 6 (e.g., Richter et al. ,2019). The models explained 18.4% of the variance, with 2.5% being explained by fixed effects.

***Does parent status predict levels of Overnight forgetting?***

Tables 36 and 37 below show the results for a multilevel model investigation the association between Parent Status and Overnight forgetting:

*Table 37. Coefficient table for hierarchical multilevel model. DV = Overnight Forgetting.*

| Predictor | | *b* | SE | *CI* | *p* |
|---|---|---|---|---|---|
| Model 1 | | | | | |
| | Encoding Error | -0.95 | 0.01 | -0.97 - -0.93 | **<.001*** ** |
| | Age | 0.00 | 0.00 | 0.00 – 0.00 | **.022*** |
| | Sex | -0.01 | 0.02 | -0.34 – 0.04 | .746 |
| | Education | -0.02 | 0.02 | -0.06 – 0.01 | .229 |
| | Employment | -0.03 | 0.02 | -0.07 – 0.01 | .177 |
| | Alertness | 0.02 | 0.01 | -0.01 – 0.01 | .884 |
| | Sleep duration | 0.00 | 0.00 | -0.00 – 0.00 | .680 |
| Model 2 | | | | | |
| | Encoding Error | -0.95 | 0.01 | -0.97 - -0.93 | **<.001*** ** |
| | Age | 0.00 | 0.00 | -0.00 – 0.02 | .162 |
| | Sex | 0.01 | 0.02 | -0.03 – 0.05 | .552 |
| | Education | -0.02 | 0.02 | -0.06 – 0.02 | .299 |
| | Employment | -0.02 | 0.02 | -0.07 – 0.02 | .317 |
| | Alertness | -0.02 | 0.01 | -0.01 – 0.01 | .865 |
| | Sleep duration | 0.00 | 0.00 | -0.00 – 0.00 | .429 |
| | YCA (2 or under) | 0.05 | 0.03 | -0.00 – 0.11 | .064 |
| | YCA (2-6) | 0.04 | 0.03 | -0.01 – 0.09 | .138 |
| | YCA (6+) | 0.02 | 0.04 | -0.06 – 0.10 | .627 |

*Table 38. Model comparison statistics for the hierarchical multilevel model described in Table 36.*

| Model | npar | Test | deviance | Chisq | df | Pr(>Chisq) |
|---|---|---|---|---|---|---|
| Model 1 (null) | 16 | | 8372.8 | | | |
| + Parent status | 19 | 1x2 | 8368.9 | 3.87 | 3 | .276 |

Surprisingly, Parent status was not associated with our measure of Overnight forgetting, even whilst controlling for the level of Encoding error. It seems plausible that an association may exist between parents with very young children (2 and under) and increased Overnight forgetting, given how close this category was to reaching the threshold for significance, but the available data can't support this claim and more research is required. This model explained a large amount of variance (48.7%) with a large amount of that coming from the fixed effects (41.5%). However, most of the explained variance is coming from inserting Encoding error into the model as it is highly negatively related to Overnight forgetting, suggesting that the more you learn, the more you can forget.

***Does parent status predict levels of Positive affect?***

Similarly, to above, Parent status also did not significantly contribute to the variance in Positive affect as shown in Tables 38 and 39 below. The model explained 16.8% of the variance in Positive affect.

*Table 39. Coefficient table for hierarchical multiple regression. DV = Positive Affect.*

| Predictor | | *b* | SE | *t* | *p* |
|---|---|---|---|---|---|
| Model 1 | | | | | |
| | Age | 0.06 | 0.05 | -0.04 – 0.16 | .269 |
| | Sex | 0.24 | 0.88 | -1.49 – 1.96 | .785 |
| | Education | 2.31 | 0.83 | 0.66 – 3.95 | **<.01**** |
| | Employment | 3.61 | 1.33 | 0.99 – 6.24 | **<.01**** |
| | Alertness | -2.43 | 0.26 | -2.94 - -1.92 | **.001***** |
| | Sleep duration | -0.00 | 0.00 | -0.01 – 0.01 | .749 |
| Model 2 | | | | | |
| | Age | 0.07 | 0.06 | -0.05 – 0.20 | .256 |
| | Sex | 0.02 | 0.89 | -1.73 – 1.77 | .982 |
| | Education | 2.25 | 0.85 | 0.57 – 3.93 | **<.01**** |
| | Employment | 3.03 | 1.39 | 0.31 – 5.76 | **.029*** |
| | Alertness | -2.42 | 0.26 | -2.93 - -1.90 | **<.001***** |
| | Sleep duration | -0.00 | 0.00 | -0.01 – 0.01 | .749 |
| | YCA (2 or under) | -1.93 | 1.27 | -4.42 – 0.57 | .130 |
| | YCA (2-6) | -1.45 | 1.14 | -3.68 – 0.79 | .204 |
| | YCA (6+) | -0.01 | 1.89 | -373 – 3.71 | .996 |

*Table 40. Model comparison statistics for the hierarchical regression described in Table 38.*

| Model | *Res.df* | *df* | *Test* | *F* | *p* | $\Delta R^2$ |
|---|---|---|---|---|---|---|
| Model 1 (null) | 505 | | | | | 0.167 |
| + Parent Status | 502 | 3 | 1x2 | 1.10 | .384 | 0.168 |

***Does parent status predict levels of Negative affect?***

And finally, Parent status also did not significantly contribute to the variance in Negative affect as shown in Tables 40 and 41 below. The model explained only 3.6% of the variance in Negative affect.

*Table 41. Coefficient table for hierarchical multiple regression. DV = Negative Affect.*

| Predictor | | *b* | SE | *t* | *p* |
|---|---|---|---|---|---|
| Model 1 | | | | | |
| | Age | -0.05 | 0.06 | -0.16 – 0.06 | .391 |
| | Sex | -1.73 | 0.94 | -3.58 – 0.11 | .065 |
| | Education | 1.58 | 0.89 | -0.18 – 3.34 | .078 |
| | Employment | -0.83 | 1.43 | -3.63 – 1.97 | .561 |
| | Alertness | 0.89 | 0.28 | 0.34 – 1.44 | **.002**** |
| | Sleep duration | -0.00 | 0.00 | -0.01 – 0.01 | .678 |
| Model 2 | | | | | |
| | Age | -0.09 | 0.07 | -0.23 – 0.04 | .185 |
| | Sex | -1.58 | 0.95 | -3.45 – 0.29 | .098 |
| | Education | 1.76 | 0.91 | -0.04 – 3.55 | .055 |
| | Employment | -0.70 | 1.48 | -3.61 – 2.21 | .637 |
| | Alertness | 0.91 | 0.28 | 0.36 – 1.46 | **.001***** |
| | Sleep duration | -0.00 | 0.00 | -0.01 – 0.01 | .695 |
| | YCA (2 or under) | 0.13 | 1.36 | -2.54 – 2.80 | .924 |
| | YCA (2-6) | 1.31 | 1.22 | -1.08 – 3.70 | .283 |
| | YCA (6+) | 1.47 | 2.02 | -2.50 – 5.45 | .467 |

*Table 42. Model comparison statistics for the hierarchical regression described in Table 40.*

| Model | *Res.df* | *df* | *Test* | *F* | *p* | $\Delta R^2$ |
|---|---|---|---|---|---|---|
| Model 1 (null) | 505 | | | | | 0.039 |
| + Parent Status | 502 | 3 | 1x2 | 0.57 | .635 | 0.036 |

Finding that being a parent across all categories of youngest child age was associated with increased encoding error even whilst controlling for sleep duration somewhat supports our assumption that recruiting from this population would greatly enrich the variation in sleep loss in our sample and is very intriguing in and of itself. However, finding that Parent status did not contribute to the variance in Overnight forgetting, positive affect, or negative affect despite finding significant associations between awakenings and these outcomes at a sample level suggests that a group suffering greater levels of sleep disruption would have been a more effective strategy towards better understanding the relationship between sleep fragmentation and declarative memory more generally.

## Discussion

The current study sought to investigate the contribution of nocturnal awakenings to variance in two key outcomes for cognition and mental health that are known to be dependent on sleep: declarative memory and affect. It also explored the associations between parent status and declarative memory and affect. We also sought to do so in a sample anticipated to be rich in awakenings. Our investigation led to two key findings relating to the role of awakenings in declarative memory. First, the number of nocturnal awakenings was positively associated with encoding error whilst controlling for sleep duration, alertness, demographics, and affect (as the number of awakenings increased so too did encoding error). Sleep duration did not predict encoding. Second, neither the number of nocturnal awakenings nor sleep duration predicted overnight forgetting. However, when encoding error was also controlled in an additional exploratory analysis, the number of self-reported nocturnal awakenings lasting 5 or more minutes on a typical night across the previous month did predict overnight forgetting, with increased awakenings being associated with increased forgetting. Relating to affect, it was found on average participants who experienced more nocturnal awakenings expressed lower

positive affect and higher negative affect when sleep duration was controlled. The present study finally found that all levels of parenthood (0-2;2-6;6+) were associated with increased encoding error relative to non-parents but did significantly differ with regards to overnight forgetting, positive affect, or negative affect.

These findings are preliminary and do not include objective measurement of sleep. Despite that clear limitation, they are relatively consistent with the sleep duration literature, to be discussed later, and may be a starting point from which to extend the sleep fragmentation literature. Namely, the finding that a measure of the number of awakenings (here self-reported number of awakenings lasting 5 minutes or longer for a typical night across the previous month) predicted encoding whilst controlling for sleep duration is in line with discussed findings from the animal literature (Lee et al., 2021; Lui et al., 2019) in that the degree of sleep fragmentation was related to learning whilst sleep duration had been accounted for. Our investigation extends these findings to humans and to declarative memory. It is already argued that encoding ability is governed by processes of global synaptic downscaling in SWS (e.g., Tononi and Cirelli, 2014; Born, 2012). These findings are indicative that these processes are benefitted by continuity independent of duration, although need much more support from studies using objective sleep measures. These findings also come with the significant caveat that not all measures of awakenings produced significant effects in the hypothesized direction. All measures of total number of awakenings were not associated with measures of declarative memory, nor were number of 5-minute awakenings in the nights prior to each session.

The finding that the number of awakenings predicted overnight forgetting whilst controlling for sleep duration only when encoding error was also controlled for is partially consistent with the meta-analyses of Newbury and colleagues (2021), who considered a much broader range of declarative memory measures than the current investigation and reported considerably larger effects of sleep disruption on encoding than overnight forgetting. Although the overnight forgetting effect observed here must be treated cautiously, given the exploratory nature of it, as in Newbury and colleagues, it suggests that the association between awakenings and encoding is stronger than the association between awakenings and overnight forgetting. One possibility is that the association with overnight forgetting is stronger under more extreme fragmentation circumstances, and/or

when encoding is more tightly controlled. Indeed, all of the other uses of our measure of overnight forgetting have trained participants to a 100% learning criterion before sleep (Cairney et al., 2016; Rudoy et al., 2009; Guttesen et al., 2022). Our models suggest that the effect sizes for awakenings and encoding error and awakenings (whilst controlling for sleep duration) and awakenings and overnight forgetting (whilst controlling for sleep duration and encoding error) are both very small. Here our decision to prioritise participant retention and rich variation in encoding likely resulted in reduced sensitivity towards measuring overnight forgetting effects, which were already likely to be difficult to capture in a naturalistic environment. In practical terms, having potentially sleep deprived volunteer participants commit to lengthy training regimens online would have been unrealistic. As well as this, we would not have been able to investigate encoding if we had done so since there would have been very little variation to work with.

Sleep duration did not significantly predict any measure of declarative memory in the current study. The literature is currently unsettled on whether sleep duration predicts measures of declarative memory consolidation. Numerous studies show group differences strongly suggesting that sleep benefits declarative memory consolidation relative to wake (e.g., Plihal and Born, 1997; Gais, Lucas and Born, 2006). Cousins et al., (2018) also found that sleep restricted individuals (<5 hours' time in bed) performed poorer on a measure of encoding across 5 night. These differences are arguably and often attributable to the time spent or continuity of SWS and don't necessarily justify the assumption that sleep duration and measures of declarative memory consolidation are related. Backhaus and colleagues (2006) measured sleep using polysomnography and measured declarative memory using a word-pair associates task amonf Insomnia patients and healthy controls. The authors observed significantly lower sleep duration among Insomnia patients, and significantly poorer retrieval, yet no significant correlation between sleep duration and consolidation for either group. The Backhaus study, however, did find that some components of sleep architecture (e.g., time spent in SWS) are related to both sleep duration and declarative memory consolidation, might make the overall pattern harder to interpret. In another study, Scullin (2013) compared pre-post sleep word-pair learning with wake control among healthy young and old adults. In this study, the authors again found the correlation between sleep duration and consolidation to be insignificant but suggests that this was possibly due to ceiling effects. The Scullin (2013) was similar to Backhaus et al (2006) in that SWS was related to consolidation.

Finally, Tucker et al., (2020) also found no differences between nap, resting wake, and wake groups on measures of declarative memory, although argued this likely due to short length of the nap (20 minutes on average vs 60-90 in comparator similar studies).In the current study, It was not clear at the point of recruitment whether our eventual sample would be significantly sleep deprived or not. In the end, our sample was not sleep deprived (Mean(sample) = 431.92 mins; SD(sample) = 90.58; Mean(parents) = 400.77; SD(parents) = 77.21; mins) and therefore the findings of the sleep deprivation literature do not apply. It is plausible that sleep duration will be more strongly related to declarative memory when there is more extreme variation in the sample. An explanation for this might be that SWS is related to measures of declarative memory, and also related to sleep duration, and therefore sleep duration is only related to consolidation to the extent that SWS is disrupted within sleep duration.

Our analyses focusing on the association between subjective number of awakenings and affect were also consistent with previous literature. Firstly, it is well established that sleep fragmentation, especially if chronic, is associated with lower positive affect and higher negative affect (e.g., Bonnet & Arand, 2003), mirroring the findings here. Secondly, our model with positive affect as the dependent variable had substantially more explanatory power than the model for negative affect (adj. $R^2$ of 0.19 vs 0.09 respectively). Since the standardised coefficients for the *positive affect ~ awakenings* ($\beta$ = -0.12) model and the *negative affect ~ awakenings* ($\beta$ = 0.13) model were very similar in magnitude, and the standardised coefficients for alertness differed considerably across models ($\beta$ = 0.13 for the model with negative affect as dependent variable and $\beta$ = -0.38 for the model with positive affect as the dependent variable) this suggests that alertness had far more of an impact on expressions of positive affect than negative affect.

It is important to point out that the effect sizes found in this study are very small. This is not to say that they could not be meaningful however as it is widely known that smaller effect sizes can be meaningful over time (see Princeton and Miller, 1992, for discussion). There is evidence to suggest that individuals who experience small but frequent sleep fragmentation over longer periods of time show subtle but meaningful signs of being negatively impacted. For example, Mucci and colleagues (2020) conducted a systematic review on the relationship between urban noise exposure and psychological

distress and found evidence of increased sleep disorders and cardiovascular problems among residents living near major roads, railways, and airports. As well as this, Lercher and colleagues conducted a telephone survey of over 1600 individuals and found increased sleep medication usage among those living near railway stations and Brink and colleagues (2010) found increased sleep medicine use. Finally, Tassi and colleagues (2013) compared three groups: chronic sleep fragmentation (living more than 10 years near to a railway station), acute sleep fragmentation (one night of comparable sleep disruption in the lab), and control (sleep in a quiet environment); and found little effect of acute fragmentation and decreased psychomotor vigilance and increased daytime sleepiness among those living near railway stations for long periods of time. This suggests that very small amounts of sleep fragmentation, even those smaller than the amount required to affect the individual in one evening, can have a meaningful negative impact if the disruption persists over long periods of time. Therefore, it is possible that even though the effect sizes discovered are small they may be meaningful if experienced chronically. Future research might seek to better understand how impactful these small effects are over longer time courses.

The effect of the number of awakenings on memory and affect in our investigation was driven by awakenings defined as the subjective estimation of the number of awakenings lasting 5 minutes or more for a typical night across the past month. This finding is consistent with work by Winser and colleagues (2013), who measured number of nocturnal awakenings both subjectively and objectively at the homes of participants and found that the average recall threshold for an awakening is 4 minutes and 19 seconds, in line with a rule of thumb in the sleep literature that a nocturnal awakening must last around 5 minutes to be remembered. This is also consistent with a growing literature that the effects of sleep fragmentation are cumulative. For example, Insana and colleagues (2013) conducted a longitudinal field study for in new mothers across the first 12 weeks postpartum and found that psychomotor vigilance continued to get worse consistently across the last 10 weeks of the study and that this was more consistently related with sleep efficiency than total sleep time. As well as this, Tassi, and colleagues (2013) reported little or no change in PVT or subjective ratings of stress, motivation and cognitive impairment in a sample acutely affected by noise disturbance whereas those who lived next to a noisy railway station showed much more pronounced signs of chronic sleep debt across these measures. Therefore, it could be the case that subjective

measures which are based on memory as opposed to gist feeling are more sensitive. It could also plausibly be the case that measuring sleep fragmentation subjectively over a longer time span is a more sensitive approach. However, our approach was exploratory with regards to which variables to include for the number of awakenings. Further research is needed to verify these tentative suggestions.

Finally, finding that all categories of parents on average experienced higher levels of encoding error is intriguing. This is because the results in tables 20-24 agree with the claims of the Richter (2019) study in that sleep seems to be restored by the time the youngest child is 6 or over. Given the weak p-value, small sample size, and wide confidence intervals, great care should be taken in interpreting this finding as it may be a Type 2 error or a confounding influence of different levels of education. Barring these, one very tentative yet intriguing possibility is that whilst sleep is restored, pregnancy-induced brain changes are not (see Martinez-Garcia et al., 2021 for one study finding persisting grey matter volume reduction in parents with youngest children aged 6). Future research might seek to image brain regions associated with declarative memory among parents with youngest children in this age band. The null results for associations between all categories of parents and Overnight forgetting, Positive affect, and Negative affect were surprising. One possibility is that levels of sleep fragmentation seen in new parents aren't materially distinctive enough from non-parents to induce deficits in consolidation and affect. Another possibility, a more likely one given the self-reported experiences of new parents, is that our sample under-represented parents with very young children and with the most severe sleep disturbance.. Indeed, the differences in typical sleep duration and number of awakenings was underwhelming, especially in the 2-6 YCA category. Following from this, it could be the case that a disproportionate number of new parents relatively high in resiliency felt compelled or able to complete what was a relatively taxing study.

The current investigation has several limitations which need to be considered when interpreting the results. First, we relied upon subjective measures of sleep, however, and this is a clear departure from most of the literature on this topic which greatly limits the ability to draw strong conclusions from the results. Literature is mixed as to the strength of association between subjective and objective measures. For instance, Insana and colleagues (2013) reported that subjective measures tend to

underestimate sleep duration (Walsh et al., 2005) and overestimate the number of nocturnal awakenings (Van Dongen et al., 2003). Studies also report weak to moderate associations between objective and subjective measures of sleep duration (with estimates ranging from 0.21 to 0.62, Zavzec et al., 2021). Furthermore, it is possible that subjective estimates of sleep parameters such as number of awakenings and sleep duration are confounded by overall subjective sleep quality. One potential reason for the poor agreement between subjectively and objectively measured awakenings is that individuals do not remember many awakenings lasting less than 5 minutes (Windser et al., 2013). There are also no studies we are aware of that test agreement between awakenings of at least 5 minutes and objective measures. Studies relying on subjective measurement ought to be regarded as complementary to those using objective measures and an important part of a broader collective effort to better understand the influence of sleep on cognition and behaviour. It is clear though that the relationship between sleep fragmentation and declarative memory would benefit from at least some objective sleep measurement, which was outside the reach of the current investigation.

With regards to the possibility that our subjective measures of sleep quality are just reflecting subjective sleep quality, the available evidence is mixed. The main evidence to suggest that the results were driven by subjective sleep quality and not sleep fragmentation is that measures of total awakenings were not significant and the measure most likely to be driven by subjective feeling (number of nocturnal awakenings that were 5 minutes or longer for a typical night across the previous month) was more consistently a significant predictor. Some studies also report associations between subjective sleep quality and amount of N1 sleep (the more fragmented an individual's sleep is, the more N1 sleep they are likely to have, O'Donnell and colleagues, 2009) and between subjective sleep quality and total number of awakenings and awakenings lasting 5 minutes or longer measured using actigraphy (Conte et al., 2022). It could also be argued however that the reason that the number of awakenings and subjective sleep quality are related is because awakenings are one of, if not the main determinants of subjective sleep quality (Zavzek et al., 2020). For example, Harvey and colleagues (2008) combined sleep diaries, free expression conditions and structured interview amongst 25 individuals with insomnia and 28 individuals without disordered sleep and found that subjective sleep quality among these individuals was designed simply as how rested they felt upon waking and the number of awakenings experienced the previous night. So, whilst it is possible that the

observed associations in this study are between subjective sleep quality and memory/affect, it is also arguable that subjective sleep is just a feeling driven by the number of awakenings experienced the night before. Other limitations include that affect was measured using a limited questionnaire. Our investigation would have benefitted from precise measures of how participants were feeling at the point of encoding and retrieval. However, this need had to be balanced against a real concern that recruitment would be challenging and that retention would be sensitive to participant burden. Finally, our eventual sample was dominated by females (85.29%) and our non-parent sample was dominated by students (83.20%).

Finally, whilst we have been tempted to infer causation from correlation at times when interpreting several the observed associations in this study due to the strong theoretical predictions underlying many of the analyses, the current investigation is inherently limited to offering only relational evidence.

There are several interesting future research avenues to explore considering the current findings. In terms of increasing confidence in our measurements, future investigations ought to seek to better understand both the relationship between 5-minute awakenings and objectively measured awakenings and between 5-minute awakenings and subjective sleep quality. It seems plausible that when an individual is asked how many times they woke up the previous night, they are encouraged to use how tired they feel to estimate how many times they woke up because they can't rely on their memory for short awakenings. However, when the individual is asked about 5-minute awakenings, it may be the case that since they can now rely upon their memories (Winser et al, 2013), this measurement will move closer to objective measures and further away from subjective sleep quality, making the 5-minute awakening a valuable measure in naturalistic sleep studies.

In the present study, no associations were found between total number of awakenings and measures of encoding or overnight forgetting. This null result is not sufficient to interpret with any confidence what this means for the Contextual Binding theory of declarative memory consolidation. Future research might seek to tightly control awakenings or recruit samples naturally rich in short, frequent awakenings and to test for associations between these short awakenings and measures of declarative memory.

Such an investigation would optimally be carried out using polysomnography and statistical means of controlling for the confounding influence of longer awakenings should be considered. It may also be useful to consider microarousal instead of awakenings, since the individual isn't even technically waking up. The more brief, intermittent, and seemingly inconsequential the sleep disruption, the harder it is for Contextual Binding to account for (see Schabus et al., 2006 for discussion on the impact of microarousals on sleep spindles).

Future investigations should also seek to explore the relationship between nocturnal awakenings and overnight forgetting in the context of higher levels of sleep deprivation. We demonstrated that the effect is very small in naturalistic conditions where sleep duration is at relatively healthy levels. One way to investigate the relationship would be through use of tightly controlled lab manipulations. However, under such conditions it is hard to effectively control for the potentially confounding effect of sleep deprivation in humans. One solution might be to compare groups high in nocturnal awakenings with greater degrees of sleep deprivation than seen in new parents (e.g., Restless Legs Syndrome or Obstructive Sleep Apnoea Syndrome) with matched controls on spatial memory and control for sleep deprivation statistically (ANCOVA). Sleep deprivation and nocturnal awakenings could be measured both subjectively and objectively (wrist actigraphy) given the lesser sampling demand compared to the current investigation. Increasing the number of trials would further ease this demand. Finally, a 100% encoding criterion would be set given what was observed in the current investigation. However, care must be taken to reconcile results with the confounding influence of atypical dopaminergic activity in Restless Legs Syndrome populations and intermittent hypoxia in OSAS populations (Natarajan, 2010; Ahuja et al., 2018).

Future research might finally seek to complement Lui and colleagues (2019) finding of a serotonin-driven sleep fragmentation circuit in fruit flies by closely investigating behavioural correlates of daytime serotonergic activity in sleep-disrupted humans. The key premise here worth investigating is that if sleep fragmentation is regulated by serotonin, then severe sleep fragmentation might exhaust serotonin reserves required for daytime functioning the following day to the extent that serotonin-driven daytime behaviour will be atypical in fragmented populations (like Goldstein and Walker's REM recalibration model of noradrenaline in emotion regulation, 2014).

Candidate behavioural measures here include the inhibition of negative emotion, eating behaviour, and social aggression (Lucki et al., 1988), maternal neglect (Baxter et al., 2020), and sexual behaviours (Angoa-Perez et al., 2015). Finding atypical daytime behaviour related to serotonergic activity in humans would support the existence of a dedicated sleep fragmentation circuit causally related to learning in humans.

With regards to the postpartum literature, future research ought to pay more attention to measures of individual resiliency as a means to ensuring that the most affected portion of the postpartum population are represented in recruitment. Having done so, it would make sense to revisit some of the null results observed in this study and test for Type 1 error. As well as this, there is at least some tentative evidence now to speculate that there are long-term or even permanent structural and/or functional changes to the postpartum brain which are associated with deficits in the encoding of spatial memories. Future investigations might take an imaging approach to explore this question more fully.

### *Conclusions*

The current investigation sought to better understand the relationships between sleep and declarative memory and sleep and affect. Specifically, the unique contributions of sleep duration and number of awakenings to the explained variance in encoding, overnight decay, positive, and negative affect were tested using hierarchical multilevel models. We also tested associations between Parent status, declarative memory, and affect. Key findings were that self-reported number of nocturnal awakenings predicted encoding, positive affect, and negative affect whilst controlling for sleep duration when sleep duration did not, and that none of our independent variables predicted overnight forgetting. As well as this, we found that positive affect predicted encoding error and negative affect did not. The effect sizes were all small to very small. These results suggest that under natural conditions, the continuity of sleep can have a small impact on next day learning of declarative information even when the duration of sleep is accounted for. It is likely therefore given the surrounding literature that synaptic global downscaling is optimal when uninterrupted. These results also suggest that under natural conditions, the continuity of sleep can have a small impact on expressions of affect. Further, awakenings did predict overnight forgetting whilst controlling for sleep duration when encoding error

was controlled for in the model, leading to the tentative suggestion that a relationship might exist between awakenings and overnight forgetting under appropriate conditions. Finally, we found that all categories of parents, regardless of youngest child age, were associated with increased encoding error, but not related to Overnight forgetting or expressions of affect. One implication of this finding is that the declarative memory deficits we observed in new parents persist for potentially much longer than was predicted given what we know about the trajectory of postpartum sleep restoration. Future research might seek to better understand the relationship between the number of awakenings and overnight forgetting through recruitment of participants who experience more frequent nocturnal awakenings and greater levels of sleep deprivation and through use of a paradigm which trains encoding to a strict criterion. Future research also might seek to investigate the longitudinal impact of these effects over time to better understand how meaningful they are. Finally, there are at least some grounds now to explore whether there are long-term or permanent structural changes associated with the encoding of declarative memories as a result of having a child. This investigation ought to be considered as a preliminary one. Conclusions drawn from the results here are tentative and weighed against the fact that sleep was not measured objectively.

# Chapter 4

# Separating the unique contributions of sleep continuity and sleep duration among a sample of individuals with Restless Legs Syndrome and healthy controls.


## Abstract

Restless Legs Syndrome (RLS) is a highly prevalent sleep disorder, yet little research has been carried out to investigate the impact of RLS on known correlates of sleep disruption. Here we sought to leverage the unique and more severe profile of sleep disruption observed in RLS relative to the sample achieved in Study 3 and to test whether continuity of sleep uniquely contributes to the explained variance in declarative memory and affect whilst controlling for sleep duration. The same pre-sleep/post-sleep online spatial memory paradigm as in Study 2 was used among a sample of unmedicated RLS (N= 64), medicated RLS (N=85) and healthy controls (N=303). As in Study 2, participants estimated their sleep parameters subjectively and completed the Positive and Negative Affect Scale (PANAS) in the post-sleep session. Hierarchical multilevel models, matched t-tests, and matched ANCOVAs were used to test these associations. RLS was significantly associated with increased encoding error. RLS was also associated with increased overnight forgetting but only when the level of encoding was controlled for within the model. Regarding affect, RLS was associated with decreased expression of positive affect but there were no group differences in expressions of negative affect. Taken together, these results are consistent with the idea that the continuity of sleep accounts for independent variance towards the formation and storage of declarative memories, in contexts of lower sleep duration than observed in Study 3. The findings also tentatively suggest that negative affect is left relatively unscathed in RLS. Future research might begin to consider whether qualitative differences exist (for example network corruption due to partial consolidation) in the memory and affective deficits associated with sleep deprivation and sleep continuity.

# Introduction

## *Background*

Restless Legs Syndrome (RLS), also known as Willis-Ekbom Disease, is a common yet poorly understood sensorimotor condition (Guo et al., 2017). Individuals coping with RLS often feel an irresistible urge to move their legs to appease physical discomfort when static with worsening symptoms at night and alongside considerable sleep disruption (see O'Regan and Anderson, 2020; Guo et al., 2017; and Natarajan, 2010, for reviews). Despite a high prevalence (approximately 5-15% in the western world have some degree of RLS, O'Regan and Anderson, 2020) and a clear individual and societal burden attached to it (Trenkwalder et al., 2021), very little psychological research has been carried out on RLS. RLS is a condition with a unique profile of sleep disruption which makes this population a valuable and untapped resource for researchers looking to better understand the association between sleep fragmentation and overnight consolidation processes. We aim to take advantage of the unique characteristics of the disorder to improve understanding of the relationship between sleep and memory more generally, by exploring whether the number of nocturnal awakenings is associated with overnight forgetting of declarative information whilst controlling for the amount of time asleep. This investigation also seeks to extend the RLS literature by establishing what if any associations exist between RLS and declarative memory and expressions of affect.

## *Restless Legs Syndrome*

**Symptoms and epidemiology.** The primary symptom of RLS is an uncontrollable need to relieve uncomfortable sensations in the legs by moving them and in such a way that cannot be accounted for by another condition. The symptoms often peak in the evening (Bogan, 2006) and can even extend to other parts of the body in some cases (O'Regan and Anderson, 2020). RLS has a primary and secondary subtype, with primary RLS occurring much earlier in life and thought to be genetically driven (Natarajan, 2010). Secondary RLS occurs later in life (around the age of 45) with no apparent genetic component and typically accompanied by other comorbidities (e.g., iron deficiency, chronic renal failure; Guo et al., 2017). A large portion (approximately 80-90%) of

individuals with RLS also have periodic limb movements (PLMs) in their sleep and will often sequentially kick with their legs in a jerky fashion upwards of 30 times an hour in more severe cases (O'Regan and Anderson, 2020; Bogan, 2006). There is a strong circadian element to RLS with symptoms beginning in the evening and persisting throughout the first half of the night before relenting for the remainder of the typical sleep cycle and into the early wakeful hours of the morning (Guo et al., 2017). The following day RLS individuals often experience daytime somnolence (Colzato et al., 2021) and these individuals are at an increased likelihood of having a comorbid mood disorder (Becker and Sharon, 2014). The most widely held theory for the cause of RLS, although debated, is that individuals with RLS have deficient iron levels in the basal ganglia (a region of the brain also associated with other motor conditions such as Tourette's Syndrome and Parkinson's Disease, Caligiore et al., 2017; Blandini et al., 2000) which in turn affects dopamine production (Ondo, 2002).

**Sleep disruption in RLS.** Not surprisingly, there exists a reasonable amount of evidence to support the assertion that sleep in RLS is significantly disrupted in most individuals who experience symptoms of it. Sleep disturbance is a common complaint of both types of RLS (Bogan, 2006) The National Sleep Foundation Poll (Phillips et al., 2006, in Bogan, 2006) surveyed 1506 US adults over the telephone and those who reported symptoms of RLS were at a greater likelihood of struggling to fall asleep, sleeping for less than 6 hours, feeling tired the next day, and missing work. As well as this, Allen and colleagues (2005, also in Bogan, 2006) interviewed 16,202 adults and reported that amongst the 416 individuals who reported RLS symptoms 61% reported fragmented sleep, 48% reported prolonged sleep onset latency (SOL), and 48% reported generally a feeling of not sleeping enough.

There have also been several polysomnography (PSG) studies of sleep architecture in RLS. In general, these studies show significantly reduced mean sleep duration (6.4 hours in Hening et al., 2007; 6.3 hours in Sergeeva et al., 2017; 5.9 hours in Cha et al., 2020), significantly increased mean number of nocturnal awakenings (26.8 in Hening et al., 2007; 19.43 in Sergeeva et al., 2017; and a significantly increased arousal index in Cha et al., 2020), increased SOL (Hornyak, 2007), and interestingly little or no difference in REM% and SWS% (Sergeeva et al., 2017, Cha et al., 2020, and

Hornyak et al., 2007 report no SWS% differences, only Hornyak et al., 2007 reported a REM% deficit in RLS).

Geng and colleagues (2022) carried out a systematic review and meta-analyis on the polysomnographic features of RLS. The authoris identified 26 studies and found significantly decreased sleep duration, increased nocturnal awakenings, lower REM %, and no difference in SWS %. . This is important for the current investigation because these are important for declarative memory consolidation, discussed below. Cha, and colleagues (2020) observed lower spindle density and poor coordination between spindles and slow oscillations in RLS relative to healthy controls who were closely matched for Age, Sex, and Education. The authors also reported that spindle power was negatively associated with wake after sleep onset in RLS and that slow oscillation duration was positively related to the arousal index in those with RLS. The authors were the first to extend a generally accepted dysfunction of the thalamus in RLS to the sleep cycle since the thalamus is also involved in spindle generation. Cha and colleagues also contributed to increasing awareness of the possibility that brain dysfunction in RLS causes sleep fragmentation, not necessarily or completely the periodic limb movements (Sergeeva et al., 2017 make comparable arguments related to motor skill consolidation discussed below) . Here the authors propose that since spindle power negatively predicted WASO, it is possible that falling below a healthy spindle power range induces cortical arousal in RLS.

Allen and colleagues (2015) further complete the emerging picture of sleep disruption in RLS by measuring glutamate levels, scanning the thalamus using MRI and monitoring sleep using PSG and found a relationship between increased glutamatergic activity (which in turn is regulated by dopamine) in the thalamus and arousal in RLS (the authors also cited Hornyak et al., 2007 and Montplaisir et al., 1997, who reported further evidence PLMs had little or no relationship with arousal in RLS).

In sum, sleep pathology in RLS is arguably best characterised by both partial sleep deprivation and sleep fragmentation, with disruption concentrating in the first half of the night with relatively little impact on the proportion of time spent in SWS or REM and with nighttime cortical arousal apparently driven by dopamine-related brain dysfunction as opposed to just periodic limb movements.

### *Is RLS associated with declarative memory deficits?*

Despite the well-established relationship between sleep and declarative memory consolidation (see Klinzing, Niethard, and Born, 2019, for recent review), surprisingly few studies have examined this association in the context of RLS. Since RLS patients often experience moderate-to-severe disruption to both sleep duration and continuity, and in a predictable circadian-driven pattern which pools in the first half of the sleep cycle, they are a highly suitable group to run a natural experiment on to better understand systems consolidation. For example, because symptoms in RLS are circadian driven, it is possible that these symptoms will inflict chronic disruption to selective components of the sleep cycle over time whilst preserving others. Specifically, we seek to test whether continuity of sleep predicts declarative memory deficits independently of sleep duration. At the same time, we also seek to verify some basic open questions in the RLS literature, namely whether belonging to the group is associated with encoding and overnight forgetting deficits as would be predicted by systems consolidation, discussed below.

**Models of sleep and memory and their key predictions.** Declarative memories pertain to facts and events which can be explicitly recalled (e.g., Camina & Guell, 2017). For a representation to be stored and available for retrieval in the long-term it needs to first be encoded and held in temporary storage during wake before being consolidated during sleep. This consolidation process involves the solidifying of an otherwise labile memory trace, and this is very broadly achieved by the integration of the knowledge temporarily held in the hippocampus to existing knowledge stored in neocortical networks (see Rasch & Born, 2013 for review). The most widely held theory of encoding is currently the Synaptic Homeostasis Hypothesis (SHY, Tononi & Cirelli, 2014), for which wakeful learning (encoding) comes at the cost of increased synaptic strength which approaches saturation after extended periods without sleep. For Tononi and Cirelli, this balance is restored in SWS through a process of synaptic down-selection (see Tononi & Cirelli, 2020, for a review of the evidence for SHY). A key prediction of SHY is therefore that disrupted sleep (either in length or in continuity) will impair next-day encoding. Several studies have found this (e.g., see Newbury et al., 2021 for a meta-analysis supporting this claim).

The dominant explanation of declarative memory consolidation is the Active Systems model (Marshall & Born, 2007; Diekelmann & Born, 2010, although see Cordi and Rasch, 2021 for a cautionary note). Within this account, sleep spindles (~11 to 16 Hz oscillations typically lasting between 0.5-1.5 s), slow oscillations (travelling bursts of < 1 Hz oscillations), and sharp-wave ripples (very fast and frequent bursts of 150-200 Hz oscillations) work in synchrony during NREM sleep (Stage 2 & SWS) to produce repeated reactivations of the to-be-stored memories thus facilitating their acquisition and integration into long-term storage networks in the neocortex (See Rasch and Born, 2015; Staresina et al., 2015 for detailed discussion).

A large body of evidence supports the proposition that sufficient duration and continuity are vital to both the encoding and consolidation of declarative memories (see Newbury et al., 2021 for a recent meta-analysis on the effect of sleep deprivation on both encoding and consolidation, see Omlin et al., 2019, and Tartar et al., 2006 for examples of studies showing declarative memory deficits following sleep fragmentation).

Alternative accounts also exist which challenge systems consolidation. The Contextual Binding theory (Yonelidas et al., 2019), proposes that the role is the hippocampus is not just a temporary one. Instead, the hippocampus continually updates and reshapes a memory representation in accordance with new contextual information. Key claims of this theory are that forgetting is contextual interference, which can occur before or after learning, and sleep, and in particular, SWS is especially protective against contextual interference and occurs at a place in the sleep phase where memory representations are most vulnerable to contextual interference.

**Declarative memory deficits in RLS.** It seems clear based on both SHY and Active Systems accounts that having RLS ought to predict declarative memory deficits. Indeed, Cha and colleagues' (2017) polysomnography study found evidence of low spindle density and poor coordination between spindles and slow oscillations among RLS participants. Given that these sleep parameters are essential to systems consolidation it is reasonable to hypothesise that their dysfunction will result in some degree of overnight forgetting. No study has directly investigated this question in RLS. However, Sargeeva and colleagues (2017) explored this issue in individuals with Periodic Limb movement disorder and found intact declarative memory consolidation despite disrupted sleep

(lower sleep duration and greater number of nocturnal awakenings than control). In their study, a PLM group, and a control group (N=14 PLM, N=15 Control, matched for age and sex) both took part in an evening-morning paired associate task in which they attempted to learn unrelated word pairs. Recall was cued by presenting participants with one word in the pair and training was capped at 60% recall. In another arm of the study, the authors did find impaired motor skill learning in PLM, but no difference in declarative memory between groups when comparing test scores in the evening with re-test scores the following morning.

There are several factors to consider when applying the Sargeeva et al study to the current investigation. The first is that despite the fact the PLMs are common in RLS, the available evidence discussed above suggests that PLMs do not drive increased arousal in RLS (Allen et al., 2015; Hornyak et al., 2007; and Montplaisir et al., 1997). As well as this, there is some meta-analytic evidence in other disorders of which sleep fragmentation is a hallmark symptom (Obstructive Sleep Apnoea Syndrome, Wallace and Bucks, 2012) that spatial memory tasks are more likely to detect declarative memory impairments than word pair tasks.

Finally, the findings of Experiment 3 suggest that the association between awakenings and consolidation is sensitive to encoding and that levels of encoding might need to be considered to observe a relationship between awakenings and consolidation. Experiment 3 also suggested that consolidation effects are very small and therefore a 60% encoding criterion may not be enough learning to create a testable amount of overnight forgetting. Study 3 took a multilevel modelling approach and had participants (a mixed sample primarily consisting of students and new parents, N= 519) complete a two-session object location task (Session 2 took part the calendar day following Session 1 at a time of the participants' convenience to allow for a delay period inclusive of a typical night of sleep) and subjectively estimate the number of awakenings and sleep duration for night of sleep in the delay period. We found a small positive relationship between one measure of awakenings (the number of awakenings lasting 5 minutes or longer for a typical night across the previous month) and overnight forgetting whilst controlling for sleep duration but only when level of encoding was factored into the model. Therefore, the data suggests that the association between awakenings and consolidation is sensitive to encoding. Encoding level was allowed to vary naturally in that study but the mean level

was still a very comparable 60%. It should also be noted that other reported uses of this object location task all found consolidation effects when encoding was set to a 100% criterion in lab-based experiments (Rudoy et al., 2009, Cairney et al., 2016, Lewis et al. ,2021). Taken together, the polysomnography data reported by Cha and colleagues (2017) strongly predicts declarative memory deficits among RLS, and it is possible that the null effect reported by Sargeeva and colleagues in PLM can be reconciled by task choice and the level of encoding. Specifically, a spatial memory task within which the level of encoding is statistically controlled may better allow for the detection of the effect if it exists.

***Does sleep fragmentation predict overnight forgetting independently of sleep deprivation in RLS?***

Relative to sleep deprivation, which has received a large amount of research attention, little is known about the extent to which sleep fragmentation is associated with declarative memory and affect. This is partially since it is very difficult to adequately control for the confounding effect of sleep duration in humans. The available animal literature supports the idea that sleep fragmentation independently predicts deficits in encoding and consolidation even when taking sleep duration into account (see Lui et al., 2019; Laharner et al., 2019). Only one study to our knowledge has investigated this in humans (namely Experiment 3 in Chapter 3). Whilst Experiment 3 complements the available animal literature and suggests that sleep fragmentation might uniquely predict overnight forgetting in humans whilst controlling for healthy sleep duration (M= 7.16 hours; SD = 1.51 hours), it is unknown whether this effect will survive when the sample is more sleep deprived. The literature is mixed with regards to whether or not sleep duration on its own is related to declarative memory consolidation. Many studies have shown group differences between sleep and declarative memory (e.g., Plihal and Born, 1997; Gais et al., 2008; see Rasch and Born, 2013, for review), yet these differences are arguably driven by SWS and not duration specifically. For example, Backhaus et al., (2008) found a significant correlation between SWS and word-pair retrieval but not between sleep duration and retrieval. Cousins et al., (2018) observed group differences between extended periods of restricted sleep and full sleep on measures on encoding and Van Dongen et al., (2003) showed similar effects of sleep restriction on another measure of declarative memory (Paced Auditory Serial Addition Task). It is perhaps

possible therefore that SWS drives the relationship between sleep and declarative memory measures, that sleep duration is related to SWS, and that sleep duration only relates to declarative memory in so much as SWS is restricted within a given sleep window. Applying this to the current study, it is clearly important to control for sleep duration given how unsettled the literature is on this topic. As well as this, it is possible that sleep duration was unexpectedly high for the sample opportunistically recruited in Experiment 3, that there was not a high degree of variability of SWS within this sample. Given that the RLS population is very rich in variability of sleep disruption, both for sleep duration and sleep fragmentation, they are an ideal group to test both whether or not sleep duration is associated with declarative memory but also whether or not sleep fragmentation is associated with measures of declarative memory in the context of this more highly variable sleep duration.

***Is RLS associated with the expression of positive and negative affect?***

There is a convincing evidence base to suggest that individuals with RLS are more likely to have or go on to develop mood disorders. For example, Becker and Sharon (2014) identified 32 epidemiological studies in a review that reported links between RLS and anxiety and depression. Panic disorder, Generalised Anxiety Disorder, and Major Depression are also reported as being strongly associated with RLS (Winkelmann et al., 2005). Despite this, very few studies have examined expressions of affect in RLS (note that affect refers broadly to an immediate expression of mood which is a longer-term description of the individual's emotional composition, Martin, 1990). This question is important both for understanding the daily lived experience of individuals with RLS and for tracking the evolution of their emotional state as the condition progresses. As well as this, daytime expressions of positive and negative affect are also a lens through which to better understand the contributions of distinct parts of the sleep cycle to daytime functioning.

As discussed above, RLS individuals suffer compromised sleep both in terms of duration and continuity. Tomasso et al., (2020) analysed effect sizes from 64 studies focussing on the relationship between sleep deprivation and affect and perhaps unsurprisingly found that sleep deprivation decreases the expression of positive affect and increases the expression of negative affect. Experiment 3 (Chapter 3) also measured

positive and negative affect using the Positive and Negative Affect Schedule (PANAS) and found that the number of nocturnal awakenings the night before measurement predicted both positive and negative affect whilst controlling for sleep duration and alertness (as awakenings go up so too does negative affect, as awakenings go up positive affect does down). Interestingly, a moderate association was found between alertness (measured using the Stanford Sleepiness Scale) and positive affect ($\beta$ = -0.38) leading to a much larger amount of explained variation in the positive affect model (R2 adjusted = 18.2%) relative to the negative affect model (R2 adjusted = 5.7%). Expressions of negative affect were only weakly related to alertness ($\beta$ = 0.13). Given that half of the sample were new parents who also experience rich variation in sleep fragmentation, this suggests that RLS individuals will express less positive affect and less negative affect than healthy controls.

Variations in the expression of positive and negative affect as discussed above are thought to be influenced by sleep. Specifically, next-day expression of positive affect is supposed to be driven by dopamine regulation in SWS (see Krause et al., 2017, for review), and next-day expression of negative affect is thought to be governed by REM sleep (see Walker and Goldstein, 2014, for review). Applying these models to RLS leads to some interesting implications. A key difference between the current study and Study 2 is that new parents experience sleep fragmentation as part of a random highly variable profile of disruption driven by the needs of the child. In RLS symptoms are driven by circadian factors and pool in the evening and first half of the night. This means that it is very plausible that the continuity of REM sleep in RLS individuals will be preserved whilst the continuity of SWS is more severely affected. This leads to the prediction that the proportion of explained variance in affect explained by sleep continuity in RLS will be higher for positive affect than for negative affect since sleep continuity in RLS is likely much more disrupted for SWS relative to REM sleep. Despite being very intriguing, that is a question better answered using polysomnography. For current purposes, it is likely that chronic pain and discomfort driven by RLS symptoms will lead to increased levels of negative affect in this group despite having preserved REM sleep. For example, Galloway and Espie (2008) found increased expression of negative affect measured using PANAS immediately following a Suggested Immobilisation Test which involves keeping the legs still for an hour. Overall, given the combination of sleep disruption and daytime discomfort

characteristic of RLS, it is very likely that expressions of positive affect will be lower, and expressions of negative affect will be higher relative to healthy controls.

***Does RLS medication use alter these associations?***

A small but considerable proportion of individuals who experience symptoms of RLS require treatment (Allen et al., 2011 estimated rates at around 1.7-2.5%). Treatment options vary with the main options including a combination of iron supplementation, screening and cessation of agitators, and dopamine agonists (O'Regan and Anderson, 2020). The evidence is mixed with regards to the efficacy of these approaches with regards to sleep, memory, and affect. Firstly, dopamine agonists, the most common and widely used therapy, has been shown on several occasions to improve sleep quality. For example, dopamine agonists have been found to increase sleep duration (Trenkwalker et al., 1995), decrease the number of awakenings, daytime somnolence, and PLMs (Collado-Seidel et al., 1999; Saletu et al., 2003, in Bogan, 2006). Two studies identified and reviewed by Jung in 2015 (Kim et. al., 2014; and Lee et. al., 2014) also reported improved cognition (e.g., word frequency, digit symbol coding, and verbal memory measured using the Korean-California Verbal Learning Test) after three months of dopamine agonist use in RLS participants.

However, despite some positive indications as to the efficacy of dopamine agonists in regulating sleep and memory in RLS, patients perpetually risk experiencing augmentation effects (i.e., worsening RLS symptoms that exceed what would typically be expected because of overuse of dopamine agonists). Augmentation causes patients to feel symptoms in more parts of their body and for a greater part of the day. Rates of augmentation are high, with around 50-60% of RLS patients requiring greater doses as their RLS progresses and about 7% per year experiencing augmentation effects (O'Regan and Anderson, 2020).

As described above, the picture of medication use in RLS is complicated and evolving. This study is the first to our knowledge to take a naturalistic snapshot of medication use in RLS without focussing on a particular type of medication and investigating how an unspecified and random sample of medication use in RLS is associated with sleep, declarative memory, and affect. Whilst it seems possible that

medication use in general in RLS may have some cognitive and affective benefit, it is unknown whether or not this will be comparable to the experience of healthy controls.

**Sleep Measures in RLS**

As described in Chapter 1, the type and combination of sleep measures used and relative to which population can have an impact when interpreting results in sleep studies. Literature dedicated towards evaluating and comparing the effectiveness of different sleep measures suggest that: (1) a combined approach is optimal where possible given that each of polysomnography, actigraphy, and self-report have their inherent strengths and weaknesses (Buysse et al., 2007); (2) polysomnography is the gold-standard for accuracy but is limited in that it is less likely to record typical sleep (Rosenberg and Van Hout, 2013); (3) actigraphy is reasonably concordant with polysomnography but underestimates awakenings (Kushida et al., 2001) and overestimates sleep duration (Conley et al., 2019); (4) self-report is insensitive to awakenings lasting less than around 5 minutes (Winser et al., 2013) but unique in that it captures subjective experience (Harvey and Tang, 2012).

Concerning Restless Legs Syndrome specifically, Hening et al., (2004) firstly establishes that it is common, especially in the lower end of the severity spectrum, to not experience all of core symptoms (about 80% experience at least one sleep symptom at least twice per week). This presents the likelihood of increased variation when using snapshot measures such as polysomnography. Saletu et al., (2007) reported that a sample of individuals with Restless Legs Syndrome had a sleep duration of 326 minutes (approximately one hour less than healthy controls) and experienced 12.2 awakenings (approximately 5 greater than healthy controls). Lee (2009) measured sleep patterns among 26 RLS patients with actigraphy and reported a mean sleep duration of 366 mins and 13.24 awakenings per night. Considerably less data is widely available on the number of self-reported number of nocturnal awakenings in RLS. Bogan (2006) reported that 60% of individuals with RLS report waking up at least once per night and about 20% report waking up 3 times or more each night and 9.7% reported sleeping less than 6 horus per night. The available data suggests that individuals with RLS self-report less nocturnal awakenings than more objective measures. However, it is unclear whether or not the subjective experience inherent to these self-reports also captures the essence of

the potential relationship between awakenings and declarative memory. In other words, it is not inherently the case that reporting less awakenings is the same as being less accurate. It is possible that self-reports capture the essence of the relationship. Finally, if the differences between the self-reported number of awakenings and objectively measured number of awakenings are large, this could have implications for statistical power when investigating the relationship between awakenings and declarative memory and affect.

### *The Current Investigation*

The current investigation has four principal aims. The first is to test whether spatial declarative memory deficits (namely related to encoding and overnight forgetting) are associated with RLS.This study is not intended to directly inform the dominant theories of sleep and declarative memory consolidation described above. Investigations capturing oscillatory patterns and objective measurements of awakenings are better suited to this. Instead, the current investigation aims to show that sleep fragmentation is an important and overlooked influencer of the processes underlying these models and as such should be explored more often and more routinely framed in the central of the debate.. Secondly, we aim to extend the RLS literature on mood and assess expressions of affect as part of the daily lived experience of those who experience symptoms. Isolating expressions of positive and negative affect separately may also help to improve understanding of the contributions of distinct parts of the sleep cycle to daytime functioning in RLS. Thirdly, we aim to build on Study 3 aimed at isolating the unique contributions of sleep fragmentation and sleep duration to the variance in declarative memory and affect, this time exploring whether the number of awakenings predict these outcome variables in the presence of the greater levels of sleep deprivation seen in RLS. Finally, we aim to take an exploratory snapshot of how an opportunistic sample of individuals experiencing RLS symptoms and taking unspecified medication for it compare with healthy controls on measures of declarative memory and affect.

To test these aims, the same pre-sleep/post-sleep memory game paradigm as in Study 3 will be used among a sample of unmedicated RLS, medicated RLS, and healthy controls. Importantly, instead of inserting a direct (subjective) measurement of awakenings into the models, it will be inferred from RLS group membership given that it

is one of the core symptoms of RLS. Direct (subjective) measurements of the number of awakenings will also be asked, however, and evaluated as part of a separate analysis as outlined below.

## Methods

### *Hypotheses*

The following hypotheses can be derived from the aims described above:

**H1**. The following set of predictions are based on a set of linear models combining data from individuals complaining of at least a moderate degree of RLS symptoms (i.e., having a score or 10 or greater on the International Restless Legs Scale; IRLS) with healthy controls from Study 2 in Chapter 3. It is predicted that complaining of RLS symptoms will be associated with increased encoding error, decreased expression of positive affect, and increased expression of negative affect. RLS will also be associated with increased nocturnal awakenings. Belonging to this cohort will also be associated with increased overnight forgetting, but only when the level of encoding error is controlled for in the model. This set of analyses is like those discussed below in H2 but allows us to consider the influence of encoding error in the association between awakenings and overnight forgetting, and to explore whether medication use offers any apparent protection against memory and affective deficits in RLS.

**H2**. The following set of predictions are based on a set of matched t-tests focused on comparing unmedicated RLS participants with healthy controls from Study 2. Unlike in H1, participants were matched on age and level of education. It is predicted that the RLS group (IRLS >= 10) will be significantly higher in encoding error, express significantly less positive effect and more negative affect, and wake up more frequently during the night than healthy controls. It is also predicted that there will be no significant differences in overnight forgetting between the groups since the level of encoding error is not controlled for. Finally, when the analyses are ran again controlling for sleep duration (matched ANCOVAs), the same pattern of results will be observed.

**H3**. In Study 2, an increased number of subjectively estimated awakenings lasting 5 minutes or longer predicted increased overnight forgetting whilst controlling for sleep duration and level of encoding error. The above analysis in H1 and H2 revisit this question but are limited in that they rely on the same healthy participants that were used in Study 2. Therefore, we sought to carry out a similar mixed model among a completely independent sample, namely our RLS participants. It is predicted that increased subjectively estimated awakenings the night preceding retrieval will be associated with increased overnight forgetting whilst controlling for sleep duration and encoding error.

***Participants***

Participants were adults aged 18 or older who scored 10 or greater on the International Restless Legs Scale (See Questionnaires section below, Walters et al., 2003) as well as healthy controls. Participants were also excluded if they disclosed symptoms of common 'mimic' disorders which are known to be similar in presentation to RLS (e.g., chronic kidney disease; Hening et al., 2009). To be included in the study, participants had to complete both sessions and complete Session 2 by the end of the calendar day following completion of Session 1. A total of 624 individuals were screened, of which 295 completed Session 1, of which 215 completed both sessions. Of these 215 individuals, 179 were eligible for inclusion in the study. Reasons for exclusion included not completing the second session in the appropriate time window (the end of the calendar day following the completion of the first session), completing the first session more than once, or declaring having a relevant comorbidity. Of these 179 individuals who fully met the inclusion criteria, 85 met the criteria for RLS and were medicated, 64 met the criteria for RLS and were unmedicated, and 30 reported no RLS symptoms and as such were added to the pool of healthy controls.

For H1 the aim was to better understand the influence of RLS on declarative memory and affect and to explore any evidence of a treatment benefit among those taking medication. A multiple regression approach was used combining all these healthy controls with the 179 individuals recruited for this study. 100 power simulations were carried out using PowerSim in R (Kumle, Vo & Draschkow, 2021). To carry out these simulations, we specified one of the models we were most interested in (namely the one with overnight forgetting as the dependent variable) and used the dataset from Study 2

to allow the package to estimate power. The simulations suggested that power with α=0.05 was around 80%. Table 42 below shows sample characteristics for H1:

*Table 43. Sample characteristics H1 (n = 452).*

| Characteristic | Healthy | RLS unmedicated | RLS medicated | Total |
|---|---|---|---|---|
| Age | M = 34.5; SD = 9.1 | M = 37.1; SD = 14.0 | M = 51.1; SD = 15.9 | M = 31.4; SD = 15.5 |
| **Sex assigned at birth** | | | | |
| Male | N = 67 (14.89%) | N = 13 (2.89%) | N = 18 (4.00%) | N = 98 (21.68%) |
| Female | N = 234 (52.00%) | N = 51 (11.33%) | N = 67 (14.89%) | N = 352 (77.88%) |
| **Education** | | | | |
| High School | N = 184 (40.89%) | N = 19 (4.22%) | N = 27 (6.00%) | N = 230 (50.88%) |
| University/college | N = 118 (26.22%) | N = 44 (9.78%) | N = 58 (12.89%) | N = 220 (48.67%) |

*Note*. Sample characteristics for multilevel models with N=452 participants. RLS = Restless Legs Syndrome. Age is in years. Percentages are for whole sample. Healthy (N) = 303 (67.03%); RLS unmedicated (N) = 64 (14.16%); RLS medicated (N) = 85 (18.81%).

For H2 we sought to investigate the influence of RLS on declarative memory and affect free from the confounding influence of treatment. This was achieved through 1:1 nearest neighbour propensity score matching using the 'Matchit' package in R (Ho, Lmai, King, and Stuart, 2011). Each of our eligible treatment naive RLS participants were matched on age and education. Our sample size is that which is required for 80% power for detection of a medium effect with α=0.05 (G*Power; Faul, Erdfelder, Lang, and Buchner, 2007). However, the analyses conducted so far in this thesis have suggested that the effect sizes we are interested in are much smaller than had first been anticipated. A post-hoc sensitivity analysis using GPower 3.1 suggested that the amount of participants included in H2 would have approximately a 30% chance of detecting effect sizes similar to that observed in Experiment 3.Details of the matching procedure are contained in the results section below. See Table 43 for sample characteristics for testing H2:

*Table 44. Sample characteristics H2 (n = 118).*

| Characteristic | Healthy | RLS Unmedicated | Total |
|---|---|---|---|
| Age | M = 33.78<br>SD = 12.78 | M = 37.03<br>SD = 13.74 | M = 35.41<br>SD = 13.31 |
| **Sex assigned at birth** | | | |
| Male | 13 (22.03%) | 13 (22.03%) | N = 26 (22.03%) |
| Female | 46 (77.97%) | 46 (77.97%) | N = 92 (77.97%) |
| **Education** | | | |
| High School | 11 (18.64%) | 17 (28.81%) | N = 28 (23.73%) |
| University/college | 48 (81.36%) | 42 (71.19%) | N = 90 (76.27%) |

*Note.* Healthy (N) = 59; RLS unmedicated (N) = 59. Percentages for total sample.

For H3 the goal was to independently test a potentially interesting exploratory finding from Study 2. Therefore, only the 149 individuals recruited specifically for inclusion in the current study were used to test H3, described in Table 44 below. 100 power simulations were carried out using Powersim in R (Kumle, Vo & Draschkow, 2021) on the data to determine the likelihood of detecting an effect of equal size to that found in the previous study. With an α=0.05, these simulations suggested that with the participants recruited, the power to detect an effect at least as large as that seen in Experiment 3 was approximately 69%.

*Table 45. Sample characteristics H3 (n = 149).*

| Characteristic | RLS Unmedicated | RLS Medicated | Total |
|---|---|---|---|
| Age | M = 37.1 | M = 51.1 | M = 44.1 |
| | SD = 14.0 | SD = 15.9 | SD = 15.0 |
| **Sex assigned at birth** | | | |
| Male | N = 13 (8.72%) | N = 18 (12.08%) | N = 38 (20.8%) |
| Female | N = 51 (34.23%) | N = 67 (44.97%) | N = 140 (79.2%) |
| **Education** | | | |
| High School | N = 19 (12.75%) | N = 28 (18.79%) | N = 52 (31.54%) |
| University/college | N = 45 (30.20%) | N = 57 (38.26%) | N = 127 (68.46%) |

*Note.* Percentages are for whole sample. RLS unmedicated (N) = 64 (42.9%); RLS medicated (N) = 85 (57.1%).

Data was collected online from August 2020 until March 2021 and participants were recruited on social media. To safeguard against attrition, a prize fund was created where participants who fully completed the study were entered into 3 draws with a chance to win a £50 Amazon voucher. Figure 16 below shows distributions of symptom severity for both RLS groups:

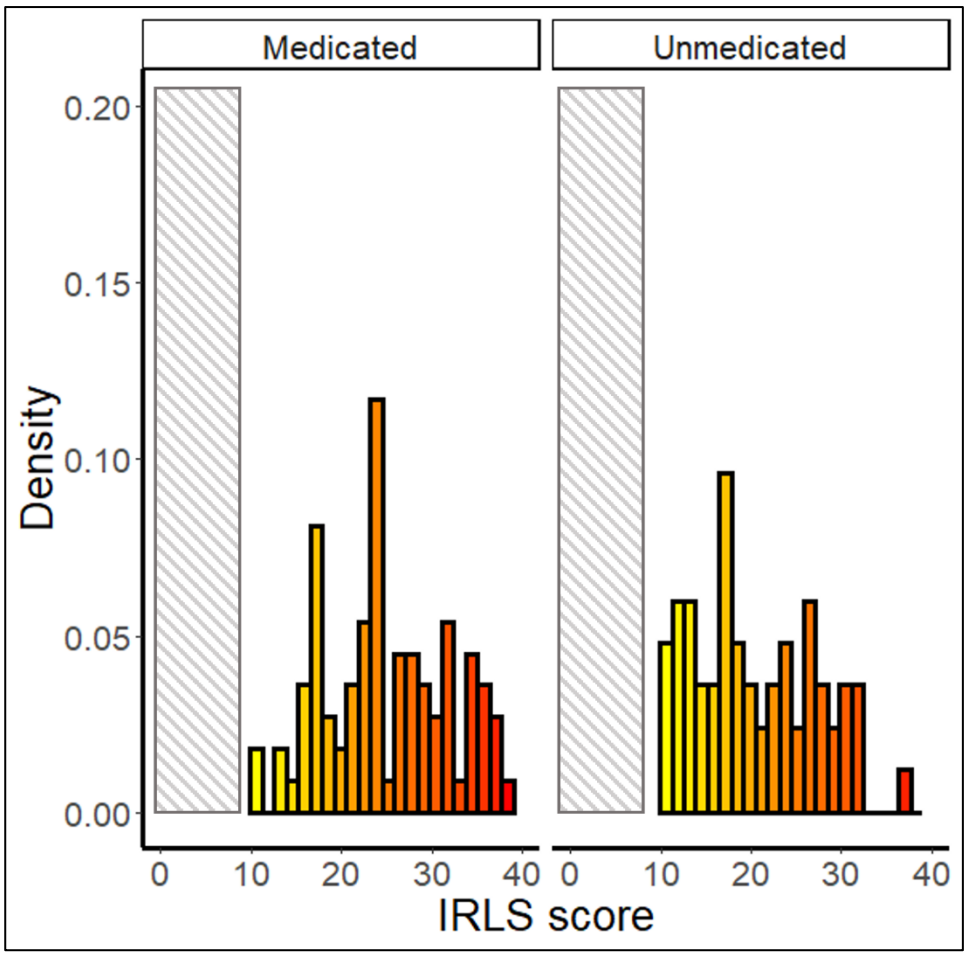



*Figure 16. Density plot of RLS severity as measured by IRLS. Density plots for International Restless Legs Scale scores for the medicated group (N = 85 M = 24.9, SD = 7.0) and the unmedicated group (N = 64; M = 20.3, SD = 6.7). 0-10 (Grey) = No -to- mild RLS; 11-20 = Moderate RLS; 21-30 = Severe RLS; 31-40 = Very severe RLS.*

### *Stimulus Presentation*

Stimulus presentation was identical to that of the study reported in Experiment 3. Participants were tasked with learning the screen locations of 20 objects (e.g., *button, key)*. Participants again completed the tasks on a device of their choice and 46.70% of participants completed the study on a smartphone, 39.09% on a laptop, and 14.21% on a tablet.

### *Procedure*

The procedure for this study is like Experiment 3 and is visualised in Figure 17 below:

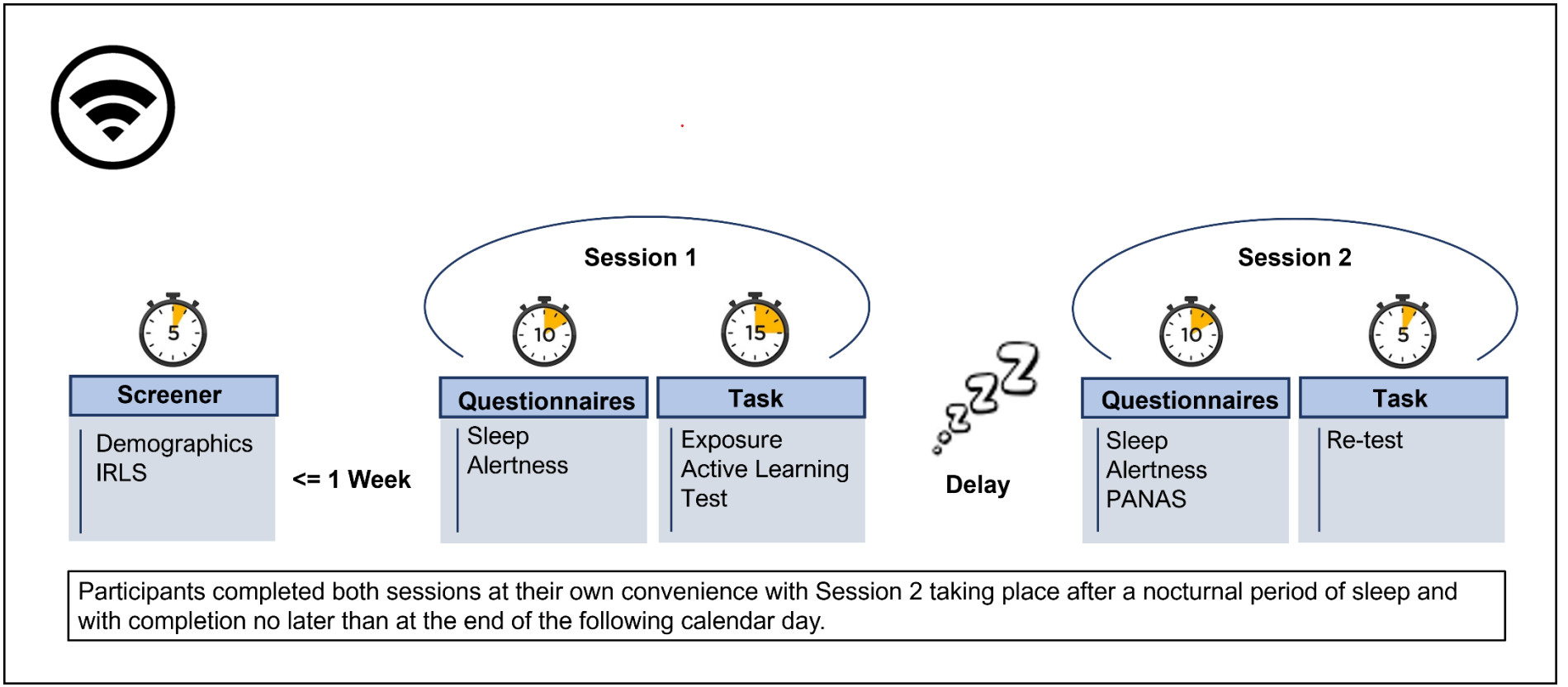



*Figure 17. Timeline, tasks, and time needed to complete each session.*

The only difference in this study was that participants completed a screener form within a week of completing Session 1 and were sent a link via email with a description of the study and links to begin it. This email address would also serve as the unique ID that allowed us to combine Google Form and Pavlovia responses across sessions. Figure 18 below demonstrated Session start times, which were at the participant's convenience. Participants were encouraged to start Session 1 in the evening and to complete Session 2 the following morning but were ultimately included if they completed both sessions on consecutive calendar days with a period of nocturnal sleep between them.

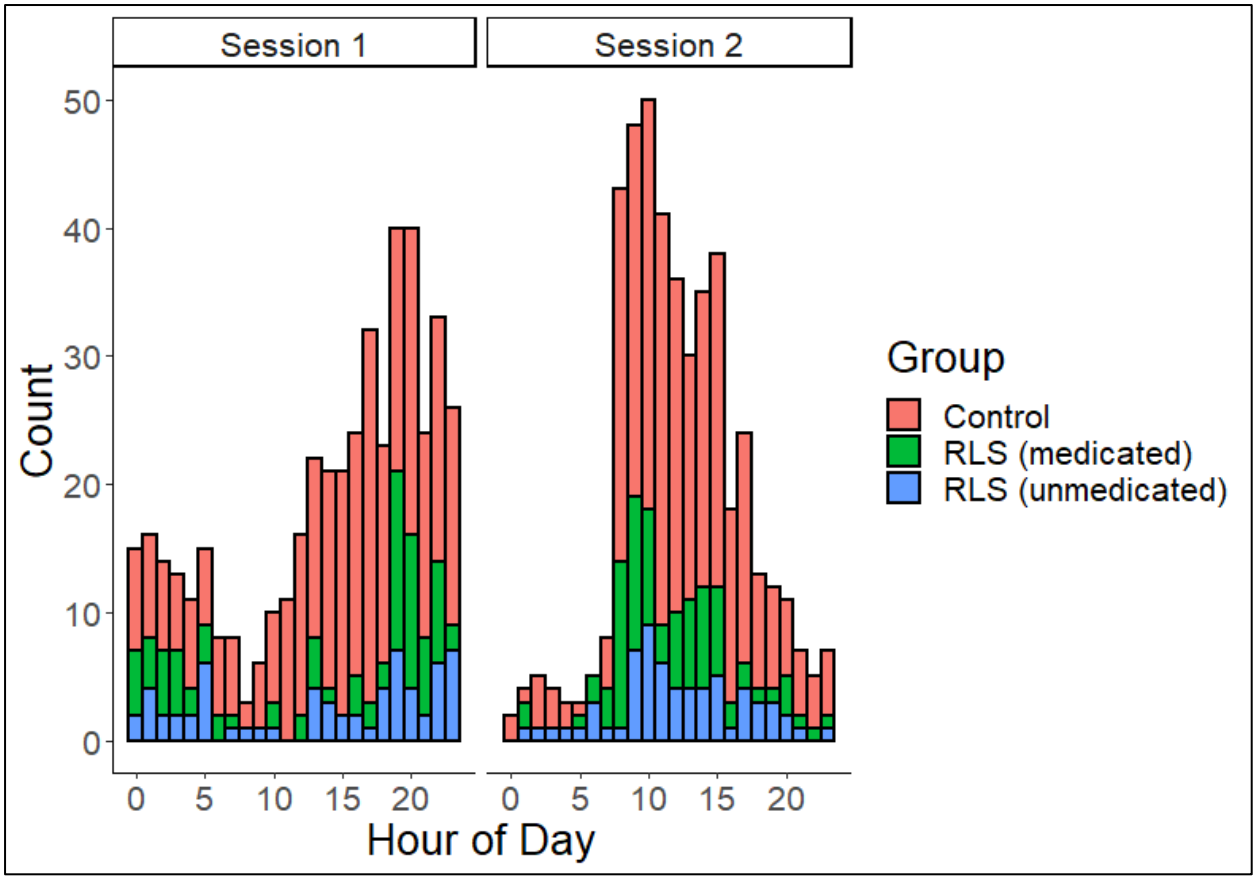



*Figure 18. Stacked histograms for participant start times for each session. X axis = Hour of day (0-23), Y axis = Count. N=452.*

### *Questionnaires*

**Screener**. The screener form was a short questionnaire completed using Google Forms in which potential participants gave some basic demographic information, completed the International Restless Legs Scale and declared whether or not they had been diagnosed with any disorders that would exclude them from participating in the study.

***Demographics***. Age (years), sex assigned at birth and years of education.

***International Restless Legs Scale.*** The International Restless Legs Scale (IRLS, Walters et al., 2003) subjectively captures the severity and frequency of symptoms of Restless Legs Syndrome across the previous week. Ten questions make up the IRLS, with answers being on a scale of 0 (No RLS) to 4 (Very severe RLS). With a total score of 40, 0-10 indicates no or mild RLS, 11-20 indicates moderate RLS, 21-30 indicates severe RLS, and 31-40 indicates very severe RLS. The IRLS score has acceptable construct validity, internal consistency reliability (α=0.81), and concurrent validity (r=-0.68) with the Restless Legs Syndrome Quality of Life questionnaire (RLSQol; Abetz et al., 2006). We also asked whether participants were taking or had recently taken medication for RLS. Unfortunately, the first question on this scale (Overall, how would you rate the RLS discomfort in your legs or arms?) was unknowingly absent from the questionnaire throughout data collection. To create a score which matched the guidelines discussed above, the items on the scale (10) were divided by the number answered (9) and multiplied by the total score of the answered items. All other questions, including those relating specifically to sleep and mood disruption remained.

***Comorbidities.*** For our RLS cohort, participants were asked to disclose whether they had a diagnosed sleep disorder other than RLS and if they were pregnant, had chronic kidney disease, were diabetic, had deep vein thrombosis, had an open wound in the lower half their body, or had gotten a fracture anywhere in the body within the past 3 months (see Hening et al., 2009, for discussion). Having another sleep disorder would confound the results and having any of the other comorbidities can mimic RLS in subjective measures. Declaring any of these resulted in exclusion from the study. Healthy controls were asked if they had a diagnosed sleep disorder and were similarly excluded

if they declared so (see Study 2 for a more detailed discussion of inclusion criteria for that sample).

**Session 1 Questionnaire. *Sleep.*** All participants subjectively estimated their sleep duration in hours and minutes (to the nearest 15 minutes) for the night prior to completion of Session 1. They also estimated the total number of times they woke up for this night, how many of these awakenings lasted 5 minutes or longer, how many times in total they woke up during the night for a typical night across the last month, and how many of these awakenings lasted 5 minutes or longer.

***Alertness.*** Alertness was measured using the Stanford Sleepiness Scale (Hoddes, Dement, and Zarcone, 1972), and as such participants described how they felt currently (1=*Feeling active, vital, alert, or wide awake; 2=Functioning at high levels, but not at peak; able to concentrate; 3=Awake, but relaxed, responsive but not fully alert; 4=Somewhat foggy, let down; 5=Foggy, losing interest in remaining awake; 6=Sleepy, woozy, fighting sleep, prefer to lie down; 7=No longer fighting sleep, sleep onset soon, having dream-like thoughts)*.

**Session 2 Questionnaire. *Sleep.*** *All p*articipants disclosed the same information with regards to the evening prior to the completion of Session 2.

***Alertness.*** Stanford Sleepiness Scale, as above.

***Affect****.* Affect was measured subjectively using the Positive and Negative Affect Schedule Short Form (PANAS-SF; Watson, Clark, and Tellegen, 1988). This measure splits mood into positive and negative affect and asks participants for a list of 20 items to indicate the extent to which they had felt that way over the last week (*Interest; Distressed; Excited; Upset; Strong; Guilty; Scared; Hostile; Enthusiastic; Proud; Irritable; Alert; Ashamed; Inspired; Nervous; Determined; Attentive; Jittery; Active; Afraid).* The value for Cronbach's alpha was reported as α=0.86-0.90 for the positive affect scale and α=0.84-0.87 for the negative affect scale by Watson and Tellegen in 1988.

### *Object location task*

The object location task used was identical to that used in Experiment 3. This task was adapted from one used by Rudoy, Voss, Westerberg, and Paller (2009). In Session 1, which lasted on average 15 minutes, participants were first exposed to 20 items one at a time which each had a unique screen location (Figures 2 and 3 above). They were then tasked with learning these object locations in an active learning phase and their memory for these screen locations was tested. In Session 2, which lasted around 5 minutes, they were tested in the exact same fashion as in Session 1.

## Statistical Analysis Plan

Broadly speaking, H1 and H3 take a hierarchical multiple regression approach and H2 involves the use of matched t-tests and ANCOVAs. For H1 and H3, if the outcome variable was a measure of memory, this allowed us to control for the random effects of participant and trial (object).

### *Dependent variables*

**Encoding error.** A mathematical index of distance was used to measure retention of object locations after the first test in session 1. This method has been used in several previous studies (e.g., Rudoy et al., 2009; Cairney, Lindsay, Sobczak, Paller, and Gaskell, 2016). Firstly, the absolute value of the subtraction of the test x-coordinate from the correct x-coordinate is calculated for each object (*S1(x)Distance = ABS(S1X-CorrectX*)) in normative screen units. The same is then carried out for the y-coordinates (*S1(y)Distance = ABS(S1Y-CorrectY*)) for each object before calculating the square root of (*S1XDistance*S1XDistance)* + (*S1YDistance*S1YDistance).* The resulting value (*S1Distance*) is one index of encoding and is in normative screen units, with higher values representing a lower amount of success in the participant's attempt to remember the object location of all the objects. Indices of encoding error in our study ranged from 0.002 (almost no error) to 2.06 (almost no accurately encoded information).

**Overnight forgetting.** The same index as for encoding error was carried out for Session 2 test scores and an index of overnight forgetting was calculated which is

*S2Distance - S1Distance.* A positive index of overnight forgetting is indicative of memory loss and a negative index of overnight forgetting is indicative of memory improvement (consolidation). Indices of overnight forgetting in our study ranged from -1.90 (substantial overnight improvement) to 1.82 (substantial overnight forgetting).

**Positive affect.** A positive affect score was calculated by adding the scores on items 1,3,5,9,10,12,14,16,17, and 19 from the PANAS-SF. Scores can range from 10-50 and a higher score corresponds to a higher degree of expressed positive affect.

**Negative affect.** A negative affect score was calculated by adding the scores on items 2,4,6,7,8,11,13,15,18, and 20 from the PANAS-SF. Scores can range from 10-50 and a higher score corresponds to a higher degree of expressed negative affect.

***Control variables***

**Age.** Age was measured in years.

**Sex assigned at birth.** Participants selected from: *Female; Male; Don't know;* and *prefer not to say*. Dummy coding was used with "Female" as the intercept.

**Level of Education.** A binary variable was created for level of education with the following levels: High School; and; Further Education. Dummy coding was used with "High School" as the intercept.

**Alertness.** Alertness was measured using the Stanford Sleepiness Scale, an ordinal variable ranging from *1 - Feeling active, vital, alert, or wide awake* to *7 - No longer fighting sleep, sleep onset soon; having dream-like thoughts*. Since RLS individuals are also likely to be experiencing deficits in attention it is important to control for this potentially confounding influence in the models.

**Random effects.** Participant number and trial (object) were random effects.

### *Independent variables*

**Sleep duration**. Sleep duration is the subjectively estimated total nocturnal sleep time for the night preceding each session. If the outcome variable is encoding error, then sleep duration refers to the evening preceding session 1. Otherwise, it refers to the evening preceding Session 2.

**Awakenings.** For H3 , Awakenings is the subjectively estimated total number of awakenings for the night preceding Session 2.

**Group.** For H1, Group was a factor with 3 levels: *Control; RLS_unmedicated; RLS_medicated*. This variable was dummy coded with “Control” as the intercept. For H2, Group was a factor with 2 levels: *Control; RLS_unmedicated*. Each group was matched for Age and Education using nearest neighbour propensity score matching.

## Results

### *H1 - Hierarchical Multiple regressions (N=452):*

**Is group related to declarative memory and expressions of affect?** This section takes a hierarchical multiple regression approach to test associations RLS, declarative memory, and affect by combining all participants recruited for this study (N=179; RLS (unmedicated) = 64; RLS (medicated) = 85; Healthy controls = 30) with all the healthy controls from Chapter 3 (N=273). Table 45 below shows descriptive statistics for sleep, memory, alertness, and affect by group:

*Table 46. Descriptive statistics for H1.*

| Variable | Healthy | RLS Unmedicated | RLS Medicated | Total |
|---|---|---|---|---|
| Encoding error | 0.27 (0.30) | 0.41 (0.34) | 0.35 (0.34) | 0.31 (0.31) |
| Overnight Retention | 0.10 (0.12) | 0.08 (0.13) | 0.09 (0.13) | 0.10 (0.45) |
| Positive Affect | 28.4 (7.67) | 25.4 (8.88) | 23.8 (8.98) | 27.08 (8.30) |
| Negative Affect | 19.8 (7.09) | 21.1 (8.40) | 20.2 (7.01) | 20.03 (7.25) |
| TST(S1) | 459.55 (91.97) | 377.86 (148.55) | 342.22 (110.52) | 425.34 (117.68) |
| TST(S2) | 464.85 (88.48) | 417.14 (108.56) | 375.36 (123.69) | 440.77 (105.09) |
| Awakenings(S1) | 1.22 (1.60) | 2.73 (2.03) | 4.29 (3.98) | 2.02 (2.58) |
| Awakenings(S2) | 1.31 (1.60) | 2.48 (2.06) | 4.07 (3.34) | 1.99 (2.35) |
| SSS1 | 3.09 (1.24) | 3.52 (1.46) | 3.65 (1.38) | 3.27 (1.31) |
| SSS2 | 2.77 (1.18) | 3.29 (1.41) | 3.56 (1.50) | 2.99 (1.31) |

*Note.* Means and Standard deviations objective variables across models. Encoding error = index of retention based on S1 test scores; Overnight Retention = Index of retention based on S2-S1 scores; TST(S1) = Subjectively estimated sleep duration for night previous to completion of Session 1 in minutes; TST(S2) = Subjectively estimated typical sleep duration for night previous to completion of Session 2 in minutes; Awakenings(S1) = subjectively estimated number of nocturnal awakenings for night previous to S1; Awakenings(S2) = subjectively estimated number of nocturnal awakenings for night previous to S2; SSS1 = Stanford Sleepiness Scale scores for Session 1; SSS2 = Stanford Sleepiness Scale scores for Session 2.

Figure 19 shows Pearson's correlations for objective measures:

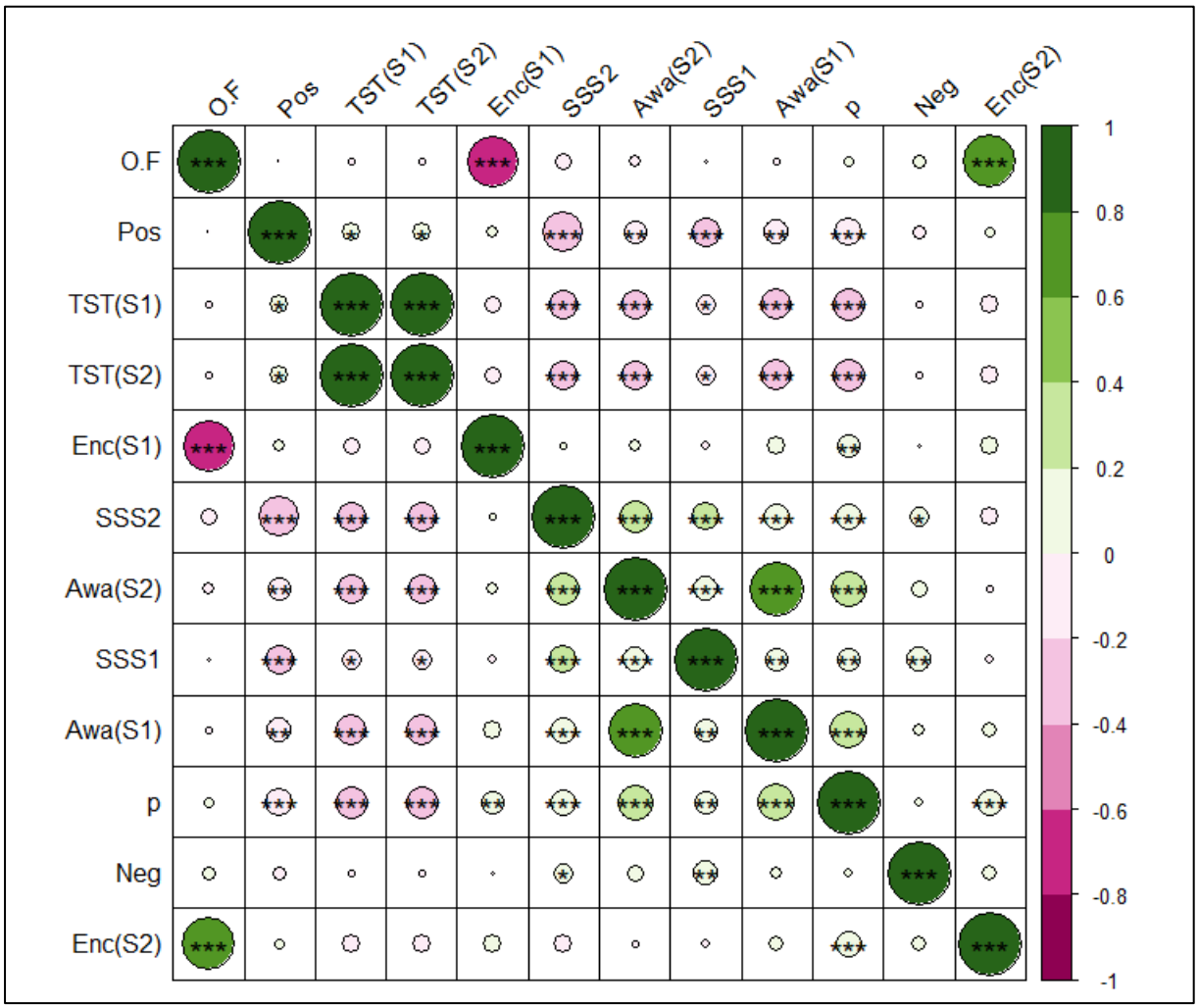
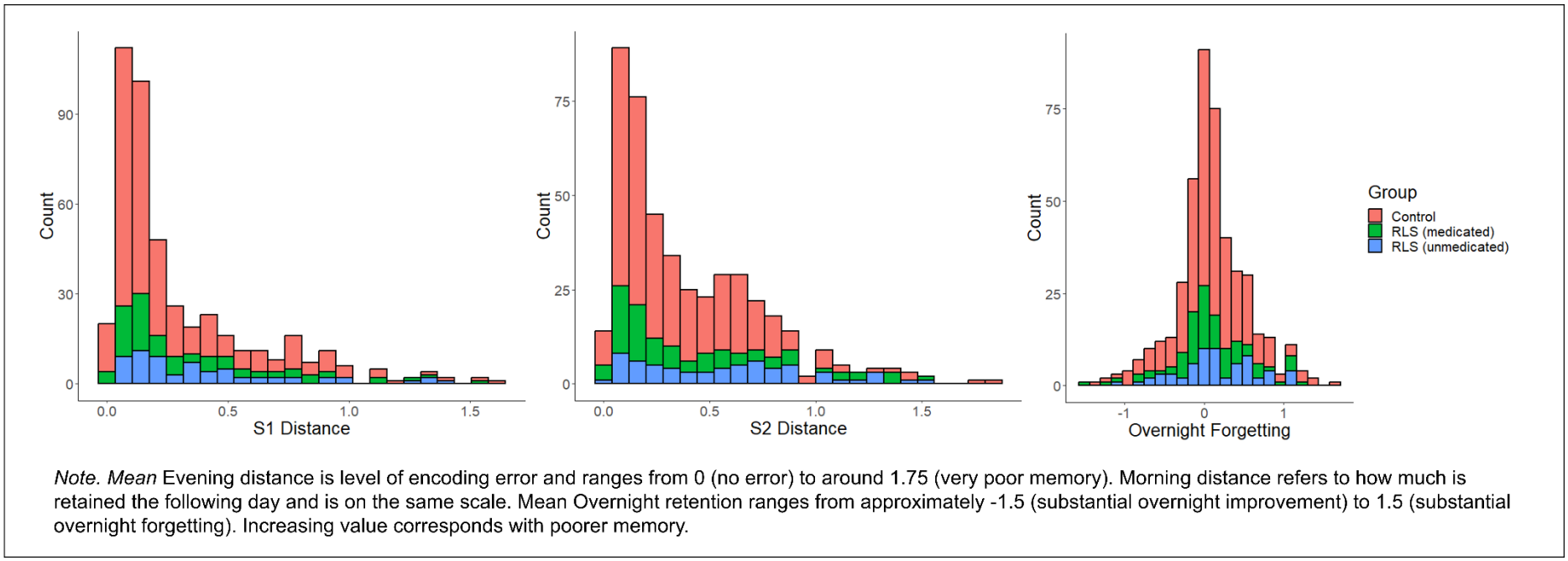



*Figure 19. Correlation matrix for objective measures for H1.*

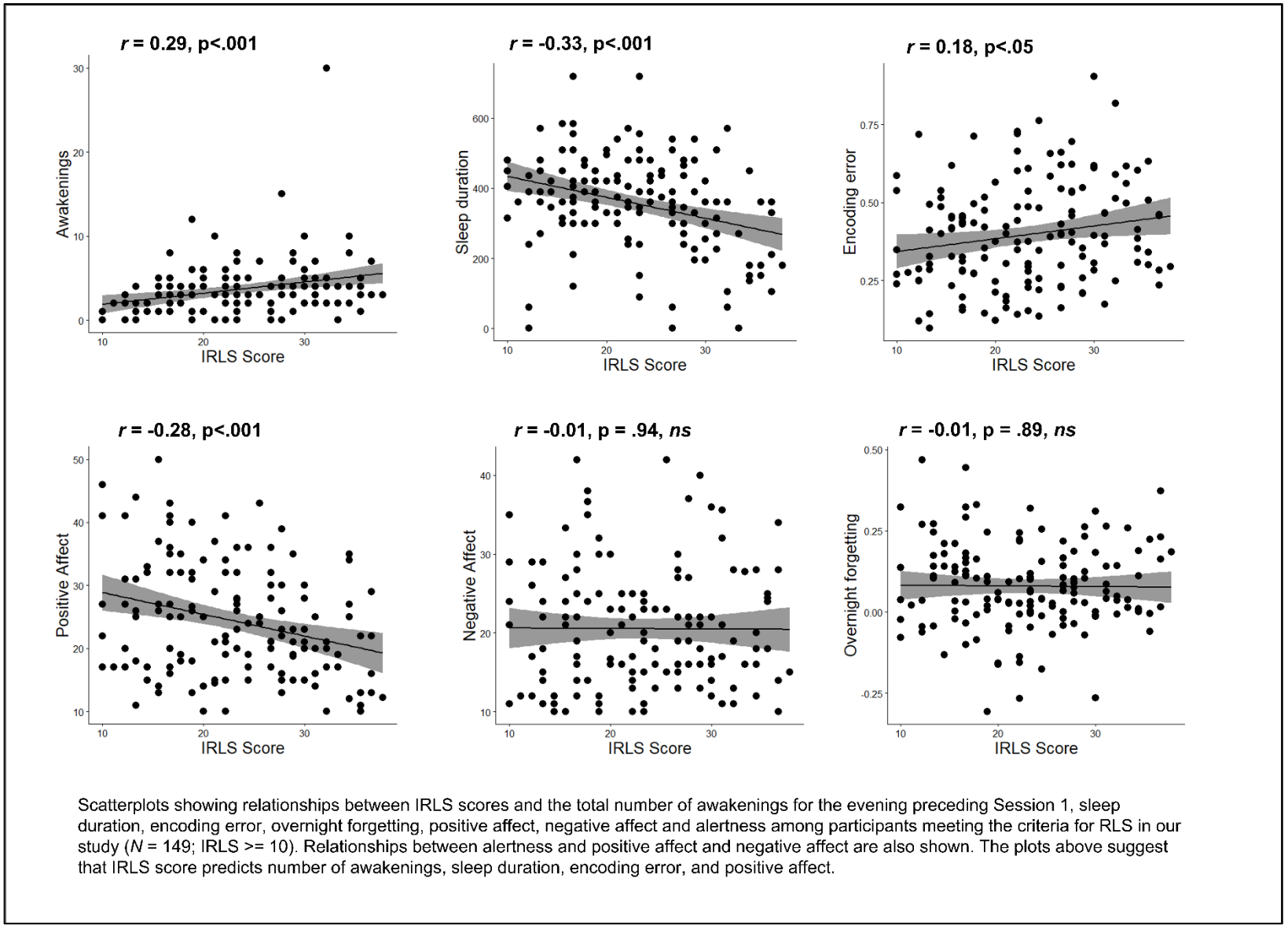
Figure 20 visualises scale and distributions for the spatial memory measures used:




*Note. Mean* Evening distance is level of encoding error and ranges from 0 (no error) to around 1.75 (very poor memory). Morning distance refers to how much is retained the following day and is on the same scale. Mean Overnight retention ranges from approximately -1.5 (substantial overnight improvement) to 1.5 (substantial overnight forgetting). Increasing value corresponds with poorer memory.

*Figure 20. Stacked histograms for spatial memory measures from Object Location task.*

Taking a similar approach to Cha and colleagues (2017), we classed a participant as having met the criteria for RLS if they scored 10 or greater on the IRLS, suggesting that they had at least moderate RLS symptoms. Figure 21 below shows associations between IRLS score and some of the studies' key dependent variables (e.g., encoding error, positive affect) among all participants who met the criteria for RLS (N = 149; Unmedicated = 64; Medicated = 85):




Scatterplots showing relationships between IRLS scores and the total number of awakenings for the evening preceding Session 1, sleep duration, encoding error, overnight forgetting, positive affect, negative affect and alertness among participants meeting the criteria for RLS in our study (*N* = 149; IRLS >= 10). Relationships between alertness and positive affect and negative affect are also shown. The plots above suggest that IRLS score predicts number of awakenings, sleep duration, encoding error, and positive affect.

*Figure 21. Scatterplot matrix focussing on associations with IRLS.*

Figure 21 suggests that as complaints of RLS symptoms increase in severity, sleep becomes more fragmented and more deprived, as would be expected in individuals with RLS. There is also some evidence of impairment in the expression of positive affect as symptom severity increases and a higher level of encoding error.

**H1: Relationships between group and encoding error.** To investigate the relationship between RLS and encoding error, Group (*RLS_unmedicated/ RLS_medicated/ Control*) was entered as a separate hierarchical step in a multilevel model with encoding error as the dependent variable and participant and trial (object) as random effects. The model also controlled for variation in age, education, alertness, and sex assigned at birth. As the data in Tables 46 and 47 show, only the coefficient for the unmedicated RLS group reached significance, suggesting that experiencing at least moderate symptoms in the past week was associated with increased encoding error relative to healthy controls for this group. The standardised coefficient was small for the unmedicated group ($\beta = 0.09$, $SE\_\beta = 0.02$) and non-significant for the medicated group, suggesting at least some benefit of medication use in alleviating memory impairment. Introducing Group as a separate step in the model uniquely contributed to the variance in encoding error, with the rest of the variance being explained by the intercept and the grouping structure (ICC = 0.20, Marginal $R^2 = 0.01$, Conditional $R^2 = 0.21$). Assumptions of linearity and independence of the variables and that the residuals were normally distributed were met.

*Table 47. Coefficient table for a hierarchical multilevel model with encoding error as dependent variable.*

| Predictor | | *b* | SE | *CI* | *p* |
|---|---|---|---|---|---|
| Model 1 | | | | | |
| | Age | 0.00 | 0.00 | 0.00 – 0.00 | **.010**** |
| | Sex | -0.02 | 0.00 | -0.06 – 0.02 | .305 |
| | Education | 0.02 | 0.02 | -0.01 – 0.05 | .196 |
| | Alertness | -0.00 | 0.01 | -0.01 – 0.01 | .696 |
| | Sleep duration | -0.00 | 0.00 | -0.00 – 0.00 | .079 |
| Model 2 | | | | | |
| | Age | 0.00 | 0.00 | 0.00 – 0.00 | .432 |
| | Sex | -0.02 | 0.00 | -0.06 – 0.02 | .271 |
| | Education | 0.01 | 0.02 | -0.02 – 0.04 | .581 |
| | Alertness | -0.00 | 0.01 | -0.02 – 0.01 | .565 |
| | Sleep duration | -0.00 | 0.00 | -0.00 – 0.00 | .272 |
| | RLS (unmedicated) | 0.10 | 0.02 | 0.05 – 0.15 | **<.001***** |
| | RLS (medicated) | 0.05 | 0.03 | 0.05 – 0.15 | .062 |

*Note.* N=452. Model comparison for hierarchical multilevel model testing group (control/RLS (unmedicated)/RLS (medicated) against the log transformation of the S1Distance whilst controlling for age, sex assigned at birth, highest level of education, alertness, participant (random effect) and object (random effect). Sleep duration is the self-reported nocturnal sleep duration for night preceding the completion of Session 1. Alertness is captured by the Stanford Sleepiness Scale.

*Table 48. Model comparison statistics for multilevel model described in Table 46.*

| Model | npar | Test | deviance | Chisq | df | Pr(>Chisq) |
|---|---|---|---|---|---|---|
| Model 1 (null) | 8 | | 4548.0 | | | |
| + Group | 10 | 1x2 | 4531.5 | 16.42 | 2 | **<.001***** |

**H1: Relationships between group and overnight forgetting.** The same multilevel modelling approach was applied with overnight forgetting as the dependent variable. One exception was that encoding error was added to the model as a control variable following the finding from Chapter 3 that the number of nocturnal awakenings only predicted overnight forgetting when encoding error was considered. Tables 48 and 49 below show the results:

*Table 49. Coefficient table for a hierarchical multilevel model with overnight forgetting as dependent variable.*

| Predictor | | *b* | SE | *CI* | *p* |
|---|---|---|---|---|---|
| Model 1 | | | | | |
| | Encoding error | -0.97 | 0.05 | -0.99 – (-0.95) | **<.001*** ** |
| | Age | 0.00 | 0.00 | -0.00 - 0.00 | .053 |
| | Sex | -0.01 | 0.02 | -0.05 – 0.03 | .476 |
| | Education | 0.01 | 0.02 | -0.02 – 0.05 | .498 |
| | Alertness | -0.00 | 0.01 | -0.01 – 0.01 | .921 |
| | Sleep duration | -0.00 | 0.00 | -0.00 – 0.00 | .301 |
| Model 2 | | | | | |
| | Encoding error | -0.97 | 0.05 | -0.99 – (-0.95) | **<.001*** ** |
| | Age | 0.00 | 0.00 | -0.00 - 0.00 | .972 |
| | Sex | -0.02 | 0.02 | -0.05 – 0.02 | .370 |
| | Education | 0.00 | 0.02 | -0.03 – 0.03 | .958 |
| | Alertness | -0.00 | 0.01 | -0.01 – 0.01 | .682 |
| | Sleep duration | -0.00 | 0.00 | -0.00 – 0.00 | **.034*** |
| | RLS (unmedicated) | 0.09 | 0.03 | 0.04 – 0.14 | **0.001**** |
| | RLS (medicated) | 0.05 | 0.03 | -0.01 – 0.10 | .114 |

*Table 50. Model comparison statistics for multilevel model described in Table 48.*

| Model | npar | Test | deviance | Chisq | df | Pr(>Chisq) |
|---|---|---|---|---|---|---|
| Model 1 (null) | 10 | | 6231.5 | | | |
| + Group | 12 | 1x2 | 6219.6 | 11.83 | 2 | **.003**** |

The data shows a very small - small positive association between unmedicated RLS and overnight forgetting (β = 0.07, SE_β = 0.02) whilst controlling for sleep duration. Interestingly, sleep duration was also negatively associated with overnight forgetting (β = -0.04, SE_β = 0.02).. As was the case for encoding error, there was no relationship between medicated RLS and overnight forgetting, suggesting at least some degree of symptom relief among RLS medication users. Introducing Group as a separate step in the model uniquely contributed to the variance, with the rest of the variance being explained by the intercept and the grouping structure (ICC = 0.19, Marginal $R^2$ = 0.43, Conditional $R^2$ = 0.54). Assumptions of linearity and independence of the variables and that the residuals were normally distributed were met.

**H1: Relationships between group and expressions of affect.** Hierarchical multiple regressions were used to test the relationships between the RLS groups and expressions of affect. Positive and negative affect were measured using PANAS in Session 2. One model was created with positive affect as the dependent variable, and another identical model was created with negative affect as the dependent variable. Control variables were age, sex, education level, alertness, and sleep duration. Tables 10 and 11 display the positive affect data and Tables 50 and 51 show the data for negative affect:

*Table 51. Coefficient table for a hierarchical multiple regression with positive affect as dependent variable.*

| Predictor | | *b* | SE | t | *p* |
|---|---|---|---|---|---|
| Model 1 | | | | | |
| | Age | -0.08 | 0.03 | -3.05 | **.002**** |
| | Sex | 1.27 | 0.93 | 1.37 | .34 |
| | Education | 1.04 | 0.78 | 1.34 | .171 |
| | Alertness | -2.57 | 0.29 | -8.93 | **<.001***** |
| | Sleep duration | -0.00 | 0.00 | -1.22 | .224 |
| Model 2 | | | | | |
| | Age | -0.03 | 0.04 | -0.88 | .379 |
| | Sex | 0.75 | 0.92 | 1.40 | .163 |
| | Education | 1.43 | 0.79 | 1.81 | .071 |
| | Alertness | -2.39 | 0.29 | -8.17 | **<.001***** |
| | Sleep duration | -0.00 | 0.00 | -1.40 | .163 |
| | RLS (unmedicated) | -2.57 | 1.16 | -2.21 | **.028*** |
| | RLS (medicated) | -3.08 | 1.33 | -2.32 | **.021*** |

*Note.* N=452. Model comparison for hierarchical multilevel model testing group (control/RLS (unmedicated)/RLS (medicated) against Positive affect whilst controlling for age, sex assigned at birth, highest level of education, alertness, participant (random effect) and object (random effect). Sleep duration is the self-reported nocturnal sleep duration for night preceding the completion of Session 1. Alertness is captured by the Stanford Sleepiness Scale. $R^2$ = 0.178, R2 adjusted = 0.169, p<.001*** for Model 1 (null); $R^2$ = 0.187, R2 adjusted = 0.174, p<.001*** for Model 2 (+ Group).

*Table 52. Model comparison statistics for hierarchical multiple regression described in Table 50.*

| Model | *Res.df* | *df* | *Test* | *F* | *p* | *ΔR²* |
|---|---|---|---|---|---|---|
| Model 1 (null) | 412 | | | | | 0.179 |
| + Group | 410 | 2 | 1x2 | 3.57 | **.029*** | 0.189 |

Here being part of either of both RLS subgroups predicted lower expression of positive affect. Being in the unmedicated group was associated with being 2.57 points lower on the PANAS positive affect scale than healthy controls and being in the medicated group was associated with being 3.08 points lower on the scale. These results suggest not only that having RLS is associated with impaired expression of positive affect, but that this association is stronger among our medicated participants. This is an interesting contrast to what we found for the RLS group and declarative memory above. Our results seem to tentatively suggest that medication use in our sample was benefitting declarative memory but not the expression of positive affect. The model explained 18.9% of the variance with 1% coming uniquely from our group variable.

*Table 53. Coefficient table for a hierarchical multiple regression with negative affect as dependent variable.*

| Predictor | | *b* | SE | t | *p* |
|---|---|---|---|---|---|
| Model 1 | | | | | |
| | Age | -0.07 | 0.02 | -0.12-(-0.02) | **<.01**** |
| | Sex | -1.60 | 0.83 | -3.23-0.02 | .05 |
| | Education | 1.43 | 0.71 | 0.03-2.83 | **.045*** |
| | Alertness | 0.58 | 0.27 | 0.06-1.10 | .**03*** |
| | Sleep duration | -0.00 | 0.00 | -0.01-0.00 | .30 |
| Model 2 | | | | | |
| | Age | -0.11 | 0.03 | -0.17-(-0.05) | **.001***** |
| | Sex | -1.53 | 0.83 | -3.15-0.10 | .07 |
| | Education | 1.19 | 0.72 | -0.22-2.61 | .10 |
| | Alertness | 0.45 | 0.27 | -0.09-0.98 | .10 |
| | Sleep duration | -0.00 | 0.00 | -0.10-4.72 | .39 |
| | RLS (unmedicated) | 1.98 | 1.09 | -0.17-4.13 | .07 |
| | RLS (medicated) | 2.31 | 1.22 | -0.10-44.72 | .06 |

$R^2$ = 0.036, R2 adjusted = 0.025, p<.01** for Model 1 (null); $R^2$ = 0.047, R2 adjusted = 0.032, p<.01** for Model 2 (+ Group).

*Table 54. Model comparison statistics for the hierarchical multiple regression model described in Table 52.*

| Model | *Res.df* | *df* | *Test* | *F* | *p* | *ΔR²* |
|---|---|---|---|---|---|---|
| Model 1 (null) | 441 | | | | | 0.027 |
| + Group | 439 | 2 | 1x2 | 2.44 | .09 | 0.028 |

In contrast to the analysis for positive affect, Group did not significantly predict expressions of negative affect, nor did it uniquely contribute to the variance, and only 2.8% of the variance was explained by the model (Tables 52 and 53, above). As for the low amount of variance explained, this is less surprising. Like what was observed in new parents in Experiment 3, a moderate - large negative relationship seems to exist between alertness and positive affect that does not for alertness and negative affect. Figure 22 below illustrates this by creating scatterplots among those who meet the criteria for RLS (N = 149 with IRLS >= 10; Unmedicated = 64; Medicated = 85) showing associations between affect and IRLS and between affect and alertness (SSS):

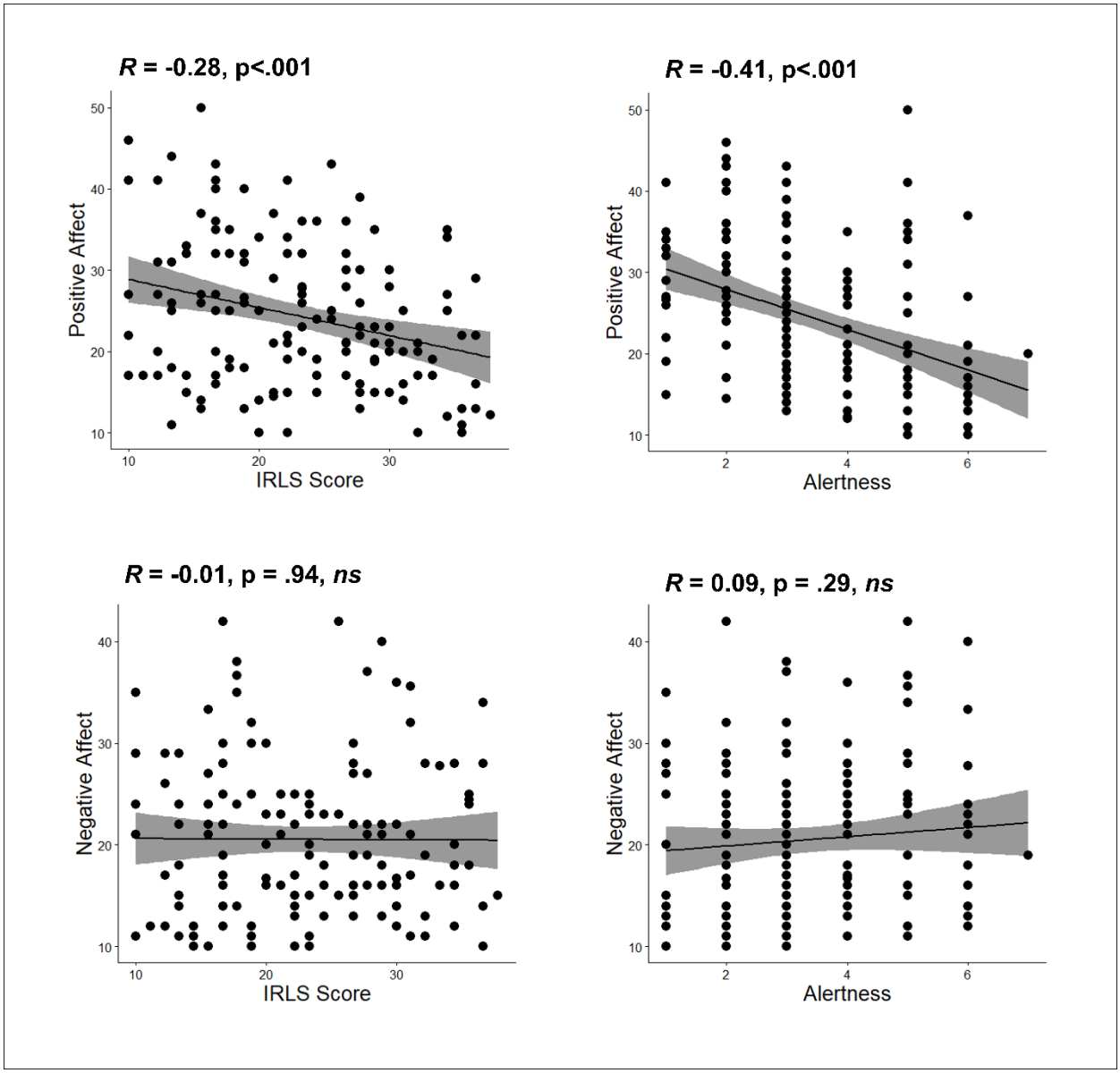



*Figure 22. Scatterplot matrix exploring associations between RLS severity, affect, and attention.*

**H1 Supplementary analysis: Relationships between group and awakenings.** One way to explain the pattern of results for H1 is that after controlling for sleep duration, the key thing which distinguishes the RLS groups from healthy controls is that their sleep is more fragmented. To test this, we carried out a similar multilevel model to those previously with the number of nocturnal awakenings lasting 5 minutes or longer for a typical night in the previous month as the dependent variable, as shown in Tables 54 and 55 below:

*Table 55. Coefficient table for a hierarchical multiple regression model with number of awakenings as dependent variable.*

| Predictor | | *b* | SE | *t* | *p* |
|---|---|---|---|---|---|
| Model 1 | | | | | |
| | Age | 0.05 | 0.07 | 6.80 | **<.001*** ** |
| | Sex | -0.05 | 0.24 | 0.10 | .83 |
| | Education | -0.10 | 0.21 | -0.68 | .64 |
| | Alertness | 0.24 | 0.08 | 2.88 | **.002**** |
| | Sleep duration | -0.00 | 0.00 | -4.22 | **<.001*** ** |
| Model 2 | | | | | |
| | Age | 0.01 | 0.01 | 1.51 | .14 |
| | Sex | -0.01 | 0.23 | -0.00 | .98 |
| | Education | -0.21 | 0.20 | -1.32 | .30 |
| | Alertness | 0.13 | 0.08 | 1.54 | .08 |
| | Sleep duration | -0.00 | 0.00 | -4.00 | **<.001*** ** |
| | RLS (unmedicated) | 0.67 | 0.31 | 6.66 | **.029*** |
| | RLS (medicated) | 2.42 | 0.34 | 2.28 | **<.001*** ** |

$R^2$ = 0.229, R2 adjusted = 0.220, p<.001*** for Model 1 (null); $R^2$ = 0.308, R2 adjusted = 0.297, p<.001*** for Model 2 (+ Group).

*Table 56. Model comparison statistics for the multiple regression model described in Table 54.*

| Model | *Res.df* | *df* | *Test* | *F* | *p* | *ΔR²* |
|---|---|---|---|---|---|---|
| Model 1 (null) | 441 | | | | | 0.220 |
| + Group | 439 | 2 | 1x2 | 25.72 | **<.001*** ** | 0.297 |

We chose this measurement of awakenings because the literature suggests that an awakening needs to last approximately 5 minutes to be remembered (Winser et. al., 2013). It is plausible that this level of measurement is less likely to be confounded by

subjective sleep quality (see O'Donnell et al., 2009, and Conte et al., 2022, for examples of significant negative associations between number of awakenings and subjective sleep quality, but also research by Harvey et al., 2008, which suggests that the reason the two constructs are related is because subjective sleep quality is driven by the number of awakenings). Firstly, both RLS groups were likely to have significantly more awakenings than healthy controls. The model explained 30.8% of the variance in awakenings, with 7.9% being directly attributable to Group. This is not in itself surprising and supports the inference that the observed deficits in declarative memory and affect are explained at least partially by sleep fragmentation. What is surprising, however, is the differences between RLS and controls for awakenings was larger for the medicated group, even when age was factored into the model (in general our RLS medicated group was significantly older than the unmedicated group). This is surprising because it would also predict significant declarative memory deficits among the unmedicated group, which were not observed.

***H2: Matched Group comparisons (N=118): Are there group differences between unmedicated RLS and matched controls for declarative memory and affect?***

The following section aims to answer similar questions to that in H1 but with a more dedicated focus on the influence of untreated RLS on declarative memory and affect. We approach these questions using matched comparisons between all our unmedicated RLS participants which were complete cases (N=59) and a control group matched on age and level of education using nearest neighbour propensity score matching (N=59). Table 56 below shows descriptive statistics for the post-matching sample and figure 23 visually compares the pre-post matching transformation of the control group to be comparable with the unmedicated RLS group:

*Table 57. Descriptive statistics for H2.*

| Variable | Healthy | RLS Unmedicated | Total |
|---|---|---|---|
| Encoding error | 0.31 (0.15) | 0.41 (0.17) | 0.35 (0.17) |
| Overnight Retention | 0.11 (0.12) | 0.08 (0.13) | 0.10 (0.13) |
| Positive Affect | 28.92 (7.86) | 25.35 (8.88) | 27.14 (8.54) |
| Negative Affect | 18.91 (7.28) | 21.11 (8.40) | 20.01 (7.91) |
| TST (S1; mins) | 452.86 (95.59) | 377.86 (110.52) | 415.36 (109.58) |
| TST (S2; mins) | 452.86 (95.02) | 417.14 (108.56) | 435.00 (103.18) |
| Awakenings (S1) | 1.37 (1.44) | 2.73 (2.03) | 2.05 (1.88) |
| Awakenings (S2) | 1.56 (1.70) | 2.48 (2.06) | 2.01 (1.94) |
| SSS1 | 3.06 (1.41) | 3.52 (1.46) | 3.29 (1.45) |
| SSS2 | 2.65 (1.18) | 3.29 (1.41) | 2.97 (1.33) |

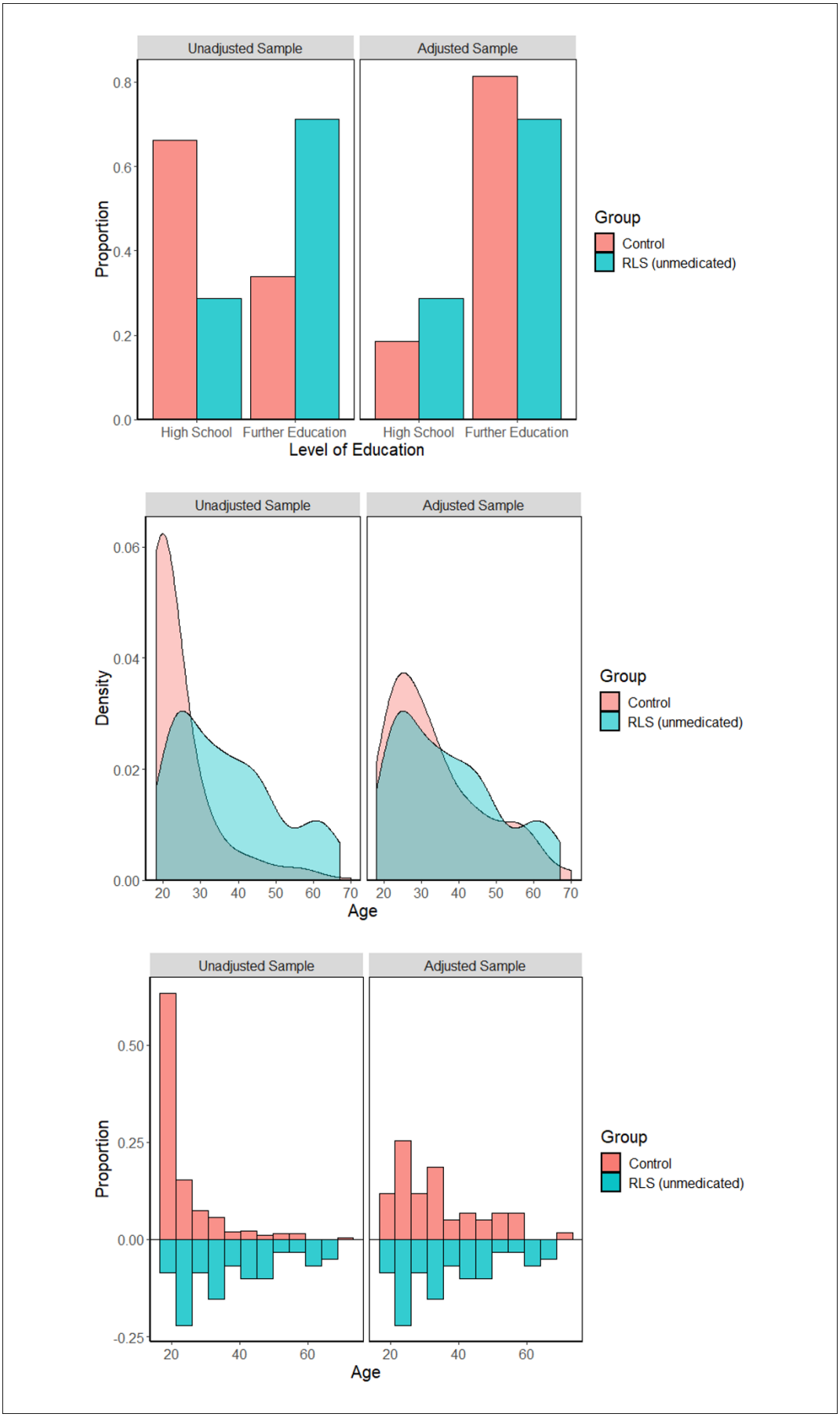



*Figure 23. Pre and post matching distributions.*

**H2: Group comparisons.** Two phases of matched-group comparisons were carried out. Firstly, matched t-tests were carried out between groups for our dependent variables of interest (namely Encoding error, Overnight forgetting, Positive affect, and Negative affect, Sleep duration, and Number of awakenings) as well as for potentially confounding variables (Age, Alertness). Table 57 below shows the results of these analyses:

*Table 58. Results from matched independent t-tests.*

| DV | IV Mean | IV SD | DV Mean | DV SD | *t* | *p* | *d* |
|---|---|---|---|---|---|---|---|
| Encoding error | 0.41 | 0.17 | 0.31 | 0.15 | -3.47 | **<.001*** ** | 0.61 |
| Overnight retention | 0.08 | 0.13 | 0.11 | 0.12 | 1.16 | .25 | *na* |
| Positive affect | 25.35 | 8.88 | 28.92 | 7.86 | 2.38 | **.019*** | 0.52 |
| Negative affect | 21.11 | 8.40 | 18.91 | 7.28 | -1.58 | .12 | *na* |
| Sleep duration | 418.10 | 90.30 | 458.57 | 63.80 | 2.91 | **.004**** | 0.56 |
| Awakenings | 1.81 | 1.67 | 0.90 | 1.17 | -3.51 | **<.001*** ** | 0.61 |
| Age | 37.30 | 14.00 | 34.22 | 12.50 | -1.30 | .19 | *na* |
| Alertness (S1) | 3.52 | 1.46 | 3.08 | 1.41 | -1.80 | .07 | *na* |
| Alertness (S2) | 3.29 | 1.41 | 2.65 | 1.18 | -2.74 | **.007**** | 0.53 |

***Note.*** Independent t-tests comparing means between an unmedicated group of individuals with Restless Legs Syndrome and healthy controls matched for age and education. *p < .05, **p < .01, ***p < .001. Encoding error in normative screen units; Overnight retention in normative screen units; Positive affect on PANAS scores; Negative affect on PANAS scores; Sleep duration in self-reported minutes for a typical night in previous month; Awakenings in self-reported number of nocturnal awakenings lasting 5 minutes or longer for a typical night in previous month; Age in years; Alertness on Stanford Sleepiness Scale scores.

The observed pattern of results conforms with those of the multiple regressions in H1. Specifically, encoding error was significantly higher and expression of positive affect was specifically lower for the RLS group. Overnight forgetting did not differ between groups which in the absence of any control of Encoding error was predicted. Like the data shown in Tables 48 and 49 above, there was no difference between groups for expression of negative affect which was not expected. This again is consistent with either Type 1 error or a combination preserved REM continuity and engaging with the study at moments of the day where negative affect is low (essentially at times when they feel like it). As predicted, The RLS group on average slept for less and woke up more which was predicted. One potentially confounding influence to consider is that Alertness differed significantly between groups for Session 2. Intuitively, the RLS group was less alert than control. The groups did not significantly differ on age which suggests that the matching process was effective.

*Table 59. Matched ANCOVAs between RLS (Unmedicated) and Control.*

| DV | EMM (IV) | EMM(DV) | *F* | *p* | $\eta_p^2$ |
|---|---|---|---|---|---|
| Encoding error | 0.40(0.02) | 0.32(0.02) | 7.95 | **<.01**** | 0.06 |
| Overnight retention | 0.09(0.16) | 0.11(0.16) | 0.88 | .349 | *na* |
| Positive affect | 25.4(1.07) | 28.9(1.07) | 5.39 | **.022*** | 0.04 |
| Negative affect | 21.1(1.12) | 18.9(1.12) | *t*(-1.9) | .054 | *na* |

***Note.*** ANCOVAs comparing means between an unmedicated group of individuals with Restless Legs Syndrome and healthy controls matched for age and education whilst controlling for sleep duration. *p < .05, **p < .01, ***p < .001. Encoding error in normative screen units; Overnight retention in normative screen units; Positive affect on PANAS scores; Negative affect on PANAS scores; EMM= Estimated Marginal Mean (Standard error in brackets). The ANCOVA for Negative affect was fitted using a robust version of standard error which is heteroscedasticity consistent the test of which returns a t-value.

The above series of ANCOVAs (Table 58) were carried out to control for the influence of sleep duration and test the inference that something about other than sleep duration (we argue sleep fragmentation) in RLS can uniquely explain observed declarative memory and affective deficits in RLS. As predicted, The RLS group on average was higher in Encoding Error and the expression of Positive affect whilst controlling for sleep duration. Also predicted, there was no difference in Overnight

forgetting in the context of varying levels of Encoding error. Negative affect was compared between groups using a robust version of standard error which is heteroscedasticity consistent and was not significant. These results suggest that memory deficits and some of the poor affect seen in RLS are related to factors other than sleep duration alone. The most likely candidate explanation here is that sleep fragmentation can influence declarative memory and affect even in the context of already existing partial sleep deprivation.

***H3: Multilevel model (N=149): Is the number of nocturnal awakenings associated with overnight forgetting whilst controlling for sleep duration?***

In Experiment in chapter 3 it was observed that the number of awakenings lasting 5 minutes or longer for a typical night in the previous month was positively associated with Overnight forgetting whilst controlling for sleep duration in a sample which consisted primarily of new parents and students. However, sleep duration was healthier than anticipated in this sample (M= 435.8; SD = 91.9, mins), leaving it unclear if the sleep fragmentation effect would still contribute to the variance when sleep duration was lower and more varied. To test this, we carried out a final multilevel model using only RLS participants who specifically took part in this study and were not involved in our study involving new parents (N=149; 64 RLS unmedicated; 85 RLS medicated). Table 59 below shows that we were somewhat successful in recruiting a sample both lower on average and more varied in sleep duration than what was observed in Experiment 3 (Experiment 3 figures: TST (S1) - Mean = 429.7 mins; SD = 90.6 mins; TST (S2) - Mean = 435.8 mins; SD = 91.9 mins). See the appendices for a histogram showing relatively increased representation of partial sleep deprivation (sleep duration < 6 hours) in the current sample.

*Table 60. Descriptive statistics for H3.*

| Variable | RLS Unmedicated | RLS Medicated | Total |
|---|---|---|---|
| Encoding error | 0.41 (0.17) | 0.37 (0.16) | 0.39 (0.17) |
| Overnight Retention | 0.08 (0.13) | 0.08 (0.13) | 0.08 (0.13) |
| Positive Affect | 25.4 (8.88) | 23.8 (8.98) | 24.6 (8.93) |
| Negative Affect | 21.1 (8.40) | 20.2 (7.01) | 20.65 (7.71) |
| TST (S1; mins) | 376.17 (110.46) | 341.47 (147.87) | 358.82 (129.17) |
| TST (S2; mins) | 416.25 (107.93) | 373.41 (124.26) | 394.83 (116.1) |
| Awakenings (S1) | 2.75 (2.02) | 4.26 (3.96) | 3.51 (2.99) |
| Awakenings (S2) | 2.47 (2.05) | 4.07 (3.32) | 3.27 (2.69) |
| SSS1 | 3.53 (1.45) | 3.67 (1.37) | 3.6 (1.41) |
| SSS2 | 3.27 (1.41) | 3.58 (1.50) | 3.43 (1.46) |

**H3: Relationships between awakenings and Overnight forgetting.** An identical model to that displayed in Tables 48 and 49 above was created, substituting the total number of awakenings for the night before Session 2 for our Group variable to test whether subjectively estimated sleep fragmentation was associated with Overnight forgetting whilst controlling for sleep duration. Tables 60-62 below show the results:

*Table 61. Coefficient table for a hierarchical multilevel model with overnight forgetting as dependent variable.*

| Predictor | | *b* | SE | *CI* | *p* |
|---|---|---|---|---|---|
| Model 1 | | | | | |
| | Encoding error | -0.98 | 0.02 | -1.01 - -0.94 | **<.001*** ** |
| | Age | -0.00 | 0.00 | -0.00 - 0.00 | .789 |
| | Sex | 0.03 | 0.03 | -0.03 - 0.10 | .299 |
| | Education | 0.02 | 0.03 | -0.04 - 0.07 | .570 |
| | Alertness | 0.02 | 0.01 | -0.00 - 0.03 | .066 |
| | Sleep duration | -0.00 | 0.00 | -0.00 – 0.00 | .915 |
| Model 2 | | | | | |
| | Encoding error | -0.99 | 0.02 | -1.02 - -0.05 | **<.001*** ** |
| | Age | -0.00 | 0.00 | -0.00 - 0.00 | .676 |
| | Sex | 0.04 | 0.03 | -0.03 - 0.10 | .246 |
| | Education | 0.01 | 0.03 | -0.04 - 0.07 | .596 |
| | Alertness | 0.02 | 0.01 | 0.00 - 0.04 | .**042*** |
| | Sleep duration | -0.00 | 0.00 | -0.00 – 0.00 | .873 |
| | Awakenings | -0.00 | 0.00 | -0.01 – 0.00 | .349 |

*Note.* Model comparison for hierarchical multilevel model testing Awakenings against Overnight retention whilst controlling for age, sex assigned at birth, highest level of education, alertness, encoding error, and participant (random effect). Sleep duration is the self-reported nocturnal sleep duration for night preceding the completion of Session 1. Alertness is captured by the Stanford Sleepiness Scale. Marginal $R^2$ = 0.469, Conditional $R^2$ = 0.559, p<.001 for Model 1; Marginal $R^2$ = 0.469, Conditional $R^2$ = 0.559, p<.001 for Model 2

*Table 62. Model comparison statistics for the multiple regression model described in Table 60.*

| Model | npar | Test | deviance | *Chisq* | *df* | *Pr(>Chisq)* |
|---|---|---|---|---|---|---|
| Model 1 (null) | 10 | | 2443.5 | | | |
| + Awakenings | 11 | 1x2 | 2442.6 | 0.88 | 1 | .35 |

*Table 63. Model comparison using different subjective measures of awakenings.*

| Model | *β* | *SE_beta* | *Marginal $R^2$* | *Conditional $R^2$* | *p* |
|---|---|---|---|---|---|
| Model 1 (null) | | | 0.468 | 0.559 | |
| + Awa_Total_S2 | -0.02 | 0.03 | 0.469 | 0.559 | *ns* |
| + Awa_Total_Month | 0.03 | 0.03 | 0.468 | 0.557 | *ns* |
| + Awa_5_S2 | -0.01 | 0.03 | 0.469 | 0.559 | *ns* |
| + Awa_5_Month | 0.00 | 0.03 | 0.468 | 0.559 | *ns* |

The key results from this set of analyses are that none of our subjective measures of awakenings predicted overnight forgetting, even when controlling for level of encoding error. This set of results contradicts H1 above, and Experiment 3, which are both consistent with the number of awakenings being significantly associated with overnight forgetting.

## Discussion

The present study aimed to better understand the association between nocturnal awakenings declarative memory and affect. It also sought to investigate these relationships among individuals with RLS. We report several key findings with regards to the encoding and consolidation of declarative memories. Firstly, among a sample of individuals with RLS both taking and not taking medication and healthy controls from Study 2, membership of the unmedicated RLS population significantly predicted deficits in encoding and overnight forgetting. Importantly, the consolidation effect only existed when the level of encoding was controlled for in the model. The effect sizes were small but significant even whilst controlling for sleep duration. Sleep duration was also positively related to overnight forgetting. Unexpectedly, when subjective sleep measures were used for the number of awakenings instead of inferring sleep fragmentation from

RLS group membership, awakenings did not predict overnight forgetting among a sample consisting only of those meeting our criteria for RLS. Considering now the relationship between awakenings and affect, both RLS groups (medicated and unmedicated, who both had significantly higher number of awakenings than control) expressed less positive affect than controls. Surprisingly, there were no group differences in negative affect.

Despite very few investigations into the association between RLS and declarative memory deficits, the current results are in line with the strong predictions suggested by SHY and systems consolidation and are the first to show clear declarative memory deficits among RLS. However, it should be pointed out early in the discussion that the results here do not include those from objective sleep measurements and are therefore preliminary and tentative in nature. For encoding, SHY suggests that since RLS individuals experience chronic sleep fragmentation, overnight synaptic restoration will be disturbed to the effect that next day learning reliant on this restoration process will be impaired (e.g., Tononi and Cirelli, 2014; Born, 2012). For consolidation, systems consolidation suggests that the chronic sleep fragmentation typical in RLS will disrupt the delicate synchrony of oscillations involved in the integration of newly encoded memories into long-term storage networks in the neocortex. This prediction is especially strong given that Cha and colleagues (2017) observed lower spindle density and poor coordination between spindles and slow oscillations in RLS using polysomnography. Our results are also in line with those reported in Study 2, in which increased nocturnal awakenings lasting 5 minutes or longer were associated with increased encoding error and overnight forgetting. Our resultsmight be used to extend these findings by suggesting that sleep fragmentation uniquely contributes to the variance in encoding and consolidation even in the context of highly variable and partially deprived sleep duration. Finally, our results compliment Experiment 3 and further support the idea that the level of encoding needs to be considered when investigating consolidation. This may also reconcile our results with those of Sargeeva and colleagues (2017), who did not find any declarative memory deficits among individuals with a similar disorder to RLS (PLMD) using a 60% encoding criterion. The guiding principle here seems to be that the more that is learned, the more that can be forgotten.

However, our final analysis (focusing on RLS participants alone) did not find any significant association between awakenings and overnight forgetting and conflicts with

the above interpretation. It was initially thought that the increased sleep fragmentation and partial sleep deprivation which characterises RLS would have created the ideal conditions to detect an association between awakenings and overnight forgetting if it existed. Indeed, it may be the case that no association between the two exists, and the association between RLS and overnight forgetting in H1 was driven by a variable other than the number of awakenings. However, another intriguing possibility which reconciles the results is that a significant association between awakenings and overnight forgetting does exist, but subjective sleep measures are not reliable predictors of individual differences in studies with clinical samples. There is evidence to suggest that there is poorer concordance between objective and subjective measures of sleep in clinical populations relative to healthier ones (e.g., Conley et. al., 2019; Biddle et al., 2015; Biddle et al., 2017; see a series of mega-analyses in Chapter 5 strongly supports this line of reasoning). If true, this would suggest that the positive results found in Study 2 were facilitated by a greater degree of concordance between objective and subjective sleep measure and that despite greater levels of sleep disruption in the current study there was also an accompanying deterioration in the accuracy of the subjective sleep measures we had to rely on. Finally, it is also possible that the subjective measure may be sufficient for detecting a group difference, but not sensitive enough to detect individual differences within the RLS sample (see Hedge, Powell, and Sumner, 2018, for discussion). It is difficult to know which explanation to prioritise without further research.

Concerning the association between RLS and expressions of positive and negative affect, our results were expected for positive affect and surprising for negative affect. For positive affect, given that individuals with RLS often have or go on to develop mood disorders (Becker and Sharon, 2014), it is not surprising that there is evidence of impaired mood at the level of daily expression of affect. As well as this, since individuals with RLS experience frequent nocturnal awakenings, and that these are associated with decreased expression of positive affect and increased expression of negative affect (as was found in Study 2, also see Bonnet & Arand, 2003), it was strongly predicted that RLS would be associated with decreased expression of positive affect in this study. This prediction was especially strong among RLS since positive affect is thought to be regulated by SWS, and the circadian component of RLS symptoms pools in the first half of the night which is dominated by SWS (Krause et al., 2017; Guo et al., 2017).

For negative affect, however, the lack of difference between RLS and control in our analyses goes against the grain of the literature just discussed. In Experiment 3, for example, increased nocturnal awakenings were associated with increased expression of negative affect in a sample of new parents and healthy controls. There are two candidate explanations worth discussing. The first is simply that we failed to detect an effect that was in fact there. In our multilevel models and ANCOVAs, the effects were trending close to the significance threshold. It is possible that with a more sensitive measure of negative affect that the effect may have been detected. Another intriguing possibility is that since negative affect is thought to be regulated by REM sleep (Walker and Goldstein, 2014), and since the majority of REM sleep takes place at a time when RLS individuals typically experience a circadian-driven relief of symptoms (Guo et al., 2017), RLS may be associated with a disproportional disruption to SWS-governed processes (here positive affect) than REM-governed processes (here negative affect). Although we were aware of this possibility, we thought that expressions of negative affect might still be more frequent among RLS due to the daytime experience of uncomfortable and painful symptoms. One possibility is that participants took part in the study at times when their daytime symptoms weren't as severe, and this was reflected in our measures of affect.

Whilst studies exist examining the effects of specific RLS treatments, this study is the first to our knowledge to test the relationship between general medication use (i.e., a random selection of the many legitimate treatments in circulation within the RLS community) and cognitive and affective outcomes. This was largely exploratory and an opportunistic use of the medicated RLS sample which would have otherwise been excluded from the study. Our finding that there were significant group differences between medicated RLS individuals and healthy controls for measures of encoding error and overnight forgetting suggests very broadly and tentatively that taking medication often given to individuals complaining of RLS symptoms is to some degree protective of the formation and storage of declarative memories. This is consistent with previous findings of improved memory in RLS after taking dopamine agonists, the most common and widely used therapy (Jung, 2015; O'Regan and Anderson, 2015). This apparent benefit was not evident when levels of expression of positive affect were compared across groups, with both RLS groups expressing significantly less positive affect than healthy controls. We are not aware of any literature looking at medication use and the expression of positive affect to compare this finding to. Future investigations might seek

to further explore whether RLS medication use protects or restores declarative memory formation without modifying the expression of daytime positive affect. If this were to be the case and supported by further and more robust empirical evidence, it would be very intriguing to better understand the underlying reasons since both processes are thought to be governed by SWS and it is unclear how one could be manipulated without also affecting the other.

Finally, like the discussion in Experiment 3, effect sizes observed were small to very small. It is therefore likely the case that if they are meaningful, it would be over the long term (see Princeton and Miller, 1992, for general discussion, and Mucci et al., 2000, Brink et al., 2010, and Tassi et al., 2010, for examples of small yet meaningful longitudinal effects within the sleep fragmentation literature).

The present study has several limitations which should be considered when taking the results into account. The first limitation is that it cannot be shown that the group differences observed in H1 and H2 were actually attributable to sleep fragmentation. There are a number of potential differences between the groups which could have driven the effects observed. Not measuring sleep objectively in this study makes it unknown whether they restless leg had less slow wave sleep than the healthy controls.However, a recent meta-analysis on the polysomnographic features of sleep in Restless Legs Syndrome carried out by Geng et al., (2022) found no statistical difference in percentage of slow wave sleep across the 26 studies it identified. Similarly, it is also entirely possible that the Restless Legs Syndrome group were also hindered in their task performance by other cognitive deficits (e.g., attention and executive function; Gamaldo et al., 2008).

Another key limitation is that when direct measurement of the number of awakenings was analysed, no association was found between them and overnight forgetting. Whilst an attempt to reconcile these results was made above the results from H3 are contradictory to those reported in H1 and H2 and show the clear need for further research with greater reliance upon objective measures of sleep, ideally at the level of oscillatory activity. Given the relatively small number of awakenings reported here subjectively compared to a typical polysomnography or actigraphy study (e.g., 13 per night in Lee et al., 2009), it is possible that the sensitivity of our awakenings variable was lower than optimal. This could explain the contradictory findings observed between H3

and H1 and H2, but there is not enough evidence to conclude this with any degree of confidence without further investigation. Building from the discussion in the introduction about the relative strengths and weakness of sleep measures in RLS, it seems to be the case that the nuance offered by objective measurement is more suited to investigating the relationship between awakenings and declarative memory than the gist offered by self-report measures.

Finally, our eventual samples were overwhelmingly female (mean across analyses = 76.76%) and attrition was high. It is possible therefore that participants who failed to complete the study systematically differed from those who did on a relevant characteristic, for example on degree of sleep disruption.

There are two broad categories of future directions worth discussing. The first relates to suggested steps taken to validate some of the intriguing results reported in this study. The second relates to future directions which build from these observations. An important first step is to marry the pattern of declarative memory impairment observed here with the PSG work conducted by Cha and colleagues (2017). This would also allow for objective measurement of sleep fragmentation variables. However, determining the unique contribution of sleep fragmentation to declarative memory deficits using PSG would be very difficult. This is because it is a major challenge to control for sleep duration whilst at the same time manipulating sleep fragmentation and controlling for sleep duration statistically requires many participants which would be time consuming and costly using a PSG paradigm. Future investigations might seek to use wrist actigraphy, which has been shown to have good agreement with PSG (but less so in clinical populations, Conley et al., 2019; McCall and McCall, 2011). A PSG paradigm, however, would be optimal for validating our analyses of affect. We were limited in that our measures had to be as concise as possible to protect against attrition. Future investigations might make use of PSG and much more comprehensive measures of positive and negative affect, e.g., PANAS-X (Positive and Negative Affect Schedules - Expanded form, Watson and Clark, 1994).

Future research might also begin to explore whether the disruption caused by sleep fragmentation and sleep deprivation are qualitatively different. There are several plausible adverse consequences to repeatedly disrupting a delicate synchrony of activity

(consolidation) which might not arise by not allowing the processes to start at all. These include but are not limited to affecting the efficiency and accuracy which slow oscillations, spindles, and sharp-wave ripples synchronise; limiting the hippocampus's capacity to operate as a short-term storage site, and; corrupting otherwise healthy neocortical networks with broken or partial memory traces. There are currently no references we are aware of in the sleep and memory literature to what we are calling *partial correlation*. As well as this, it is entirely plausible that rebooting active systems consolidation processes repeatedly within a sleep fragmentation context will come at a cost, for example neurotransmitter resources also used in the daytime (like Walker and Goldstein's 2014 model of emotion calibration). Finally, the halt-mechanism of consolidation that sleep fragmentation triggers is not currently understood. As such, it is not yet known whether consolidation can be partial or whether it is an all-or-nothing process. Such questions are likely out of the reach of human behavioural paradigms and will require an animal neurobiological approach.

### *Conclusions*

Here we aimed to investigate the associations between sleep and declarative memory and affect. Specifically, the current investigation evaluated the unique contribution of sleep fragmentation to explained variance in declarative memory (encoding and consolidation) and affect (positive and negative). We also aimed to test strong theoretical predictions that memory and affect would be impaired in RLS. Finally, we opportunistically explored the influence of unspecified medication use within our RLS sample. Hierarchical models, matched t-tests and matched ANCOVAs were used to explore these questions. For declarative memory, unmedicated RLS was associated with increased encoding error and increased overnight forgetting when the level of encoding was controlled for, suggesting that the continuity of sleep can impact upon the formation and storage of declarative memories even when the duration of sleep is controlled for. Importantly, the current results also suggest that this small effect persists even in the context of lower levels of sleep duration as opposed to the relatively healthy sleep duration observed in Experiment 3, adding to the argument that sleep fragmentation can be uniquely detrimental to declarative memory, but to the extent that the group differences observed were in fact driven by sleep fragmentation, which cannot be conclusively stated in this study. As well as this, when subjective measures of

awakenings were used to approach this question, no association was found between the number of awakenings and overnight forgetting. This was surprising given the degree of sleep disruption observed among our RLS participants and highlights the need to apply caution when drawing too strong conclusions from these results. Intriguingly though, these apparently contradictory results may be reconciled by the poor concordance between subjective and objective sleep measures in clinical samples. An interesting dichotomy was observed with regards to the association between RLS and affect, such that the expression of positive affect was impaired in RLS whilst the ordinary expression of negative affect seemed unchanged compared to healthy controls. This contrasts with the results of Study 2 in which increased awakenings significantly predicted both decreased positive affect and increased negative affect. One explanation for this is that the regulation of negative affect is left unscathed in RLS since the circadian component of the disorder may disproportionately target SWS relative to REM, which is thought to govern negative affect. Finally, we also found that the random snapshot of unspecified medication use in our RLS sample seemed to be at least somewhat protective or restorative towards declarative memory formation and storage since only the unmedicated group differed from healthy controls. Interestingly, there was no evidence to suggest that medication use was helping to regulate the expression of positive affect, which is interesting from the point of view of sleep research since declarative memory and positive affect are thought to be governed by SWS and it therefore unclear how one could be manipulated without affecting the other. Future research might seek to validate some of the potentially very informative trends identified in this study using objective measurements. As well as this, future research might also seek to explore whether there are qualitative differences in the deficits caused by sleep deprivation and sleep fragmentation respectively and what these mean for active systems consolidation. Overall, these results are a promising yet tentative first step towards better understanding the link between sleep fragmentation and declarative memory formation and storage in humans.

# Chapter 5

# Relatedness and agreement of subjective and objective sleep measures in healthy adults and those with sleep disorders

## Abstract

Measuring sleep using subjective report is cost-effective, efficient, and common in large-scale studies. Yet, the consensus is that more objective measures, such as polysomnography and wrist actigraphy, measure sleep with the greatest accuracy (Lubas et al., 2022; Baillet et al., 2016). As well as this, and despite the above advantages associated with the use of sleep diaries, there is an insufficient empirical foundation to support the assumption that they are capturing the same information as with the more objective approaches. And if they are, it is unclear that they do so accurately enough to warrant their widespread use in sleep research and clinical settings. Regarding the latter, recent meta-analytic evidence also exists to suggest that sleep measurement is less accurate using actigraphy compared to polysomnography amongst clinical relative to healthy populations, begging the question as to whether the same deterioration in accuracy exists between subjective measures and more objective ones among those with sleep disorders. The current investigation sought to establish the strength of relatedness (extent to which two variables influence each other) and agreement (extent to which two measurements agree with one another) between sleep diaries and more objective measures (polysomnography and actigraphy) among healthy participants and those with sleep disorders for Total Sleep Time (TST), Number of Awakenings (NA), Sleep Onset Latency (SOL), Wake After Sleep Onset (WASO), and Sleep Efficiency (S.E). A rapid review was conducted which identified 13 studies from which we accessed and pooled raw data to calculate effect sizes for relatedness using a mega-analytic mixed effects paradigm. Intraclass Correlation Coefficients and Bland-Altman plots were used to assess agreement among these method-parameter pairs. Relatedness and agreement between subjective and more objective measures was poor, particularly among clinical populations. One exception is that we found a large effect size for the relationship between actigraphy, and diary measured total sleep time among healthy participants and a moderately sized intraclass correlation coefficient for the two measures. Subjective

sleep measures should be used in combination with more objective measures where possible, and alternative strategies should be considered among clinical populations. The low level of agreement among measures is consistent with the possibility that subjective sleep measures are capturing qualitatively different information, and future research should explore this possibility further.

## Introduction

A consensus is held among clinicians and sleep researchers that polysomnography (PSG), which is the systematic physiological measurement of sleep using electroencephalography to capture brain activity and electrooculography to capture chin movements (Rundo and Downey, 2019), is the most accurate tool for the measurement of sleep (Lubas et al., 2022; Baillet et al., 2016; Conley et al., 2019). However, PSG is costly, labour intensive and time consuming. Therefore, alternative approaches are used in contexts requiring cheaper, quicker, or more scalable output. Notable among these include wrist actigraphy devices, which use movement to estimate sleep parameters (Martin, and Hakim, 2011), and self-report in the form of questionnaires or diaries, which involve the patient or participant using their memory or record-keeping to estimate their sleep parameters (e.g., Biddle et al., 2015). PSG and actigraphy are commonly found to show high degrees of statistical agreement across the most common measures of sleep parameters (e.g., Total Sleep Time, Number of Awakenings, Sleep Onset Latency, Wake After Sleep Onset, and Sleep Efficiency; Conley et. al., 2019; McCall and McCall, 2011).

However, despite often being used in large-scale studies, very little research has been carried out to investigate the underlying assumption driving the use of subjective sleep measurements, namely that they are indeed related to (i.e., the extent to which two outcomes are influenced by one another, e.g., sunshine and happiness) and in agreement with (the extent to which two measurements capture the same construct, e.g., a clock and a sundial) more objective measures. As well as this, there is some evidence to suggest that the degree to which subjective and objective measures of sleep are related and in agreement with one another is influenced by individual characteristics, most notably of which is whether the individual has a chronic clinical condition such as a sleep disorder (Bianchi et. al., 2013; De Francesco et al., 2021; Werner et al., 2016). This

study aims to bring together an important yet inconsistent literature on the use of subjective sleep measures and quantify both the strength of relationship and the extent of statistical agreement between subjective and objective measures. We seek to do so for a wider range of variables than most studies typically consider, and to quantify the influence of having a sleep disorder on these observed relationships.

It is important to note that two phenomena can be strongly related to each other without agreeing and vice versa (see Lui et al., 2016, for discussion). Statistical relatedness refers to the extent to which changes in variables are associated with each other and if this association is linear, metrics like the Pearson's Correlation Coefficient can be used to assess the strength and direction of this association. Statistical agreement, however, refers to the extent to which two variables, typically measurements, concord with each other (see Liu, Tang, Lu, Feng, and Tu, 2016, for detailed discussion). Statistical agreement can be measured using Intraclass Correlation Coefficient, which is commonly used to assess whether one measure (typically a cheaper or more efficient one) can replace another (typically a more expensive or laborious one). Koo and Li (2016) published commonly used guidelines for the interpretation of the ICC which we have adopted for both discussion of the studies below and for our eventual analyses. As such, ICCs range between 0 and 1, with 0 indicating no agreement and 1 indicating complete agreement. Within this, ICCs below 0.5 are typically regarded as poor, between 0.5-0.75 as moderate, between 0.75-0.90 as good, and above 0.90 as excellent. Therefore, for the interest of consistency and interpretability of the remainder of the discussion, we will report the relevant ICCs as reported in each study in the literature and then interpret that ICC based on Koo and Li's guidelines even a different set of guidelines (e.g., Cicchetti's 1994 guidelines) was used to interpret the ICC in that study.

A particularly relevant publication to the current discussion is that of Conley and colleagues in 2019. The authors carried out a systematic review and meta-analysis of the concordance between sleep parameters measured by actigraphy and by polysomnography. 96 studies were included in their analyses, and concordance between actigraphy and PSG was calculated for Total Sleep Time (TST), Number of Awakenings (NA), Sleep Onset Latency (SOL), Wake After Sleep Onset (WASO), and Sleep Efficiency (S.E). As well as this, effect sizes (ICCs) were compared between healthy and clinical populations. Their analyses included 14 studies within which the samples had

chronic conditions, for example low back pain, depression, insomnia, and heart failure. The key findings of their investigation are that actigraphy overestimates sleep, underestimates wake, and that these differences are exacerbated by having a chronic condition. General overestimation of sleep was evident when considering TST and S.E, and underestimation of wake was evident when considering SOL and WASO. Measurements for both groups, however, were within clinically accepted guidelines as outlined by the American Academy of Sleep Medicine (Smith et al., 2018). Conley and colleagues' investigation is a highly useful reference point for researchers and clinicians looking to measure sleep parameters. The current investigation seeks to expand on it by also considering subjective sleep measures, which are a common, cost efficient and quick way to collect sleep data.

Although available evidence broadly suggests that subjective and objective measures of sleep are at least to some degree statistically related and in agreement with each other, the exact size of these effects across different sleep parameters is less clear. Campanini and colleagues evaluated the relatedness and agreement between sleep diary and actigraphy estimates of TST, SOL, and S.E among 163 Brazilian school teachers and found a moderate degree of relationship and agreement for TST (ICC = 0.70; r = 0.60) and poor agreement for SOL (ICC = 0.49, R = 0.38) and S.E (0.16; r = 0.22). However, the authors also acknowledged the wide inconsistency in the literature and cited examples of highly variable findings to their own (e.g., Arora et al., 2013; McCrae et al., 2005; Kolling et al., 2015). In another study, Lauderdale, and colleagues (2008) reported a correlation of 0.45 between actigraphy and sleep diary estimated TST among 669 young adults between the ages of 18-30 but found poor agreement, with subjective estimations systematically overestimating TST by 48 minutes on average. The literature at present is missing a central reference point which has measured relatedness and agreement across a wide range of the most measured sleep parameters among healthy adults.

Cespedes and colleagues (2016) compared actigraphy and self-report sleep measures among 2086 Hispanic/Latinos. In this study, participants wore actigraphy wrist watched for 5 nights and filled out sleep diaries each night. The authors reported a moderate correlation between actigraphy and self-report for sleep duration, and

specifically that actigraphy measured sleep duration was one hour shorter than self-reported sleep duration.

Girschik and colleages (2012) investigated agreement between actigraphy and self-reported sleep duration and found poor agreement between subjectively and objectively measured sleep. The results from this study are not directly comparable, however, as the authors opted to condense their self-reported sleep duration and actigraphy sleep duration measures into categorical variables and measure agreement using Cohen's kappa (with agreement ranging from -0.19 to 0.14).

Thurman and colleagues (2018) investigated statistical agreement among numerous sleep measures including sleep duration, number of awakenings, and wake after sleep onset. The authors reported two key trends. The first is that agreement was strong for sleep parameters reliant on sleep onset and offset (i.e., sleep duration) but week for sleep variables reliant on sensitive detection of wakefulness (i.e., number of awakenings, wake after sleep onset). The second is that statistical agreement dwindled over time particularly among less compliant participants. Concerning statistical agreement, bland altman plots revealed that 41% of differences were within 30 minutes of each other and 78% of differences were within 90 minutes of each other, which was argues to be a reasonable degree of agreement. For the number of awakenings, 46% percent of differences were within 2 awakenings of each other, and 92% were between 6 awakenings of one another. This constutes a low level of agreement. The Intraclass Correlation Coefficient was not calculated in this study, however. The authors also measured statistical relatedness, in which it was found that sleep duration was 10.3 minutes shorter than by sleep logs ($r$ = 0.62, p < 0.00001). For number of awakenings, however, subjective and objective measures were not statistically related ($r$ = 0.03, p = 0.5). Overall, whilst there are several high-quality studies on the topic, the literature at present is missing a central reference point which has measured relatedness and agreement across a wide range of the most measured sleep parameters among healthy adults.

There exists consistent literature which suggests that relatedness and agreement between subjective and actigraphy is poor among various clinical groups. Examples include HIV (De Francesco, 2021 showed poorer agreement among HIV positive

individuals compared to HIV negative controls following a similar lifestyle), obesity (O'Brien et al., 2016), insomnia (Bianchi et al., 2013), cancer (Lubos et al., 2022), and PTSD (Werner et al., 2016). Here we narrowed our investigation to focus on sleep disorders since it is intuitive that relatedness and agreement are especially likely to be affected among conditions within which sleep is directly affected.

***The current investigation***

Here I seek to help guide the use of and interpretation of data resulting from the use of subjective sleep measures by carrying out a rapid review designed to return a snapshot of the sleep literature focussing on studies which measured sleep parameters both objectively and subjectively and to calculate the effect sizes for relatedness and agreement for all variable and measurement combinations for which we had a powerful enough raw dataset for. Our approach was intended to be as flexible and as comprehensive as possible and the dataset which resulted from this process led to us calculating effect sizes using a mega-analytic mixed-effects paradigm (Boedhoe et al., 2019) within which relatedness and agreement between objective and subjective measures of Total Sleep Time (TST), Number of Awakenings (NA), Sleep Onset Latency (SOL), Wake After Sleep Onset (WASO), and Sleep Efficiency (S.E) were calculated. Ascertaining the relatedness and agreement of subjective and more objective measurements of specific sleep variables is important to several fields of research. For example, fully understanding the results observed in Chapters 3 and 4 of this thesis can only be achieved by ascertaining the relatedness and agreement of self-reports of night waking with one of the more objective approaches (actigraphy or PSG).

Mega analysis involves pooling raw data across studies and is far less common than meta-analysis (pooling effect sizes from publications) largely due to the common lack of availability of raw data to pool and the considerable time and resources it takes to collect and standardise raw data sets from many different studies. However, two previous studies have compared mega- and meta-analysis, and both have reported and have found mega-analysis to perform as well as meta-analysis (Steinberg et al., 1997; Boedhoe et al., 2019). More specifically, Steinberg et al. (1997) reported very high agreement between the methods using cancer patient data, and Boedhoe et al (2019) reported favourable performance from mega-analysis over meta-analysis (i.e., better fit

indices along with similar standard errors and confidence intervals) using neuroimaging data. It is not the purpose of the current investigation nor is there enough evidence to argue that mega-analysis is superior to meta-analysis. The available evidence does, however, seem to justify the use of mega-analysis when convenient to do so. This mixed model approach also allowed for us to examine the influence of having a chronic sleep disorder on the relationship between sleep measures by including a simple population binary (*Healthy/Clinical)* as an interaction term, and comparison of the Intraclass Correlation Coefficient (ICC) between sleep measures across clinical and healthy samples allowed us to examine statistical agreement. This approach was intended to be exploratory and observational, and no hypotheses were made.

## Methods

### *Inclusion/exclusion criteria*

The inclusion and exclusion criteria were as follows: Studies had to contain at least two measures (PSG / Actigraphy / Subjective) of at least one of the following variables (Total sleep time, Number of awakenings, Sleep onset latency, Wake after sleep onset, and Sleep efficiency). The study population must either be healthy, meet the diagnostic criteria for a sleep disorder (either through medical diagnosis or questionnaires administered as part of the study), or include both groups. Sleep measures had to be measured across at least one full night. For actigraphy, the study must have used scientific grade wrist actigraphy (i.e., non-commercial, e.g., Fitbit). No manipulations could be conducted which interfere with natural sleep (e.g., the administration of a drug or a forced waking protocol). We did not consider having participants complete an experimental or research-based task as a “manipulation”. Participants had to be at least of an adolescent age, specifically no younger than 13 years old, to exclude parent-report.

## *Literature Search*

| **Web of Science (2006 – 2021)** | **PubMed (2006 – 2021)** | **PsycINFO (2006 – 2021)** |
|---|---|---|
| *polysomnography OR PSG*<br>**AND**<br>*actigraphy OR actimetry OR Actiwatch OR Somnowatch OR Action-W OR Sleepwatch OR WatchPAT 100 OR Motionlogger OR Actillume OR wGT3X-BT OR acceleromet OR accelometer OR actigraph*<br>**AND**<br>*PSQI OR Pittsburgh Sleep Quality Index OR sleep habits OR (sleep AND self-report) OR sleep diary OR sleep-diary OR (sleep AND questionnaire) OR (sleep AND form)*<br>**NOT**<br>*children OR pediatric* | *(((Polysomnography OR PSG) AND (actigraphy OR actimetry OR Actiwatch OR Somnowatch OR Action-W OR Sleepwatch OR WatchPAT 100 OR Motionlogger OR Actillume OR wGT3X-BT OR acceleromet OR accelometer OR actigraph)) AND (PSQI OR Pittsburgh Sleep Quality Index OR sleep habits OR (sleep AND self-report) OR sleep diary OR sleep-diary OR (sleep AND questionnaire) OR (sleep AND form))) NOT (Children OR pediatric)*<br><br>***Filters applied:*** *Journal Article, Humans, English, Adolescent: 13-18 years, Adult: 19+ years, Young Adult: 19-24 years, Adult: 19-44 years, Middle Aged + Aged: 45+ years, Middle Aged: 45-64 years, Aged: 65+ years, 80 and over: 80+ years, from 2006/1/1 - 2021/4/1.* | *(((Polysomnography or PSG) and (actigraphy or actimetry or Actiwatch or Somnowatch or Action-W or Sleepwatch or WatchPAT 100 or Motionlogger or Actillume or wGT3X-BT or acceleromet or accelometer or actigraph) and (PSQI or Pittsburgh Sleep Quality Index or sleep habits or (sleep and self-report) or sleep diary or sleep-diary or (sleep and questionnaire) or (sleep and form))) not (Children or pediatric)).af.*<br><br>***Filters applied:*** *1 and "Journal" [Publication Type]* |
| Search terms for Web of Science which returned 307 results. | Search terms for PubMed which returned 559 results. | Search terms for PsycINFO which returned 615 results. |

*Figure 24. Search terms for each of the three databases used.*

Figure 24 above shows the exact searches used across three databases: Web of Science, PubMed, and PsychINFO. We were not convinced that there was an adequate number of publications in the literature which specifically investigated the concordance between distinct types of sleep measures. Instead, we took a flexible approach which sought to return a manageable number of studies for which we anticipated in advance that we would have to request data from the authors. Searching for studies which contained a measure of PSG, actigraphy, AND subjective sleep measure facilitated this flexible approach since it brought down the number of results to a manageable level. Although we were most interested in the agreement between subjective and objective sleep measures, we intended to conduct analyses for whichever comparisons (e.g., PSG vs Actigraphy/ PSG vs Subjective/ Actigraphy vs Subjective) returned adequate amounts of data for meta- or mega-analysis. If a returned study only had two measures of sleep (e.g., actigraphy and subjective) this would not in itself bar the study from inclusion in the final analyses.

***Screening***

Search results were imported in Endnote and title and abstract screening was conducted by one member of the research team. The same researcher then screened full texts. Two other members of the research team were available to review texts which were hard to categorise, but they were not needed.

***Data requests to authors***

The first and last authors for each of the studies who met the search criteria were contacted by email and asked to provide all data relevant to the current investigation at the level of variable means by night (i.e., not epoch by epoch) for each participant. If the contact information was not available for either the first or last author, any other authors of the study whose contact information was available were contacted. The authors were also provided with a suggested data template to help them meet our request more efficiently but were also advised that we were happy to accept data in whatever way was most convenient to them. As detailed in Figure 2 below, a total of 330 author pairs were contacted in this way.

### *Heterogeneity and Measurements Obtained Across Multiple Nights*

Taking a mixed-effects approach allowed for the potential influence of methodological heterogeneity across studies and any clustering effects of taking several measures across different nights from the same participants to be controlled for by including them in the models as random effects.

### *Variable coding*

For each model, the dependent variable was mean-centered, and the Population variable was dummy coded with 'Healthy' as the intercept (Healthy =0; Clinical =1). See Figure 1 (below) for description of the remainder of the variables used which were all continuous.

### *Effect size calculation*

Each multilevel model had the more objective sleep measure as the dependent variables. Then the respective subjectively estimated sleep measure, a grouping variable (Population: Healthy/Clinical) and the interaction term of the two were added as fixed effects. Within the same step, Study was added into the model as a random effect (e.g., *TST (Act) ~ TST (Sub)*Pop (Clinical)*). Beta coefficients (β) and Cohen's d were calculated for the effect of the subjective sleep measure for each of the six mega-analyses (summarised in Table 16 below). These statistics allowed for comparison of how related each subjective sleep measure was to its more objective counterpart. Cohen's d was calculated using the following equation: *d = estimate of fixed effect/ sqrt(sum of the variances of random effects),* as outlined in Brysbaert and Stevens (2017).

### *ICC Selection*

We used McGraw and Wong's (1996) classification system for the Intraclass Correlation Coefficient, and Koo and Li's 2016) guidelines for choosing which version of the ICC is suitable for the current study. As such, the version of the ICC used in the analyses below is the one-way random effect ICC for absolute agreement with multiple

raters. The ICC was calculated using the ICC function of the R package 'psych' and reporting data from the single-random raters' version of the ICC.

### *Differences in relatedness and agreement between healthy and clinical samples*

One aim of this study was to assess whether the strength of the relationship between objective and subjective sleep measures is poorer in populations with sleep disorders. To inform this question, we tested the interaction between Population *(Healthy/Clinical)* and subjective sleep measure in each multilevel model. Another aim of the current investigation was to evaluate the degree of agreement (whether each sleep measure was the same as opposed to just being related) between each pair of measurements and to compare across healthy and clinical populations. We used Bland-Altman plots to visualise patterns in agreement and bias among comparison pairs (e.g., TST (Act) ~ TST (Sub) [Healthy]) and we calculated and compared the Intraclass Correlation Coefficient (ICC) between sleep measures for both healthy and clinical samples (see Table 17 below for a summary of these analyses). For interpreting ICCs, we used Koo and Li's (2016) guidelines. As such, an ICC of less than 0.5 is to be regarded as poor, 0.50 - 0.75 is moderate, 0.75 - 0.90 is good, and above 90 is excellent.

### *Search Results*

The above searches returned 307 results from Web of Science, 559 from PubMed, and 615 from PsychINFO (n= 1481 in total). Detailed in Table 2 below, 330 remained after the removal of duplicates (n = 362), discarding of reviews and meta-analyses (n=91), and title and abstract screening (n = 711). All first and last authors were contacted, and raw data was requested any relevant variables (Total sleep time (TST), Number of nocturnal awakenings (NA), Sleep onset latency (SOL), Wake after sleep onset (WASO), Sleep efficiency (S.E)) for which the authors had measured using any two of the following (PSG/ Actigraphy, Subjective). Authors from 40 of these publications provided data, 13 of which were usable in at least one analysis. The beginning of each subsection of the results details which studies were included in each analysis. These studies allowed for 6 mega-analyses to be conducted (TST (Act) ~ TST (Sub), TST (PSG) ~ TST (Sub), NA (Act) ~ NA (Sub), SOL (Act) ~ SOL (Sub), WASO (Act) ~ WASO (Sub), S.E (Act) ~ S.E (Sub)). See Figure 2 below for a visual summary of the screening process.

### *Study Populations*

Table 3 below summaries the participants from each of the 13 included studies in the analyses below. These studies had a total of 960 participants (Age: M = 36.81, SD = 12.45; Female = 57.81%), many of whom provided data on several nights and across several of the analyses below. Of these 960 participants, 438 were healthy and (Age: M = 32.54; SD = 6.18; Female = 52.2%), and 522 reported symptoms of a sleep disorder (Age: M = 45.34; SD  13.17; Female = 62.45%). Of those with sleep disorders, 387 (74.12%) had Insomnia and the remaining participants (n = 135 from Peter-Derex et al., 2020) had one of the following: Narcolepsy Type 1; Narcolepsy Type 2; Idiopathic Hypersomnia (a total of n=76 for both types of Narcolepsies and Idiopathic Hypersomnia with the exact breakdown in numbers unknown); Insufficient Sleep Syndrome (n = 24) or Obstructive Sleep Apnoea Syndrome (n = 36).

### *Experimental settings*

All the PSG studies identified by our review were carried out in sleep labs (n = 4, see Table 3 below). The remainder were actigraphy studies in which participants went about their daily routines and slept at home. All the PSG studies took place over one night, except Roberts et al., (2020), which took place over 4 nights. Two of the PSG studies had adaptation nights (Kobayashi et al., 2012; and Roberts et al., 2020), and the other two did not (Hermans et al., 2020; and Peter-Derex et al., 2020). All the actigraphy studies took place over many nights, ranging from 7 to 112.

### *Sleep Measurements*

Six types of wrist actigraphy devices were used across studies and all were scientific grade (See Table 3 below). These devices were the Actiwatch 2, the Actiwatch Spectrum (n=2), the Actigraph Motionwatch 8 (n=2), the wGT3X-BT (n=2), the Micro-Mini Motionlogger Actigraph (n=2), and the Readiband Actigraph SVB2 (n=1).

For subjectively measured sleep, all the studies included in the analyses below used a diary approach. Figure 25 below describes how each of our variables of interest were defined and calculated:

| Measure | Definition | Calculation |
|---|---|---|
| Total Sleep Time | Total amount of sleep (*mins*) from sleep onset to sleep offset (self-reported wake time). | Sleep Offset – Sleep Onset |
| Number of Awakenings | Number of awakenings, no matter how brief, during the nocturnal sleep phase (*n*). | Subjective estimate |
| Sleep Onset Latency | Time taken between getting into bed and sleep onset (*mins*) | Subjective estimate |
| Wake After Sleep Onset | Time spent awake between sleep onset and sleep offset (*mins*) | Subjective estimate |
| Sleep Efficiency | Amount of time spend sleeping since going to bed and trying to sleep and sleep offset (*%*). | Total Sleep Time/Time in bed*100 |
| Sleep Onset | Time the nocturnal sleep phase begins. | Bedtime + Sleep Onset Latency |
| Sleep Offset | Morning wake up time. | Subjective estimate |
| Time in Bed | Time the participant got in to bed and began attempting to sleep until Sleep Offset. | Sleep Offset – Bedtime. |
| Bedtime | Time the participant got in to bed and began attempting to sleep | Subjective estimate |

Total Sleep Time, Number of Awakenings, Sleep Onset Latency, Wake After Sleep Onset, and Sleep Efficiency were all included in analyses, the remainder were used to calculate them.

*Figure 25. Descriptions of key terms needed to interpret the Results section.*

Figure 26 shows the flow of identification of studies:

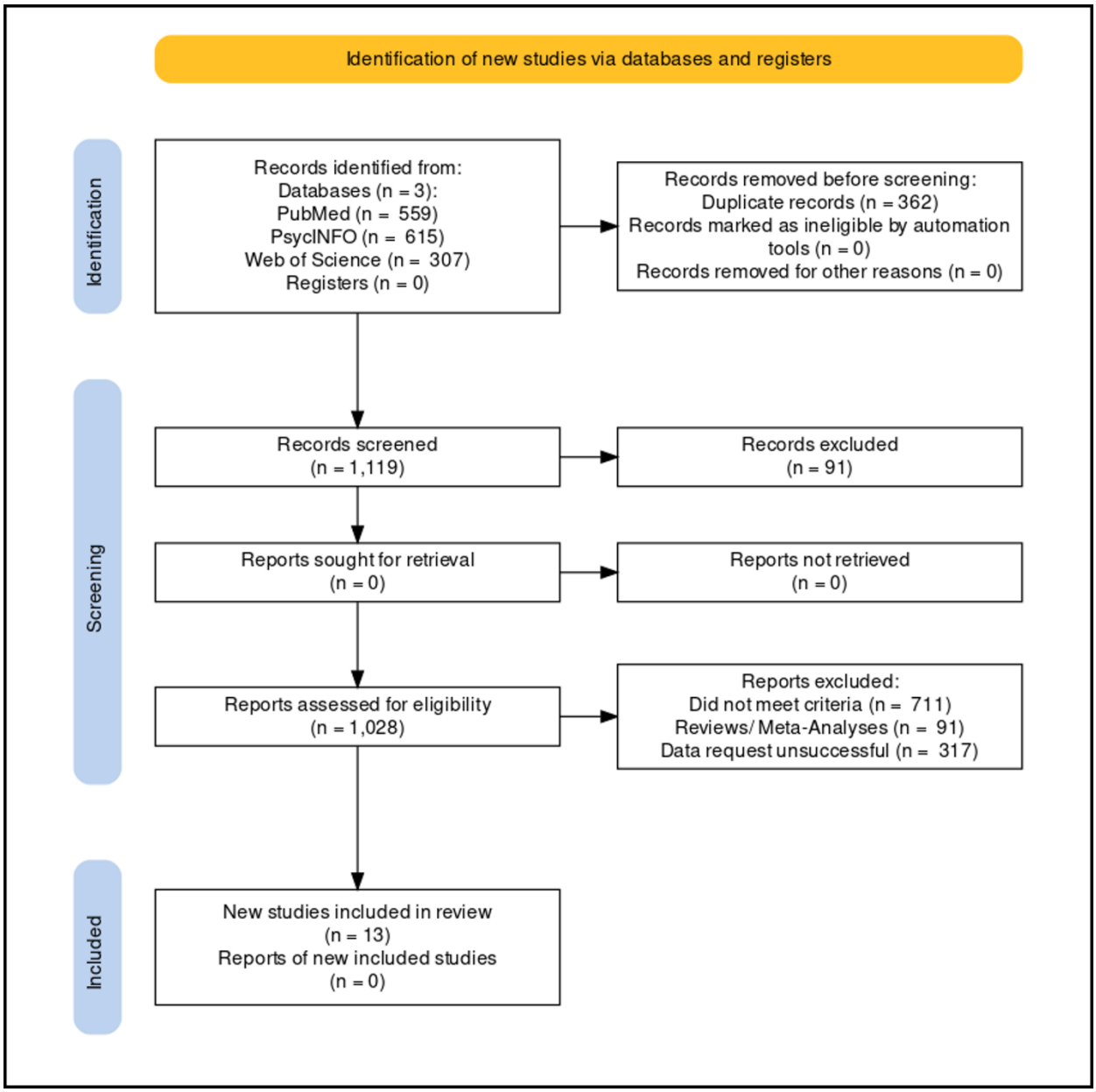



*Figure 26. PRISMA flow chart for the identification of studies.*

Table 63 describes the final studies included:

*Table 64. Key attributes of included studies (N =13).*

| Authors | Sample | Age M (SD) | Gender N (%) Female | Environment | Measures |
|---|---|---|---|---|---|
| Gieselmann & Pietrowsky., 2019 | N= 71, Insomnia | 40.6 (12.45) | 38 (52.8) | 7 nights of actigraphy and sleep diary as part of ordinary routine (Actigraph Motionwatch 8) | TST; NA; SOL; WASO; S.E |
| Goelema et al., 2019 | N=50, Healthy Adults | 51.8 (6.6) | 26 (54) | 14 nights of actigraphy and sleep diary as part of ordinary routine (Actiwatch Spectrum) | TST; SOL; WASO; |
| Gokce et al., 2020 | N=85, Healthy Adults | 22.7 (3.0) | 64 (86.7) | 7 nights of actigraphy and sleep diary as part of ordinary routine (wGT3X-BT) | TST; NA; SOL; WASO; S.E |
| Hermans et al., 2020 | N=234 (95 Healthy Adults and 139 with Insomnia) | Healthy (M = 36.0; SD =13.7); Clinical (M = 46.7; SD = 13.7) | Healthy (N = 57; 62.0%); Clinical (N = 95; 68.3%) | 1 night of in-lab PSG and sleep diary. | TST; |
| Hoang et al., 2020 | N=141, Insomnia | 53.1 (11.3) | 85 (60.0) | 3 nights of actigraphy and sleep diary as part of ordinary routine (wGT3X-BT) | TST; |
| Janku et al., 2020 | N=36, Insomnia | 46.7 (13.9) | 22 (61.1) | 42 nights of actigraphy and sleep diary as part of ordinary routine (Motionwatch 8) | TST; NA; SOL; WASO; S.E |
| Kobayashi et al., 2012 | N=23, Healthy Adults | 22.6 (5.0) | 8 (34.8) | 1 night of in-lab PSG with actigraphy and sleep diary (Micro Mini-Motionlogger) . | TST; SOL; WASO; |
| Lutz et al., 2018 | N=55, Healthy Adults | 25.0 (NA; Range = 20-37) | NA(~ 50) | 1 night of actigraphy and sleep diary as part of ordinary routine (Actiwatch 2) | TST; SOL; WASO; |
| Medina et al., 2015 | N=33, Healthy Adolescents | 16.5 (0.89, Range = 15-18) | 19 (57.6) | 9 nights of actigraphic monitoring (Micro-Mini Motionlogger Actigraph) along with sleep diaries. | TST; SOL; WASO; S.E |
| Peter-Derex et al., 2020 | N=135, Insomnia | 39.87 (14.48) | 86 (63.7) | I night of in-lab PSG and sleep diary. | TST; |
| Roberts et al., 2020 | N=8, Healthy Adults | 40.75 (4.84) | 5 (62.5) | 4 nights of in-lab PSG coupled with Actigraphy (Actiwatch Spectrum) and sleep diary | TST; SOL; |
| Slightham et al., 2017 | N=60, Healthy Adults | 54.5(9.2) | 5 (8.83) | I night of actigraphy as part of normal routine along with sleep diaries. | TST; NA; S.E |
| Thurman et al., 2018 | N=29, Healthy Adults | 23.0 (NA; Range = 18-35) | 17 (58.6) | 112 nights of actigraphy and sleep diary as part of ordinary routine (Readiband Actigraph SBV2) | TST; SOL; WASO; |

## Results

For each of the 6 comparisons that our systematic review allowed for the analysis of, multilevel models were built to investigate the strength of relationship between comparison pairs (e.g., TST (Act) ~ TST (Sub)). The influence of Population (Healthy/Clinical) on the strength of relationship between the pairs was investigated by adding Population as an interaction term into each multilevel model. We calculated the Intraclass Correlation Coefficient and created Bland-Altman plots to assess agreement between these pairs. Separate datasets were created for healthy and clinical groups when calculating ICCs and Bland-Altman plots and the results compared.

### *M1:Total sleep time (TST (Act) ~ TST (Sub))*

Table 64 below lists the studies which measured TST using both actigraphy and self-report:

*Table 65. Study characteristics for those included in M1.*

| Study | Year | Nights | *n* | Population |
|---|---|---|---|---|
| Lutz et al., | 2018 | 1 | 55 | Healthy |
| Slightham et al., | 2017 | 1 | 60 | Healthy |
| Kobayashi et al., | 2012 | 1 | 23 | Healthy |
| Roberts et al., | 2020 | 4 | 8 | Healthy |
| Gieselmann et al., | 2019 | 7 | 71 | Clinical |
| Gokce et al., | 2020 | 7 | 85 | Healthy |
| Medina et al., | 2015 | 9 | 33 | Healthy |
| Goelema et al., | 2019 | 14 | 50 | Healthy |
| Hoang et al., | 2020 | 3 | 141 | Clinical |
| Thurman et al., | 2018 | 112 | 29 | Healthy |
| Janku et al., | 2020 | 42 | 36 | Clinical |

**Relatedness.** Table 65 (below) shows the results of the multilevel model created for M1. The data shows a strong statistically significant relationship between subjectively estimated total sleep time and its corresponding actigraphy measure (Cohen's d = 0.89). The interaction between population and subjective sleep duration was also statistically significant, such that the relationship between subjectively estimated sleep duration and actigraphy-measured sleep duration is stronger when the population is healthy as opposed to clinical (Cohen's d = -0.58). Figure 27, below, visualises the interaction:

*Table 66. Results for a mixed-effects model with TST (Act) as dependent variable and Study (random effect), ID (random effect), and TST (Sub; fixed effect) as predictors*

| Predictor | *CI* | β | *SE_*β | *p* |
|---|---|---|---|---|
| TST (Sub) | 70.68 – 75.27 | 0.72 | 0.13 | **<.001*** ** |
| Pop (Clinical) | -75.82 – 18.88 | -0.14 | 0.05 | .239 |
| TST*Pop | -52.12 - -43.05 | -0.26 | 0.12 | **<.001*** ** |

*Note.* TST is mean-centred and in minutes. Population is a variable composed of two groups (*Healthy/Clinical)*. Marginal $R^2$ = 0.38; Conditional $R^2$ = 0.64; ICC = 0.42.

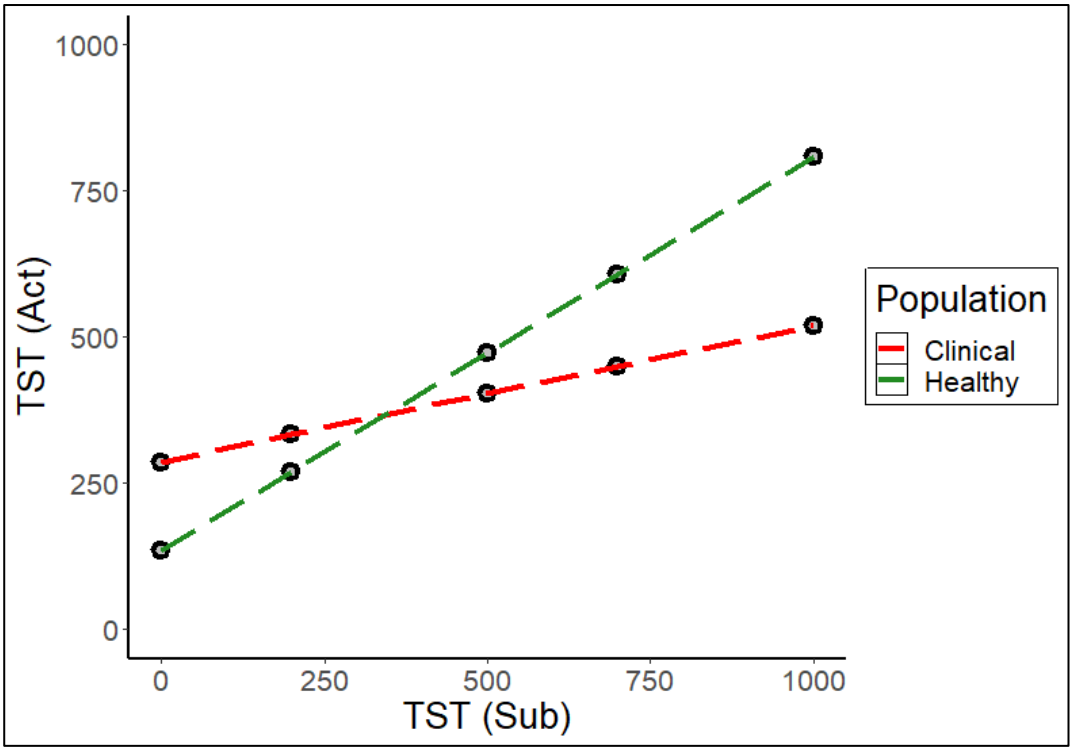
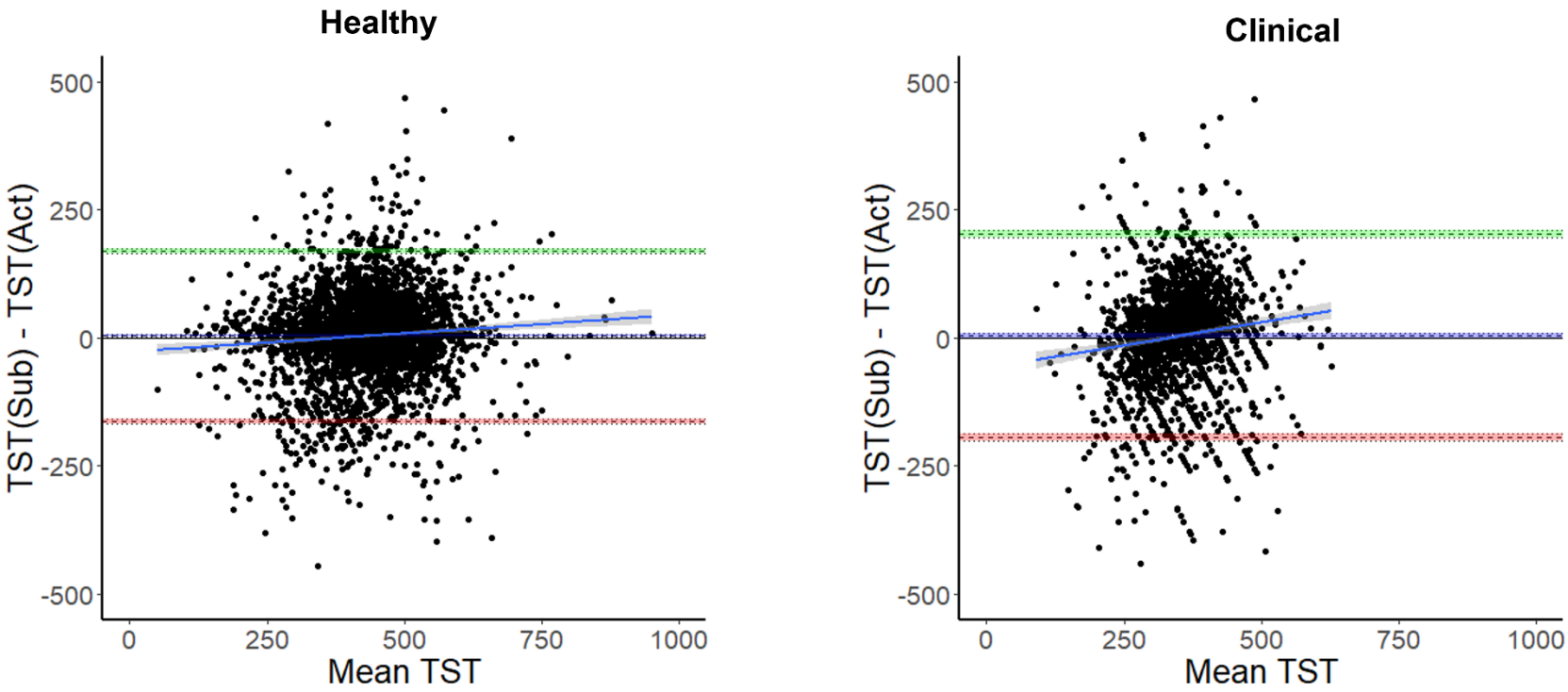



*Figure 27. Interaction plot for M1 (β = -0.26, SE_β = 0.12, p <.001***). Axes in mins.*




Healthy (mins; ULoA = 165.25, LLoA = -166.59, outliers = 5.91%). Clinical (mins; ULoA = 202.83, LLoA = -194.53, outliers = 5.97%).
ULoA = Upper Level of Agreement (Green). LLoA = Lower Level of Agreement (red).

*Figure 28. Bland-Altman plots visualising concordance between TST(Act) and TST(Sub) in amongst healthy (left) and clinical (right) participants.*

**Agreement.** Figure 28 above shows Bland-Altman plots for healthy and clinical subgroups. For healthy participants, subjectively estimated sleep duration overestimates sleep duration on average 3.81 minutes compared to 4.15 minutes for the clinical subgroup (i.e., small levels of constant bias). One sample t-tests showed that the mean difference was significantly above 0 among healthy participants (*M = 3.8, SD = 84.6, t(3992) = 2.8, p<.01, d = 0.05*) but not significantly different among clinical participants (*M= 4.1, SD = 101.3, t(2210) = 1.9, p = .054, 0.04*). Both groups show a small proportional bias trend such that mean differences increase as sleep duration increases. This means that subjective measures overestimate sleep duration more so at longer sleep durations. This trend is steeper in the clinical subgroup. These positive trends were statistically significant for both groups (Healthy: $R^2 = .005, F(1, 3991) = 18.35, p < .001$; Clinical: $R^2$ = .02, F(1, 2209) = 38.6, p < .001). Finally, both groups had a similar proportion of outliers (Healthy = 5.91%; Clinical = 5.97%), close to Bland and Altman's (1983) recommendation of 5% being an acceptable amount.

ICCS were calculated to assess the level of agreement statistically, with both groups showing statistically significant levels of agreement. For the Healthy group, the level of agreement was moderate according to Koo and Li's (2016) guidelines for interpreting the ICC (ICC (H) = 0.67, p <.001) and poor for the clinical group (ICC (C) = 0.31, p<.001). Overall, subjectively and objectively estimated (actigraphy) sleep duration is closely related and show good levels of agreement in healthy populations and are less related and poorly concordant in clinical populations.

### *M2: Total sleep time (TST (PSG) ~ TST (Sub))*

Table 66 below lists the studies which measured TST using both PSG and self-report:

*Table 67. Study characteristics of those included in M2.*

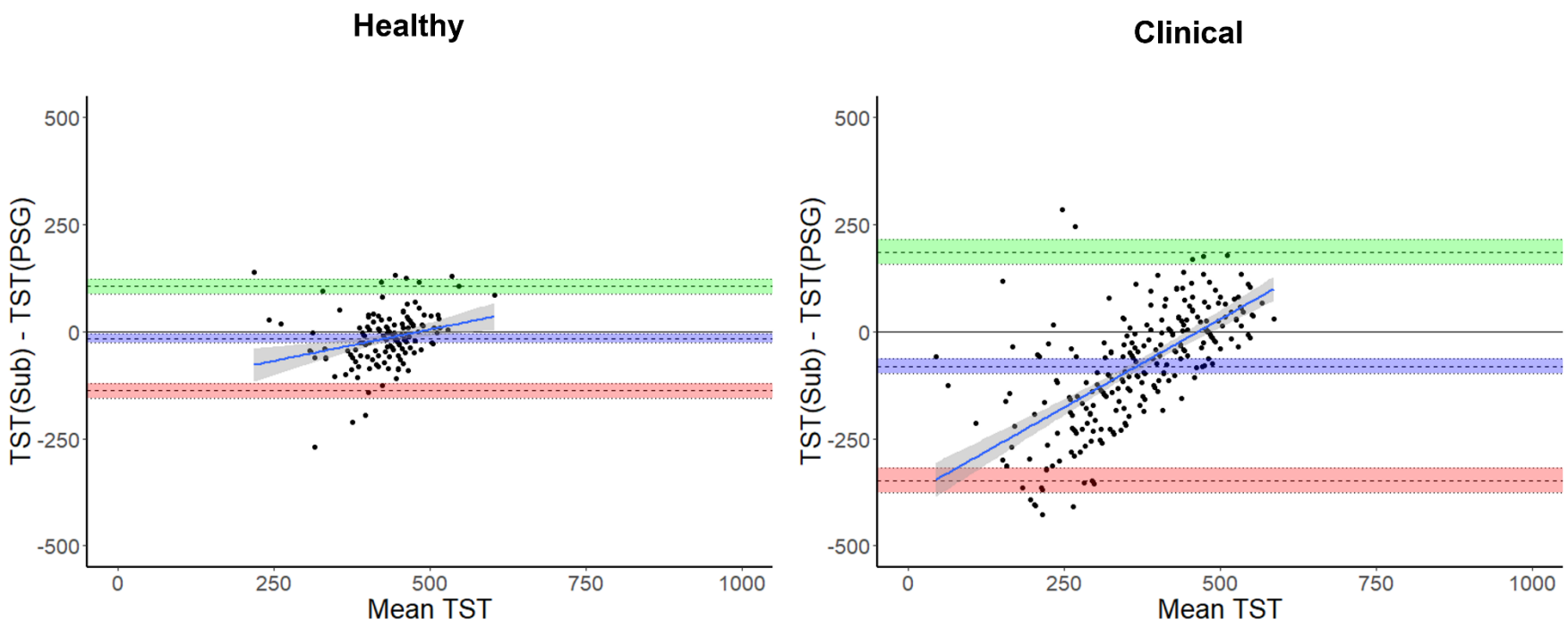
| Study | Year | Nights | *n* | Population |
|---|---|---|---|---|
| Kobayashi et al., | 2012 | 1 | 23 | Healthy |
| Roberts et al., | 2020 | 4 | 8 | Healthy |
| Hermans et al., | 2020 | 1 | 234 | Both |
| Peter-Derex et al., | 2020 | 1 | 135 | Clinical |

*Table 68. Results for a mixed-effects model with TST (PSG) as dependent variable and Study (random effect), ID (random effect), and TST (Sub; fixed effect) as predictors.*

| Predictor | *CI* | β | *SE_*β | *p* |
|---|---|---|---|---|
| TST (Sub) | 26.93 – 63.26 | 0.61 | 0.13 | **<.001*** ** |
| Pop (Clinical) | -11.57 – 19.05 | 0.02 | 0.05 | .239 |
| TST*Pop | -26.81 – 12.08 | -0.09 | 0.12 | .457 |

*Note.* TST is mean-centred and in minutes. Population is a variable composed of two groups (*Healthy/Clinical)*. Marginal $R^2$ = 0.66; Conditional $R^2$ = NA.

***Relatedness*.** The data from Table 67 above show a very strong statistically significant relationship between subjective and PSG measures of sleep duration (Cohen's d = 0.72) and no significant main effect or interaction for Population. This is unexpected given the significant interaction found above for subjective sleep duration and Population in predicting sleep duration measured by actigraphy.




Healthy (mins; ULoA = 105.48, LLoA = -137.22, outliers = 7.64%). Clinical (mins; ULoA = 186.16, LLoA = -347.53, outliers = 5.65%). ULoA = Upper Level of Agreement (Green). LLoA = Lower Level of Agreement (red).

*Figure 29. Bland-Altman plots visualising concordance between TST(PSG) and TST(Sub) in amongst healthy (left) and clinical (right) participants.*

***Agreement.*** The first observation which stands out from Figure 29 above is that there is greater bias and wider limits of agreement for the agreement between TST(PSG) ~ TST(Sub) than compared to TST(Act) ~ TST(Sub). This is unexpected and may be due to a smaller sample size and fewer studies being included in the TST(PSG) analyses. Subjective estimates significantly underestimate sleep duration by 15.87 minutes on average among healthy participants (*M = -15.9, SD = 61.9, t(143) = -3.07, p<.01*) and 80.69 minutes among clinical participants (*M = -80.7, SD = 136.1, t(247) = -9.33, p<.001*). Constant bias was much higher in the clinical group (Healthy: *d* = -0.25; Clinical: *d* = -0.59). Both plots show positive trends as the magnitude of the means increases (i.e., proportional bias) such that self-reports increasingly underestimate sleep duration as mean sleep duration increases. These positive trends were statistically significant for both groups (Healthy: $R^2 = .006, F(1, 142) = 11.03, p < .01$; Clinical: $R^2 = .43, F(1, 246) = 188.3, p < .001$) but with a much larger effect size in the clinical group. Finally, 7.64% of observations fall outside of the limits of agreement lines for the healthy group, and 5.65% for the clinical group.

ICCs suggest that agreement between PSG and subjectively estimated sleep duration is significant and moderate among healthy participants (ICC = 0.53, p<.001) and significant and poor among clinical participants (ICC = 0.34, p<.001). In sum, PSG and subjectively estimated sleep durations are strongly related but are only in moderate agreement among healthy participants and in poor agreement among clinical participants. As well as this, there is constant and proportional bias in both groups which is more severe among participants with sleep disorders.

### *M3: Number of awakenings (NA (Act) ~ NA (Sub))*

Four studies were included in our analyses of the number of awakenings, as shown in Table 68 below. One limitation of these analyses is that close inspection of the published articles and attempted correspondence with the authors failed to establish the exact algorithms used to detect the number of nocturnal awakenings using actigraphy. Histograms (included in Appendix D) revealed that the studies below were likely using the same or very similar algorithms and were therefore grouped together. Two studies (Hoang et al., 2020; and Thurman et al., 2018) were excluded from the following analyses because histograms revealed them to be most likely using different algorithms.

*Table 69. Study characteristics for all studies included in M3.*

| Study | Year | Nights | *n* | Population |
|---|---|---|---|---|
| Slightham et al., | 2017 | 1 | 60 | Healthy |
| Gieselmann et al., | 2019 | 7 | 71 | Clinical |
| Gokce et al., | 2020 | 7 | 85 | Healthy |
| Janku et al., | 2020 | 42 | 36 | Clinical |

*Table 70. Results for a mixed-effects model with NA (Act) as dependent variable and Study (random effect), ID (random effect), and NA (Sub; fixed effect) as predictors.*

| Predictor | *CI* | β | *SE*_β | *p* |
|---|---|---|---|---|
| NA (Sub) | 1.91 – 3.83 | 0.22 | 0.37 | **<.001*** ** |
| Pop (Clinical) | -48.77 – 28.65 | -0.38 | 0.65 | .610 |
| NA*Pop | -3.22 – -0.76 | -0.12 | 0.04 | **.001** |

*Note.* NA is mean-centred. Population is a variable composed of two groups (*Healthy/Clinical)*. Marginal $R^2$ = 0.05; Conditional $R^2$ = 0.84; ICC = 0.84.

***Relatedness*.** A small significant positive relationship was observed between the number of awakenings estimated subjectively and using actigraphy (Cohen's d = 0.13) which interacted with Population such that this relationship was still positive but weaker among clinical participants (Cohen's d = -0.09). Figure 30 below visualise these relationships:

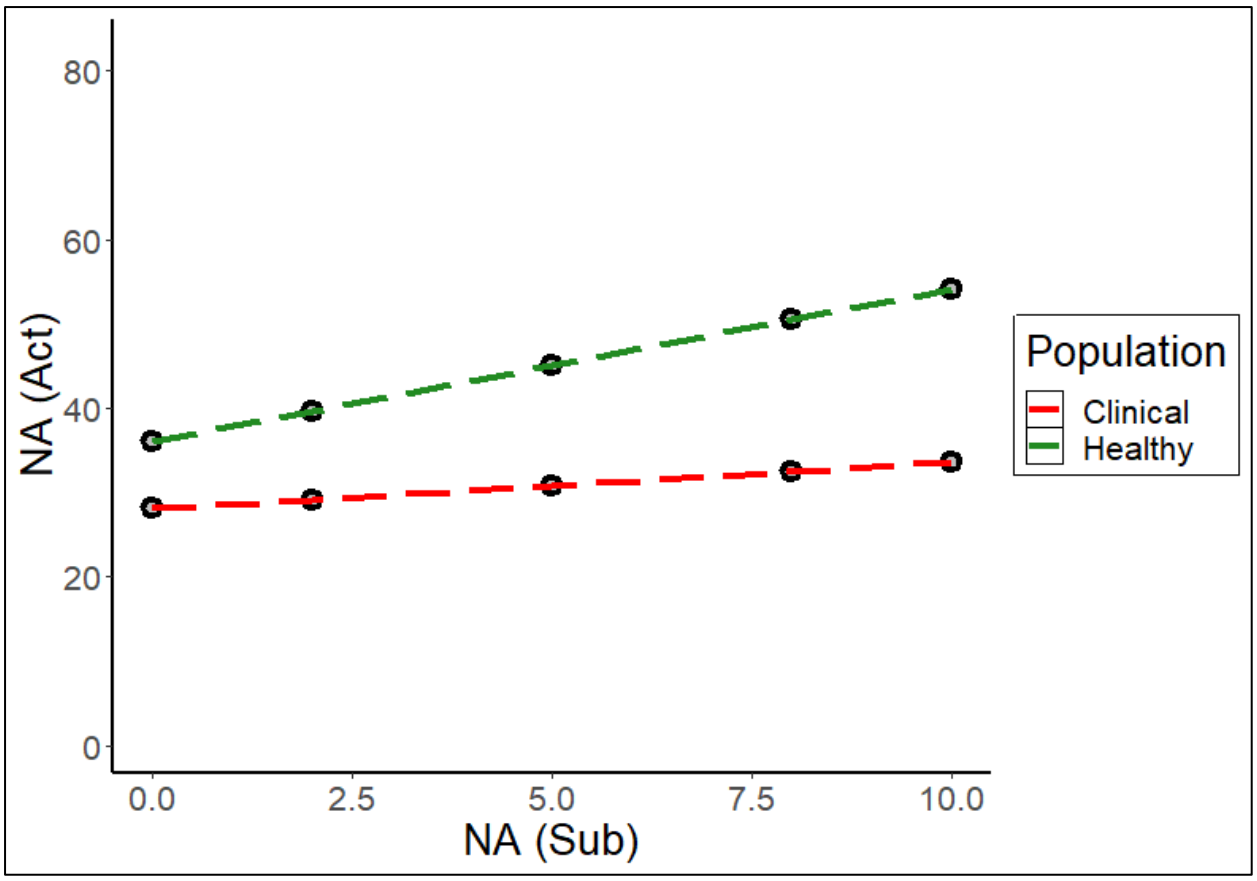



*Figure 30. Interaction plot for M3 (β = -0.12, SE_β = 0.04, p = .001***). X axis is total NA (i.e., not mean-centred) to aid interpretation.*

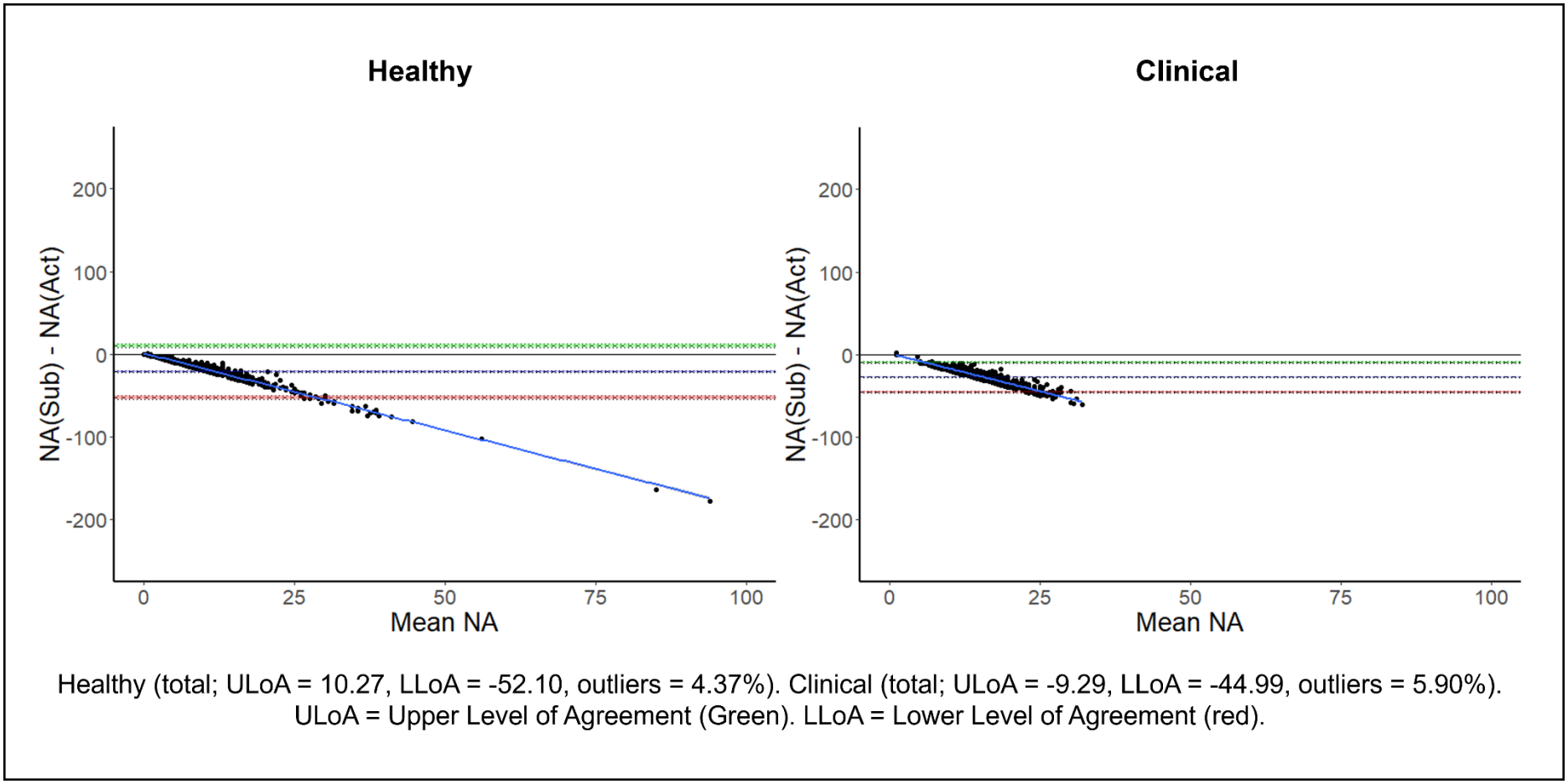



*Figure 31. Bland-Altman plots visualising concordance between NA(Act) and NA(Sub) in amongst healthy (left) and clinical (right) participants.*

***Agreement*.** As figure 31 above shows, there is very poor agreement between measures. Firstly, there is significant constant bias in both groups (Healthy: *M = - 20.9, SD = 15.9, t(594) = -32.1, p<.001*; Clinical: *M = -27.1, SD = 9.1, t(830) = -85.9, p<.001*), such that subjective estimations underestimate actigraphy measures. Bias was larger amongst clinical participants (Healthy: d = -1.31; Clinical; d = -2.98). There was also significant proportional bias in both groups (Healthy: $R^2 = .97, F(1, 593) = 2.06, p < .001$; Clinical: $R^2 = .88, F(1, 829) = 188.3, p < .001$), such that mean differences increased as the number of awakenings increased. Finally, ICCs showed no statistically significant agreement between subjective and actigraphy measures of awakenings in both groups (Healthy: ICC = 0.02, p = .069; Clinical; ICC = 0.00, p = .14).

Overall, there is a small positive relationship between subjective and actigraphy measures of nocturnal awakenings which is weaker among clinical participants. As well as this, there is no statistical agreement between the two amongst healthy or clinical participants.

### *M4: Sleep onset latency (SOL (Act) ~ SOL (Sub))*

Table 70 below shows the 9 studies which were included in the M4 analyses:

*Table 71. Study characteristics for those included in M4.*

| Study | Year | Nights | *n* | Population |
|---|---|---|---|---|
| Lutz et al., | 2018 | 1 | 55 | Healthy |
| Kobayashi et al., | 2012 | 1 | 23 | Healthy |
| Roberts et al., | 2020 | 4 | 8 | Healthy |
| Gieselmann et al., | 2019 | 7 | 71 | Clinical |
| Gokce et al., | 2020 | 7 | 85 | Healthy |
| Medina et al., | 2015 | 9 | 33 | Healthy |
| Goelema et al., | 2019 | 14 | 50 | Healthy |
| Thurman et al., | 2018 | 112 | 29 | Healthy |
| Janku et al., | 2020 | 42 | 36 | Clinical |

*Table 72. Results for a mixed-effects model with SOL (Act) as dependent variable and Study (random effect), ID (random effect), and SOL (Sub; fixed effect) as predictors.*

| Predictor | *CI* | β | *SE_*β | *p* |
|---|---|---|---|---|
| SOL (Sub) | 3.25 – 5.41 | 0.15 | 0.02 | **<.001*** ** |
| Pop (Clinical) | -20.67 – 10.56 | -0.07 | 0.11 | .526 |
| SOL*Pop | -3.69 – -0.43 | -0.05 | 0.02 | **.013*** |

*Note.* SOL is mean-centred and in minutes. Population is a variable composed of two groups (*Healthy/Clinical)*. Marginal $R^2$ = 0.02; Conditional $R^2$ = 0.18; ICC = 0.17.

***Relatedness*.** The data in Table 71 shows a small positive relationship between subjectively estimated SOL and that measured by actigraphy (Cohen's d = 0.16) which interacted with Population such that this relationship slightly weakened among clinical participants (Cohen's d = -0.07). See Figure 33 below for interaction plot:

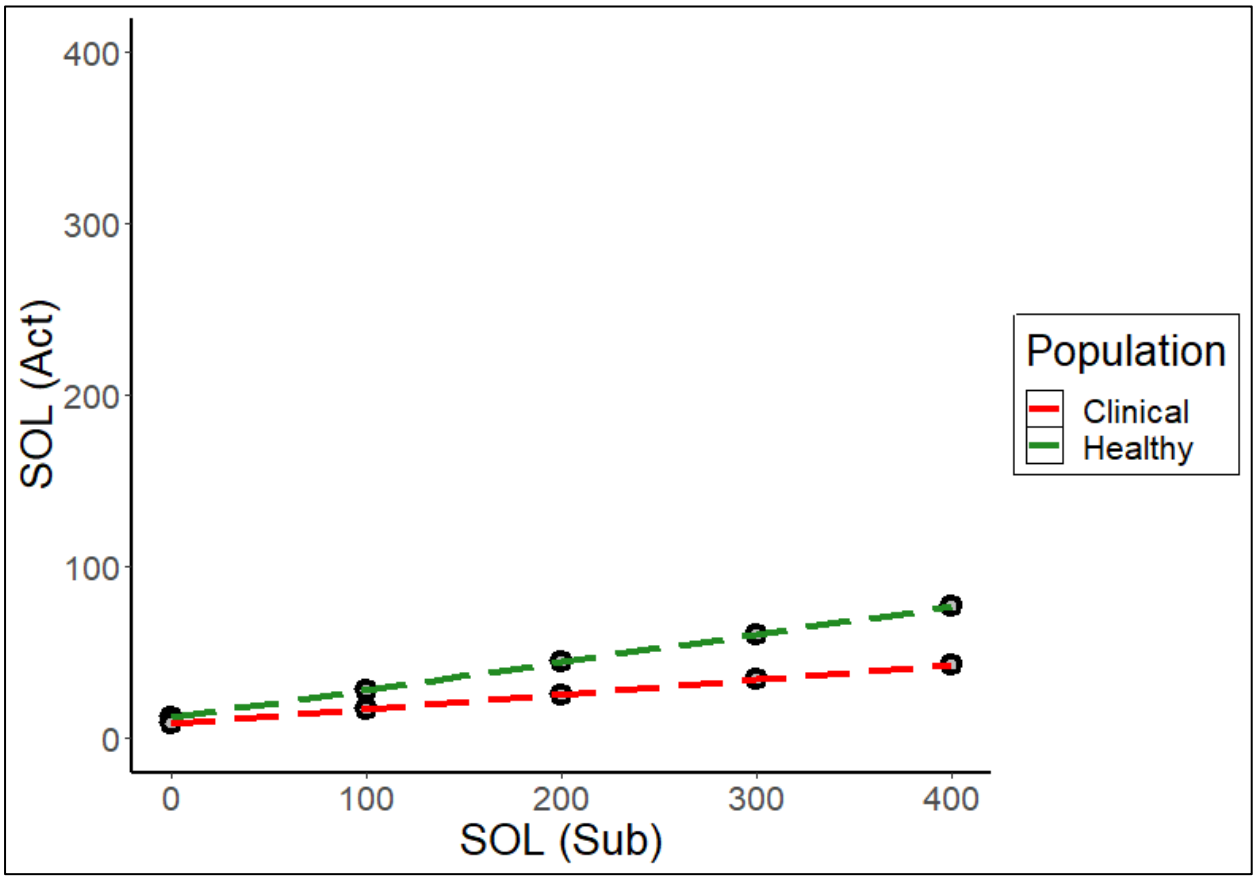
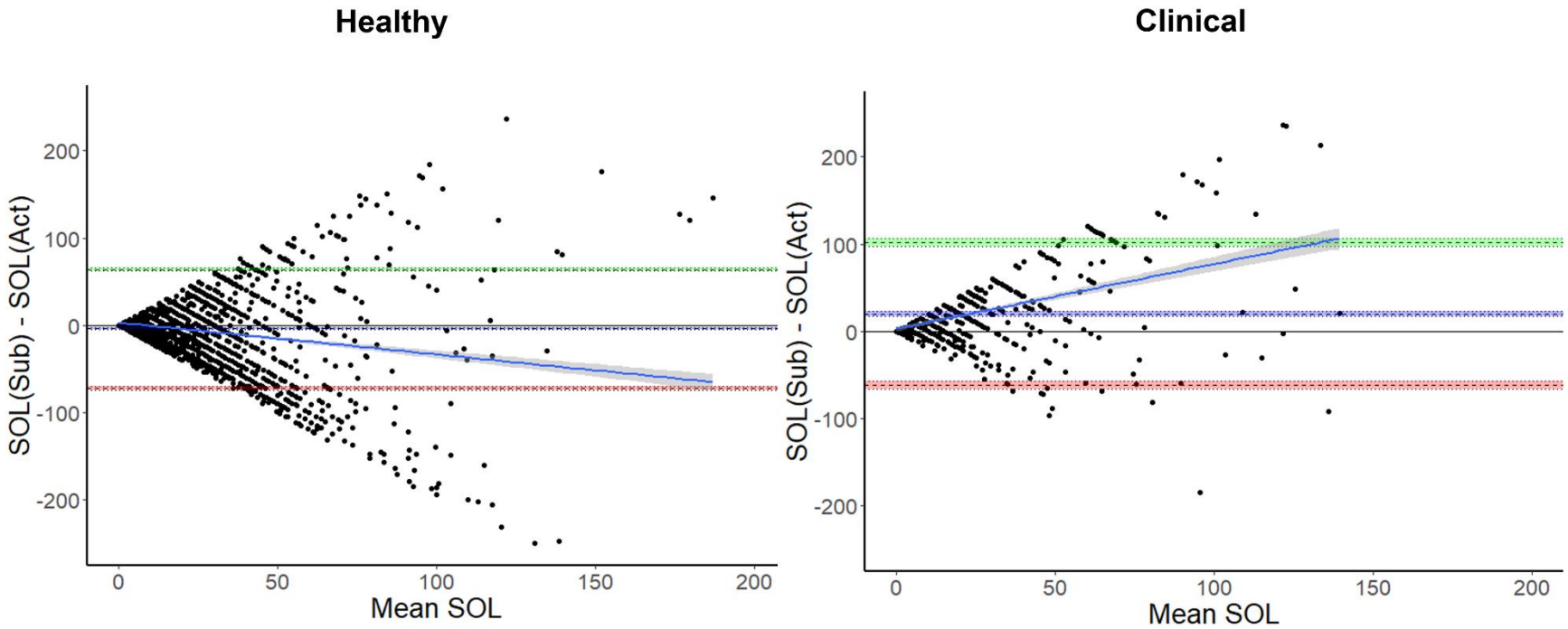



*Figure 32. Interaction plot for M4 (β = -0.05, SE_β = 0.02, p =.013*). Axes in mins.*




Healthy (mins; ULoA = 64.00, LLoA = -72.15, outliers = 5.00%). Clinical (mins; ULoA = 101.69, LLoA = -61.61, outliers = 5.78%). ULoA = Upper Level of Agreement (Green). LLoA = Lower Level of Agreement (red).

*Figure 33. Bland-Altman plots visualising agreement between SOL(Act) and SOL(Sub) in amongst healthy (left) and clinical (right) participants.*

***Agreement.*** Figure 34 above visualises agreement among subjectively estimated SOL and that of actigraphy measures for both Healthy and Clinical participants. Significant constant bias was observed in both groups (Healthy: *M = - 4.07, SD = 34.73, t(3940) = -7.37, p<.001*; Clinical: *M = 20.04, SD = 41.66, t(847) = 14.01, p<.001*) such that healthy participants on average underestimated SOL relative to actigraphy and clinical participants overestimated SOL. Constant bias was much larger among clinical participants (Healthy: $d$ = -0.11; Clinical: $d$ = 0.48). Significant non-systematic bias was also found in both groups (Healthy: $R^2 = .06, F(1, 3939) = 271.80, p < .001$; Clinical: $R^2 = .32, F(1, 846) = 405.20, p < .001$) but more so among clinical participants, such that the mean difference decreased as SOL increased among healthy participants and the mean difference increased as SOL increased among clinical participants. The magnitude of this effect was larger for the clinical group.

ICCs revealed significant but poor statistical agreement among both samples, more so in the clinical group (Healthy: ICC = 0.17, p<.001; Clinical: ICC = 0.11, p<.001). Subjective and actigraphy measured SOL show both low but significant levels of relatedness and agreement with weaker and more biassed measurement among clinical participants.

In sum, subjectively and actigraphy measures of SOL weakly but significantly correlate with one another and this relationship is weaker among clinical participants. The two measures also poorly (but significantly) agree with one another.

### *M5: Wake after sleep onset (WASO (Act) ~ WASO (Sub))*

See table 72 below for included studies:

*Table 73.Study characteristics for all studies included in M5.*

| Study | Year | Nights | *n* | Population |
|---|---|---|---|---|
| Lutz et al., | 2018 | 1 | 55 | Healthy |
| Kobayashi et al., | 2012 | 1 | 23 | Healthy |
| Gieselmann et al., | 2019 | 7 | 71 | Clinical |
| Gokce et al., | 2020 | 7 | 85 | Healthy |
| Medina et al., | 2015 | 9 | 33 | Healthy |
| Goelema et al., | 2019 | 14 | 50 | Healthy |
| Thurman et al., | 2018 | 112 | 29 | Healthy |
| Janku et al., | 2020 | 42 | 36 | Clinical |

*Table 74. Results for a mixed-effects model with WASO (Act) as dependent variable and Study (random effect), ID (random effect), and WASO (Sub; fixed effect) as predictors.*

| Predictor | *CI* | β | *SE_*β | *p* |
|---|---|---|---|---|
| WASO (Sub) | 10.87 – 16.02 | 0.23 | 0.02 | **<.001*** ** |
| Pop (Clinical) | 21.56 – 70.41 | 0.30 | 0.08 | **<.001*** ** |
| WASO*Pop | -14.52 – -6.47 | -0.09 | 0.02 | **<.001*** ** |

*Note.* WASO is mean-centred and in minutes. Population is a variable composed of two groups (*Healthy/Clinical)*. Marginal $R^2$ = 0.13; Conditional $R^2$ = 0.27; ICC = 0.16.

***Relatedness*.** The data from Table 73 suggests that our subjective and objective measures of WASO were significantly positively related (Cohen's d = 0.24). As well as this, being from a clinical group greatly increases the likelihood that actigraphy WASO will be greater (Cohen's d = 0.84). Finally, the two significantly interacted such that the statistical relationship between the subjective and objective measures weakened among clinical participants (Cohen's d = -0.19). This trend is illustrated in Figure 35 below:

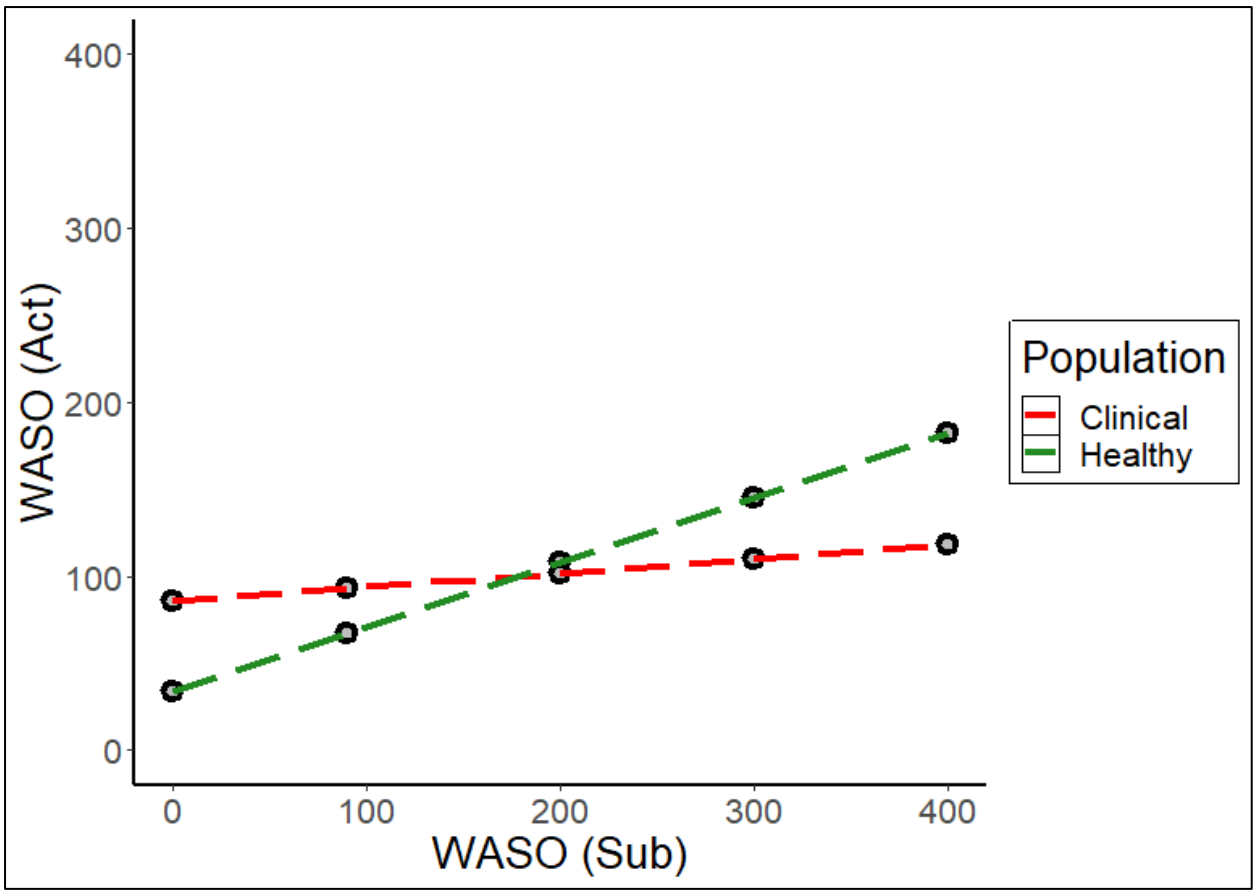



*Figure 34. Interaction plot for M5 (β = -0.09, SE_β = 0.02, p <.001***). Axes in mins.*

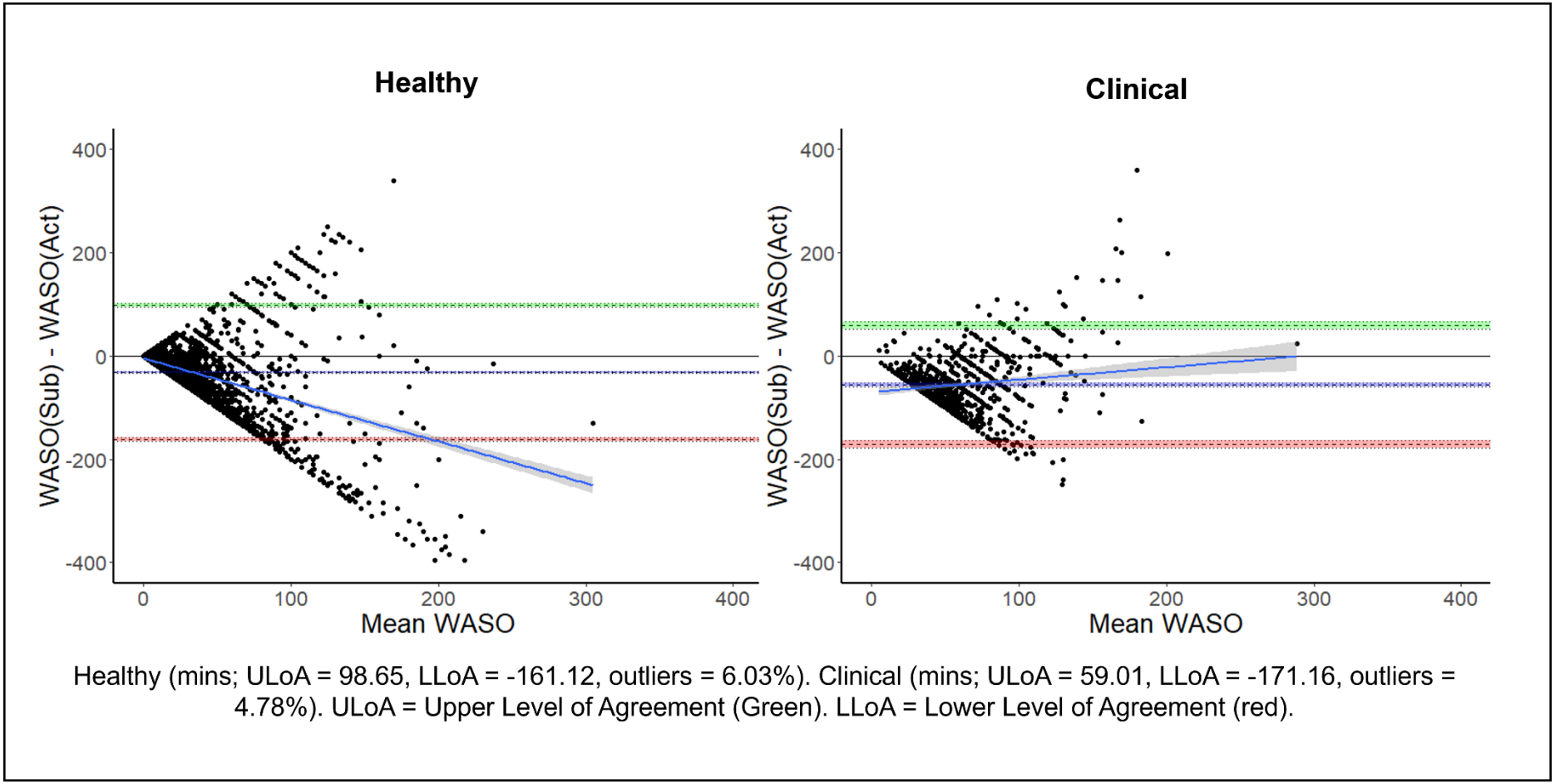



*Figure 35. Bland-Altman plots visualising agreement between WASO(Act) and WASO(Sub) in amongst healthy (left) and clinical (right) participants.*

***Agreement.*** Figure 36 above shows patterns of agreement between subjective and actigraphy WASO for Healthy and Clinical groups. On average, subjective estimates in both groups underestimate WASO measured using actigraphy (Healthy: *M = - 31.23, SD = 66.27, t(3849) = -29.24, p<.001*; Clinical: *M = -56.07, SD = 58.72, t(835) = -27.61, p<.001*). This effect was larger amongst clinical participants (Healthy: *d* = -0.47; Clinical: *d* = -0.97). Significant proportional bias trends going in the opposite directions were observed across groups such that the mean difference decreased as WASO increased among healthy participants and increased among clinical participants as WASO increased (Healthy: $R^2 = .26, F(1, 3848) = 1354, p < .001$; Clinical: $R^2 = .01, F(1, 834) = 7.31, p < .01$). Proportional bias among healthy participants was a much larger effect.

ICCs revealed very poor but significant overall statistical agreement in both groups (Healthy: ICC = 0.03, p<.01; Clinical: ICC = 0.05, p<.01). In sum, subjectively estimated and actigraphy WASO do not appear to be statistically related and show very poor levels of agreement.

### *M6: Sleep efficiency (S.E (Act) ~ S.E (Sub))*

See Table 74 for final M6 studies:

*Table 75. Study characteristics for all studies included in M6.*

| Study | Year | Nights | *n* | Population |
|---|---|---|---|---|
| Slightham et al., | 2017 | 1 | 60 | Healthy |
| Gieselmann et al., | 2019 | 7 | 71 | Clinical |
| Gokce et al., | 2020 | 7 | 85 | Healthy |
| Medina et al., | 2015 | 9 | 33 | Healthy |
| Janku et al., | 2020 | 42 | 36 | Clinical |

*Table 76. Results for a mixed-effects model with S.E (Act) as dependent variable and Study (random effect), ID (random effect), and S.E (Sub; fixed effect) as predictors.*

| Predictor | *CI* | β | *SE_β* | *p* |
|---|---|---|---|---|
| S.E (Sub) | -0.06 – 1.09 | 0.05 | 0.03 | .078 |
| Pop (Clinical) | -24.17 – 3.86 | -0.48 | 0.34 | .156 |
| S.E*Pop | 0.27 – 1.78 | 0.08 | 0.03 | **.008**** |

*Note.* S.E is mean-centred and is expressed as a %. Population is a variable composed of two groups (*Healthy/Clinical)*. Marginal $R^2$ = 0.21; Conditional $R^2$ = 0.79; ICC = 0.74.

***Relatedness*.** The data presented in Table 75 above suggests that there is no significant relationship between subjective and actigraphy measures of Sleep Efficiency , nor were actigraphy measures influenced by Population. However, an interaction was observed such that the effect of a one unit increase in the dependent variable was greater among clinical participants (i.e., subjective, and objective measures were more closely related among healthy participants; Cohen's d = 0.09). See Figure 37 below for trends by group:

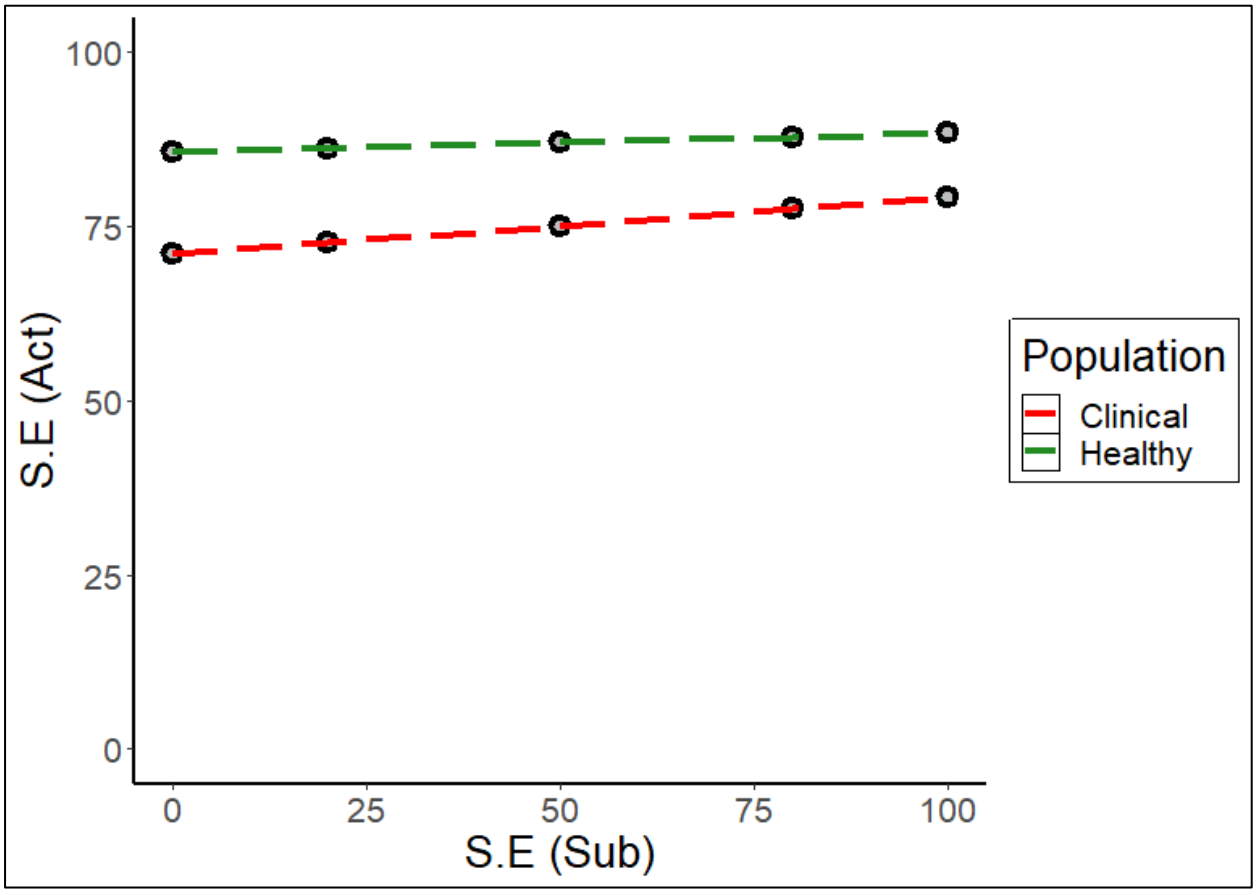
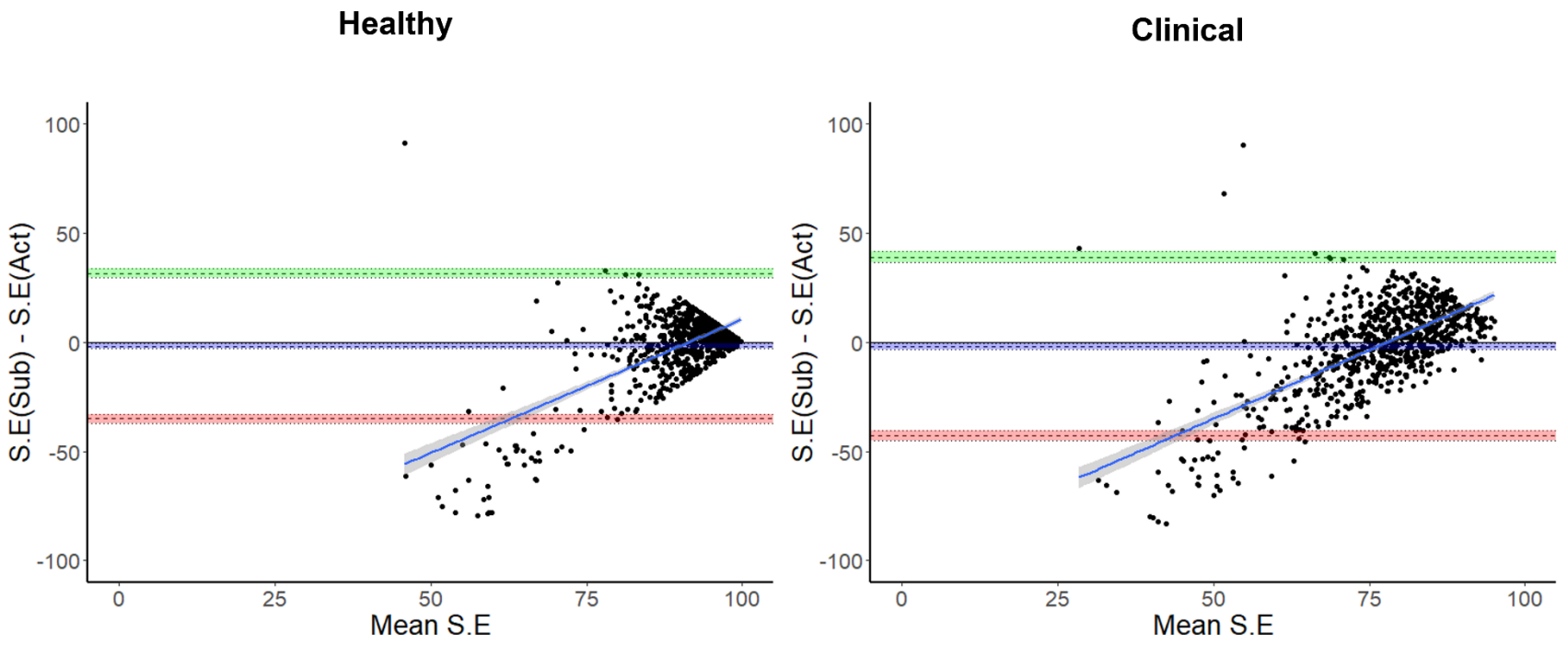



*Figure 36. Interaction plot for M1 (β = 0.08, SE_β = 0.03, p =.01**). Axes in %.*




Healthy (%; ULoA = 31.56, LLoA = -35.03, outliers = 5.38%). Clinical (%; ULoA = 39.05, LLoA = -45.04, outliers = 4.56%).
ULoA = Upper Level of Agreement (Green). LLoA = Lower Level of Agreement (red).

*Figure 37. Bland-Altman plots visualising agreement between S.E(Act) and S.E(Sub) in amongst healthy (left) and clinical (right) participants.*

***Agreement.*** Figure 12 above shows trends in agreement between subjective and actigraphy measures of Sleep Efficiency among Healthy and Clinical samples. There was a small amount of significant systematic bias among the Healthy group (Healthy: *M = -1.72, SD = 16.99, t(743) = -2.76, p = <.01, d = -0.10 )* such that subjective estimates underestimated actigraphy measures by 1.72% on average. There was also a small amount of significant systematic bias in the Clinical group (Healthy: *M = - 1.78, SD = 20.83, t(832) = -2.47, p < .05, d = -0.09)* with subjective estimates underestimating actigraphy measures by 1.78% on average. Both groups showed similar significant positive trends in proportional bias with mean differences increasing as Sleep Efficiency increased (Healthy: $R^2 = .40, F(1, 742) = 495.8 , p < .001$; Clinical: $R^2 = .44, F(1, 831) = 663.9, p < .001$).

ICCs show that statistical agreement is non-significant amongst Healthy participants and significant and very poor among clinical participants (Healthy: ICC = 0.03, p = .18, *ns*; Clinical: ICC = 0.06, p <.05). Overall, Sleep Efficiency measures do not significantly agree among healthy participants and very weakly amongst Clinical participants.

***Summary***

Tables 77 and 78 below summarise relatedness and agreement statistics report across analyses M1-M6, and Figure 38 visualises the pattern of agreement observed.

*Table 77. Relatedness statistics for each comparison M1-M6*

| Comparison | β | *SE_*β | *p* |
|---|---|---|---|
| TST (Act) ~ TST (Sub) | 0.72 | 0.13 | **<.001*** ** |
| TST (PSG) ~ TST (Sub) | 0.61 | 0.13 | **<.001*** ** |
| NA (Act) ~ NA (Sub) | 0.22 | 0.37 | **<.001*** ** |
| SOL (Act) ~ SOL (Sub) | 0.15 | 0.02 | **<.001*** ** |
| WASO (Act) ~ WASO (Sub) | 0.23 | 0.02 | **<.001*** ** |
| S.E (Act) ~ S.E (Sub) | 0.05 | 0.03 | *ns* |

*Table 78. Agreement statistics for each comparison M1-M6.*

| Comparison | ICC (Healthy) | *p* | LOA | ICC (Clinical) | *p* | LOA |
|---|---|---|---|---|---|---|
| TST (Act) ~ TST (Sub) | 0.67 | <.001*** | Moderate | 0.31 | <.001*** | Poor |
| TST (PSG) ~ TST (Sub) | 0.53 | <.001*** | Moderate | 0.34 | <.001*** | Poor |
| NA (Act) ~ NA (Sub) | 0.02 | ns | None | 0.00 | ns | None |
| SOL (Act) ~ SOL (Sub) | 0.17 | <.001*** | Poor | 0.11 | <.001*** | Poor |
| WASO (Act) ~ WASO (Sub) | 0.03 | <.01** | Poor | 0.05 | <.01** | Poor |
| S.E (Act) ~ S.E (Sub) | 0.03 | *ns* | None | 0.06 | 0.03* | Poor |

*Note. LOA = Level of Agreement. ICC = Intraclass Correlation Coefficient.*

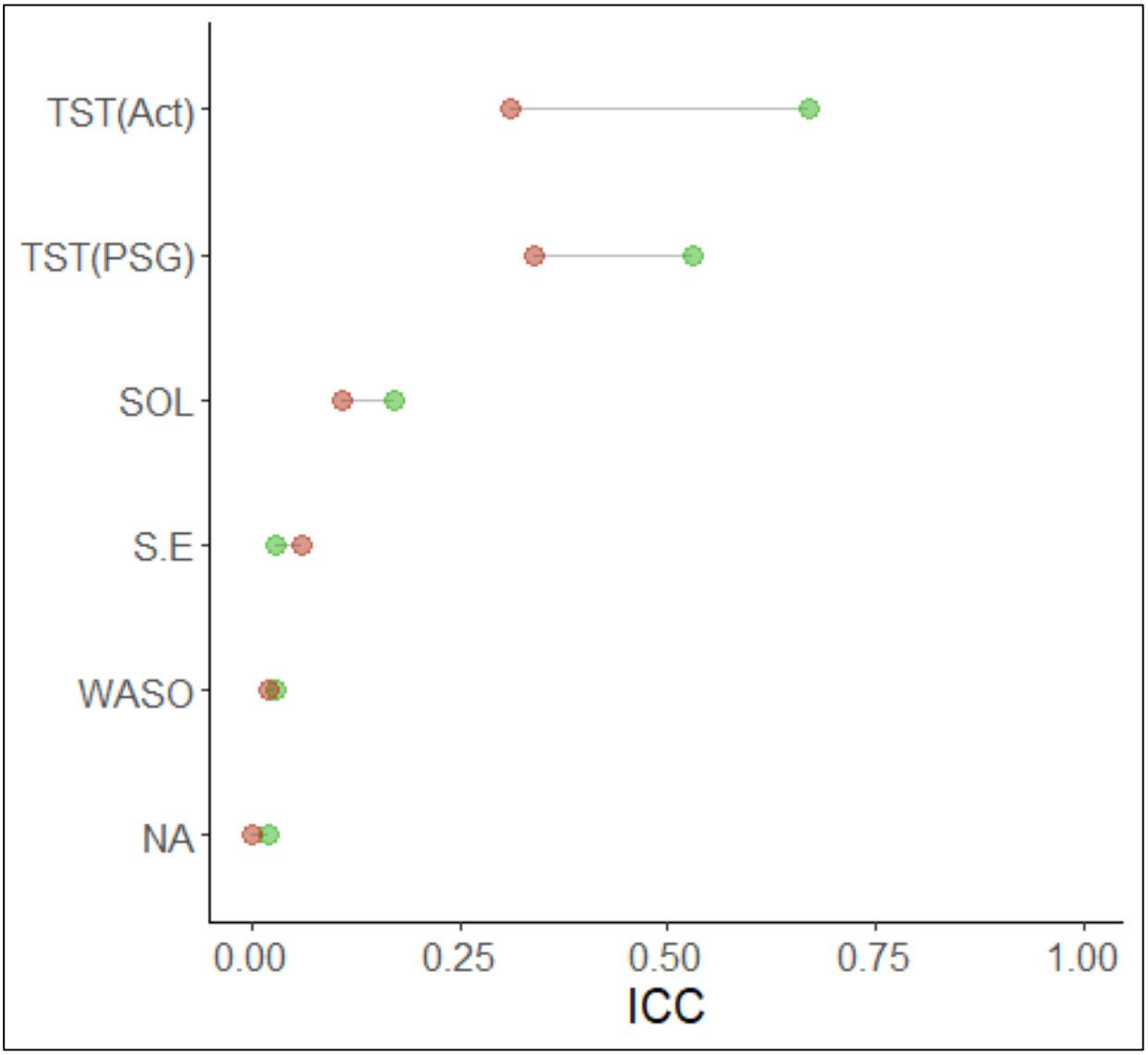



*Figure 38: Cleveland dot plot. All values plotted are for strength of agreement (ICC) between objective and subjective measures of sleep for M1-M6. Healthy participants (green); Clinical participants (red).*

# Discussion

This study sought to combine data across studies and calculate effect sizes for the relatedness and agreement between subjective and objective measures of sleep both among healthy adults and those with sleep disorders. Considering both of these dimensions are useful for understanding the quantitative strength of the association between subjective and objective sleep measures, but also whether or not they are qualitatively different constructs. Considering relatedness, all variable pairs were significantly related to one another except for sleep efficiency. The effect size was large for the relationship between actigraphy and diary measures of Total Sleep Time and small to very small when comparing Number of Awakenings, Sleep Onset Latency, and Wake after Sleep Onset using the same measures. Agreement was significantly compromised among clinical participants for all comparisons of actigraphy and sleep diary, with a medium effect size being observed for Total Sleep Time and small to very-small effect sizes for the rest. Considering statistical agreement, a similar pattern was observed such that a moderate level of agreement existed between measures of Total Sleep Time in healthy participants, a poor level of agreement for Sleep Onset Latency and Wake After Sleep Onset, and no agreement for Sleep Efficiency and Number of Awakenings. The moderate degree of agreement observed among measures of Total Sleep Time only held among healthy participants and was poor among clinical ones. Finally, these analyses suggest that subjective sleep measures overestimate Total Sleep Time among healthy and clinical individuals, underestimate Sleep Onset Latency among the healthy and overestimate it in those with sleep disorders, and underestimate Wake after Sleep Onset and Sleep Efficiency in both groups.

These findings are like those previously reported in several ways. The clearest consistency with our findings and that of the available literature is that having a chronic condition, here a sleep disorder, significantly and meaningfully limits the extent to which different measurements of sleep agree with and are related to one another. Our results are highly consistent with the meta-analyses conducted with Conley et al., (2019) in this regard, as well as the individual studies discussed above (Bianchi et al., 2013; DeFrancesco et al., 2021; O'Brien et al., 2016; Lubos et al., 2022; and Werner et al., 2016). Our results extend Conley et al's findings comparing healthy and clinical populations to subjective measures of sleep and bring together and centralise the

scattered comparisons of subjective and objective sleep measures that do exist for individual or smaller subsets of sleep parameters.

Concerning the individual sleep parameters, Total Sleep Time captured by diary seems to be the most reliable of the subjective measures both moderately related to and concords with objective measures. This is broadly consistent with the figures reported by Campanini et al., (2017) and Lauderdale et al., (2008) except that the effect sizes we observed were even larger. As well as this, observing weaker relationships and weaker agreement for most analyses among clinical participants is consistent with the many studies which report the poorer performance of subjective measures among clinical subtypes (e.g., Bianchi et al., 2013; De Francesco et al., 2021). Rather than any finding standing out, however, these results serve as a useful reference point to consider when using subjective estimates of sleep parameters and when interpreting subjective sleep data.

Also, like Conley et. al., (2019) is that the perceived-to-be-less accurate measure of sleep overestimated sleep and underestimated wake. Whilst it is obviously not our claim that humans apply the same erroneous wake/sleep classifications as actigraphy, a general theme can be gleaned across the two sets of analyses that detecting wake is something at present which is only accurately being achieved by polysomnography. Actigraphy struggles to detect wake due to difficulties classifying body movement and humans can detect wake but struggle to remember and quantify it after sleep. Our results show that the measure least reliant on human computation (diary measured Total Sleep Time) was the closest both in relatedness and agreement to their more objective comparator.

It is not surprising therefore that the subjective measure requiring the most human input, the total number of awakenings, was the least concordant with more objective measurement. The finding that there seems to be a small statistical relationship between actigraphy measured Number of Awakenings and those measured by diary and no significant concordance is not surprising but worth considering in isolation. This is because it is possible that both measurements (e.g., NA(Act) ~ NA(Sub) are capturing different things. It is possible that subjective estimates of total number of awakenings are just a measure of subjective sleep quality, and awakenings and subjective sleep quality

are clearly related (O'Donnell et al., 2009; Conte et al., 2022). Conversely, it is also possible that subjective sleep quality is driven by the number of awakenings, as argued by Harvey and colleagues (2008). Either way, subjective estimation of this construct is quite distinct from PSG or actigraphy algorithms. However, the possibility that the lack of agreement and low relatedness observed between actigraphy, and self-reported total number are awakenings is also partially explainable by the weaknesses of actigraphy. It is not inherently the case that more objective equates to being more accurate. Actigraphy is known to have problems misclassifying sleep and wake due to its reliance on physical activity (Ancoli-Israel et al., 2003). Considering awakenings specifically, actigraphy is less able than polysomnography to correctly classify patterns such as movement within sleep or physical stillness within wake (Marino et al., 2013). Despite being the gold-standard, polysomnography also has limitations. For example, it is a more intrusive, resource-intensive process which may not also capture typical sleep due to it most often only being conducted over a night or two (see Rosenberg et al., 2013 for an evaluation of polysomnography as a measurement tool for sleep). Generally, actigraphy tends to underestimate the number of awakenings compared to polysomnography (Kushida et al., 2001), especially when the awakenings are less than 5 minutes (Paquet, Kawinska, and Carrier, 2007). These findings advocate both for the combined use of objective and subjective sleep measures where possible.

If one were seeking to reduce the likelihood of confounding NA with subjective sleep quality and to measure NA on a large scale and in a practical fashion, then it is worth considering measuring awakenings that last 5 minutes or longer. In Experiment 3 (Chapter 3), for example, subjectively estimated awakenings lasting 5 minutes or longer across the previous month were positively associated with encoding error, overnight forgetting, and negative affect, respectively, as well as being negatively associated with positive affect. When the total number of awakenings was considered instead of those lasting 5 minutes or longer, almost all associations disappeared (all except for the association between awakenings and expressions of negative affect). One explanation for these results relates to research carried out by Winser and colleagues (2013) which suggests that it takes approximately 5 minutes of wakefulness for the experience to be remembered. Measurement at this threshold increases the likelihood that the computation will be more accurate than driven by feeling.

The key message to sleep researchers seems to be to apply caution when using subjective sleep measures in isolation. Except for Total Sleep Time, statistical relatedness and agreement is generally evident but effect sizes are small. When interpreting data on sleep parameters measured subjectively, researchers must also be careful to consider the heightened possibility of type 1 errors, even when studying healthy adults. When the study involves populations with sleep disorders, subjective sleep measurements will likely be greatly discordant and researchers might consider using grouping variables (e.g., *treatment/Control*) in statistical models or between-subjects designs and drawing inferences from any effects observed, taking care to control for any confounds as robustly as possible. For example, in Study 3 (Chapter 4) we conducted similar analyses to those just discussed in Study 2 but among participants with a chronic sleep disorder. In that study, we observed that having Restless Legs Syndrome was positively associated with encoding error and expressions of positive affect. We argued that these group differences were attributable to increased nocturnal awakenings in RLS since we had already controlled for Total Sleep Time and that sleep fragmentation is one of the defining characteristics of Restless Legs Syndrome. However, when we conducted a similar analysis substituting our grouping variable (RLS/Control) for a more direct measure of awakenings (subjective number of awakenings), no significant association was found (namely between number of awakenings and overnight forgetting whilst controlling for sleep duration and level of encoding in those with RLS).

The present study has several limitations. The first is that due to the flexible way that data was collected, we were unable to control for the eventual datasets we ended up with. This meant that in real terms that we were only able to conduct one analysis where PSG was the objective measurement and that almost all of our clinical samples were of insomnia patients. This limits the generalisability of the findings. Given that PSG and actigraphy are consistent, however, and that our intention was just to provide some basic rules of thumb for relatedness and agreement between objective and subjective measures, the data we gathered and analysed were largely appropriate towards facilitating this end.

The results are also limited by the fact that we were unable to determine the exact algorithms used in the actigraphy studies we included. It is possible therefore that some of the effect sizes calculated might vary depending on the algorithm used. However, only

studies with scientific grade wrist actigraphy were used so it is unlikely that the algorithms used were wildly inappropriate. It is better to characterise the results as a typical snapshot of actigraphy use within the sleep literature.

Finally, our results are limited by the fact that it was convenient to define 'healthy' as just an individual who was not diagnosed with or to our knowledge had complained of symptoms of a clinical condition. This is a limitation in that there is evidence to suggest that relatedness and concordance is affected by individual characteristics within the broader banner of 'healthiness' (e.g., race, Jackson et al., 2018; mood, Baillet et al., 2016; gender, Girshik et al., 2012). Those looking to conduct research within a specific race or gender, for example, might seek more tailored estimates from the literature. We argue though that our results are useful for those looking to study a broad range of healthy adults.

Concerning this thesis, the results support the use of subjectively measured sleep duration in Experiment 3 and Experiment 4 but advocate caution for the use of subjectively measured number of awakenings. This is because whilst the results suggest that subjectively measured awakenings are to some degree (weakly) related to more objective measures, they are not in agreement (i.e., are not measuring the same thing). Building on discussions in Chapter 1, objectively measured awakenings are defined as 30-second epochs of Stage W (Polysomnography; AASM, 2020) or increased physical activity over time (Actrigraphy), whereas subjectively measured awakenings are based on memory and subjective experience. The low levels of relatedness and null agreement between objective and subjective measures of awakenings observed in this Chapter leave it difficult to interpret some of the results in Experiments 3 and 4. For example, it is unclear whether or not the null results observed in Experiment 3 between the total numbers of awakenings and memory task scores reflect a true null result or a Type II error.

There are several future directions stemming from our results. Our results suggest that since subjective and objective measures are generally speaking significantly related to one another, but that agreement between the two is mostly underwhelming, that the two are at least to some extent capturing something qualitatively distinctive. For researchers interested in better understanding what the qualitative differences between

subjective and more objective sleep measures are, a clear starting point seems to be to better understand the extent to which subjective measures of sleep are biassed by the individual's impression of how well they slept. Similar analyses to the current investigation could be conducted with the inclusion of Subjective Sleep Quality as a fixed effect. If multicollinearity was a concern between e.g., subjective number of awakenings and subjective sleep quality in a model measuring the relationship between actigraphy and diary-measured number of awakenings, then this would be indicative that two are capturing something similar. Finally, it would be straightforward and informative to calculate the Intraclass Correlation Coefficient for the agreement between subjective sleep quality and each of the five sleep parameters included in the current set of analyses.

The current study sought to opportunistically combine raw data from several studies within the sleep literature identified by rapid review and calculate guideline effect sizes for the relatedness and agreement of subjective and objective sleep measures among both healthy and clinical adults. The general pattern of results is that subjective and objective sleep measurements are to some degree statistically related and in agreement, but effect sizes are generally small or insignificant except for Total Sleep Time. As well as this, having a sleep disorder weakened relatedness in most instances and agreement in all cases where statistical agreement was observed among healthy participants. We argue that these results serve as a useful reference point for researchers looking to use and interpret data from subjective estimations in insolation and researchers working with clinical populations should seek to consider complementary strategies to offset against the deterioration in relatedness and agreement between subjective and objective measures in these groups.

# Chapter 6: General Discussion

This thesis sought to inform models of sleep and declarative memory and models of sleep and affect respectively by trying to isolate and quantify the relationship between sleep fragmentation and deficits in these domains whilst controlling for sleep duration. A rich literature exists detailing the negative influence of sleep deprivation on declarative memory and affect, yet relatively little research has been carried out which has examined sleep fragmentation in isolation in humans. One reason for this is that it is difficult to adequately control for the potentially confounding influence of sleep deprivation on declarative memory and affect among human participants.

A review of the literature in Chapter 1 identified several populations with unique profiles of sleep disruption which made them ideal to include in naturalistic observational studies. These target groups were adults with Obstructive Sleep Apnoea Syndrome (OSAS), those with Restless Legs Syndrome (RLS), and parents in the postpartum period. These studies, albeit limited by not having objective sleep measures, have improved understanding of the impact of sleep fragmentation on declarative memory and affect and as such have also extended knowledge of sleep and declarative memory and sleep and affect generally. As well as this, important practical implications stem from these studies for those with RLS, new parents, and those considering using subjective sleep measures in research. After summarising and reflecting on the key findings from each experimental chapter, I will outline the theoretical contributions of my research and consider the remaining open questions, practical implications and future directions that stem from it.

## Summary of experimental findings

### *Chapter 2*

As mentioned above, the thesis originally started as an investigation into the influence of Continuous Positive Airway Pressure (CPAP) therapy on declarative memory and affect among adults with OSAS. Therefore, the first year was spent designing the MASC-IT study, a longitudinal PSG study among OSAS patients in collaboration with the NHS. In preparation for this study, two methodological pilot studies were carried out to

empirically justify the task choice and combination for the larger MASC-IT study. Unfortunately, the COVID-19 pandemic made completion of this study impossible and as such the focus of the thesis shifted from then on to a more general exploration of the unique influence of sleep fragmentation on declarative memory and affect using data collected from online experiments.

With the MASC-IT study in mind, the two experiments in Chapter 2 sought clarity on two research questions. The first was to establish the optimal degree of neighbourhood (the degree of similarity between the novel word and previously established entries in the lexicon) between novel word stimuli for use in novel word-pair tasks in pre-sleep/post-sleep declarative memory experiments (i.e., whether closely vs distally related novel word-pairs were more sensitive to consolidation effects). To inform this question, 40 healthy undergraduates from the University of York were split into two groups and trained to learn 20 novel words that were derivatives of existing English words (e.g., *cathedruke*). Participants carried out a training phase in the evening, were tested using a cued-recall approach before leaving to go about their typical routine, and then re-tested the following morning. One group attempted to learn novel words closely related to their derivatives (*cathedral - cathedruke*) whilst the other group attempted to learn words novel words further removed from their derivatives (e.g., *cathedral - [cathedruke] - yothedruke*). It was observed as part of this exploratory investigation that whilst both groups improved to a similar degree overnight, there was a general advantage for closely related word pairs at encoding, making them more suitable for use in studies where encoding might already be relatively low (as might be expected among OSAS patients).

The second experiment sought clarification on whether interference would occur in the MASC-IT study if two tasks from different domains of declarative memory (verbal and non-verbal/ lexical and spatial) were used on the same nights. The same procedure as in the previous study (Experiment 1) was extended to include a spatial memory task among 80 healthy undergraduates in a counterbalanced fashion and it was hypothesised that no interference would occur since sleep would halt communication between the two memory systems. A sample wide overnight improvement was observed for both tasks and no evidence of any interference between the two was observed.

It was concluded that the two experiments justified use of closely related word pairs for the word task in the MASC-IT study and that a spatial memory task could also be carried out on the same nights without any significant risk of interference. The remainder of the chapter outlined the design and analysis plan for the MASC-IT study.

***Chapter 3***

Chapter 3 signified the beginning of a new phase in the thesis, pivoting towards focussing on isolating the unique contribution of sleep fragmentation to variance in declarative memory and affect whilst controlling for sleep duration. The overall strategy for the remainder of the data collection period was to collect data from large numbers of participants online and to take a multilevel modelling approach to control for sleep duration statistically. Experiment 3 was a pre-sleep/post-sleep online spatial memory study which attempted to tease apart the unique contributions of sleep fragmentation and duration among new parents, who are often considered to experience a rich degree of variation in both sleep fragmentation and sleep duration, and healthy controls. I measured encoding and overnight forgetting of spatial memory by gamifying a popular object location task (that of Rudoy et al., 2009) and captured a snapshot of positive and negative affect among participants using PANAS. Sleep parameters were also measured by self-report.

It was hypothesised that the number of nocturnal awakenings would be positively associated with the level of encoding error, the level of overnight forgetting, and expressions of negative affect, and negatively associated with expressions of positive affect, whilst controlling for sleep duration. These hypotheses were all supported by the results observed, but only when the number of awakenings at the level of the number of awakenings lasting 5 minutes or more for a typical night across the past month. As well as this, the positive association between the number of awakenings and overnight forgetting was only observed when the level of encoding error was included as a control variable in the multilevel model. Encoding deficits were also observed in new parents relative to healthy controls, regardless of parental age, suggesting that the postpartum period has a longer impact on cognitive function than was previously thought.

The results had several limitations, most notably the lack of objective sleep measurement, but were a good starting point towards robustly evidencing that continuity of sleep is important for the formation and consolidation of declarative memories and for the expression of affect even when controlling for time slept. Since our eventual sample in Experiment 3 slept in general better than we were expecting and perhaps had hoped for when designing the study, it remained unknown whether sleep fragmentation would still uniquely contribute to the variance in declarative memory and affect in the context of more significant disruption to and variation in sleep duration.

***Chapter 4***

Experiment 4 in Chapter 4 was complementary to Experiment 3 in that it asked many of the same questions under the same conditions as in Experiment 3 but among adults with RLS and healthy controls. RLS has a unique profile of sleep disruption characterised by increased sleep fragmentation and decreased sleep duration relative to healthy controls and new parents. As well as this, RLS symptoms are circadian-driven, and pool in the first half of the sleep cycle, with the intriguing implication that they might naturalistically experience chronic disruption to Slow Wave Sleep at the same time and having relatively preserved REM sleep.

It was hypothesised again that the number of nocturnal awakenings would be positively associated with the level of encoding error, the level of overnight forgetting, and expressions of negative affect, and negatively associated with expressions of positive affect, whilst controlling for sleep duration. Importantly, it was anticipated that the sleep duration observed in this study would be more disrupted and richer in variation than what was observed in Experiment 3, and as such would inform the question of whether sleep fragmentation is uniquely associated with deficits in declarative memory and affect whilst controlling levels of sleep disruption more akin to sleep deprivation. Several important findings arose from Experiment 4. Firstly, having unmedicated RLS was associated with increased encoding error and overnight forgetting (only whilst controlling for level of encoding error as in Experiment 3) whilst controlling for sleep duration. Importantly, sleep duration also significantly contributed to the variance in overnight forgetting. This, in combination with the fact that sleep duration was lower and more varied than observed in Study 3, and that histograms showed greater representation at levels of sleep duration

commonly described as partial sleep deprivation, suggested that sleep fragmentation (one of the defining characteristics of RLS) uniquely contributes to variance in deficits in declarative memory in the context of greater disruption to sleep duration than seen in Experiment 3.

Interestingly, the results in Experiment 4 deviated from those observed in Experiment 3 in two important ways. The first of these is that having RLS (who were significantly higher in nocturnal awakenings than control) was associated with reduced expression of positive affect than control but there were no group differences in the expression of negative affect. Finally, Experiment 4 deviated from Experiment 3 in that no significant associations were observed which relied on subjective measurement of nocturnal awakenings. This observation is consistent with the possibility that subjective sleep measures were disproportionately less concordant (statistical agreement) among RLS, a clinical population, than among new parents and healthy controls. However, the exact degree of statistical relatedness and agreement between subjective and more objective measures of sleep was unknown, leaving it hard to draw strong conclusions from the results. As well as this, there is not enough evidence to conclusively state that the group differences observed were driven by sleep fragmentation, nor was there enough evidence to conclude that the contradictory findings between H3 and H1 and H2 was due to concordance differences between RLS and parents.

***Chapter 5***

In the final experimental chapter, clarification was sought as to which constructs subjective sleep measures are capturing, how closely they are related to and in agreement with more objective measures, and ultimately when and when not to use them in a research setting. We took a flexible approach to informing this question and conducted a rapid review designed to return a manageable amount of data and to pool this data to calculate effect sizes for the relatedness and agreement between subjective and more objective sleep measures for the most measured sleep parameters in research. This led to 6 mixed-effects mega-analyses being carried out to test the relatedness between subjective and more objective measures of Total Sleep Time (TST), Number of Awakenings (NA), Sleep Onset Latency (SOL), Wake After Sleep Onset (WASO), and Sleep Efficiency (S.E). Statistical agreement was also calculated using the Intraclass Correlation Coefficient. Previous research also suggested that relatedness and

agreement might be poorer among clinical subgroups when comparing PSG and Actigraphy, so this was also tested in Chapter 5 as part of the same piece of research.

A general pattern emerged within which relatedness and agreement between subjective and more objective measures of sleep parameters was poorer among clinical subgroups. These results were consistent with the interpretations made across Experiments 3 and 4 but in general advocate caution when using subjective sleep measures among clinical participants in a research setting, especially the total number of nocturnal awakenings, which this thesis largely depended on.

Overall, the studies in this thesis were an attempt to better understand the influence of sleep fragmentation on the encoding and consolidation of declarative memories and on the expression of affect. As such, I focussed on untangling the contributions of sleep duration and sleep fragmentation to variation in declarative memory and affect and opted for large-scale naturalistic online experiments to inform this question. Across Experiments 3 and 4, the results are a preliminary step towards suggesting that sleep fragmentation does contribute to the variance in deficits in declarative memory and affect whilst controlling for sleep duration, with there even being some evidence to suggest that this association between fragmentation and cognitive impairment persists within the context of partial sleep deprivation. In the final sections I will discuss in more detail how these results have informed theories of the encoding and consolidation of declarative memories, models relating to sleep and the expression of affect, and findings of practical importance to postpartum parents, those with RLS, and those considering using subjective sleep measures.

## Contributions to models of sleep and declarative memory

The studies in this thesis were not intended to directly inform the dominant theories of sleep and declarative memory consolidation described above. Investigations capturing oscillatory patterns and objective measurements of awakenings are better suited to this. Instead, they aimed to show that sleep fragmentation is an important and overlooked influencer of the processes underlying these models and as such should be explored more often and more routinely framed in the central of the debateSHY is particularly influential towards informing the encoding of declarative memories, suggesting that

learning comes at the cost of increased synaptic saturation which must be restored in sleep as part of a process commonly referred to as synaptic downscaling. Once the memory representation has been temporarily stored in the hippocampus and somehow tagged as a priority, the active systems model suggests that sleep also plays the role of transforming the to-be-remembered representation from weak and isolated to strong, durable, and integrated within frontal long-term memory stores. A strong prediction of both theories is that depriving the individual of sleep will impair the encoding and consolidation of new declarative memory traces. For SHY, an inadequate degree of synaptic downscaling in sleep will send the individual into the next day with already saturated synapses which are unsuitable for new learning opportunities. For Active Systems, a limited opportunity to co-ordinate slow-oscillation-spindle coupling events during the night will send the individual into the next day with material learned the day prior still highly vulnerable to forgetting. Despite similarly strong predictions also applying to sleep fragmentation as opposed to deprivation, the evidence base is less robust for this form of sleep disruption, especially true for studies controlling for the potentially confounding effect of sleep deprivation among human adults. The studies in this thesis were dedicated towards informing this area of research.

All four experiments in this thesis were interested in the association between sleep and declarative memory broadly. The primary goal of the thesis for this domain, which was addressed across Experiments 3 and 4, was to investigate whether sleep fragmentation uniquely contributed to the variance in encoding and overnight forgetting whilst controlling for sleep duration. Experiment 1 used novel-word learning, Experiment 2 used a combination of novel-word and object location learning and Studies 3 and 4 used object location learning along with subjective measures of sleep online. The mega-analyses in Chapter 5 aided interpretation of the results across Experiments 3 and 4. It was found across these studies that: closely-related word pairs have an advantage over distally-related ones at encoding but not for overnight forgetting (Experiment 1 and Experiment 2); no overnight interference occurs between the consolidation of verbal declarative and non-verbal declarative learning (Experiment 2); subjective measures of sleep fragmentation or group membership characterised by sleep fragmentation were associated with encoding and overnight forgetting such that increased sleep fragmentation was associated with higher levels of encoding error and overnight forgetting whilst sleep duration was controlled for (Experiment 3 and Experiment 4); the

observed association between sleep fragmentation and overnight forgetting uniquely contributed to the variation in overnight forgetting even when the distribution of sleep duration was considerably richer in variation and more representative of partial sleep deprivation (Experiment 4). Importantly, only subjectively measured number of awakenings lasting 5 minutes or more for a typical night across the past month were associated with measures of declarative memory across both studies, although membership of both groups verified as having fragmented sleep (i.e., new parents and RLS) was associated with deficits in declarative memory. Finally, across Experiments 3 and 4, sleep fragmentation was only associated with overnight forgetting when the level of encoding was also included in the model as a control variable.

No result in this thesis was a conclusive demonstration of any of the theories discussed above. Taken together, however, they tell a story largely consistent with what is believed to be the case about sleep fragmentation and the encoding and consolidation of declarative memories. They broadly suggest that sleep and declarative memory are associated; that sleep fragmentation specifically and declarative memory are associated even when taking into account sleep duration; and that encoding deficits are easier to detect than those of overnight forgetting. It is also evident that each individual claim is in clear need of objective verification and replication.

The first question that arises out of results observed across Experiments 3 and 4 is what the mechanism is for sleep fragmentations' disruptive relationship with deficits in declarative memory. Given that the current focus of the literature is on slow-oscillation-spindle coupling events as being crucial for consolidation (e.g, Staresina et al., 2015, 2018), a starting point might be to consider how sleep fragmentation affects these. There are several possibilities. Firstly, repeated awakenings might only affect the incidence of coupling events, but not the overall system pertaining to the coordination of them. The result of this sequence of disruption would simply be that there would be a net loss of retention directly correlating with the incidence of aborted coupling events. Another more troublesome possibility is that the consolidation could become temporarily impaired (thrown off track) at a more general level. Similar to what Cha and colleagues observed in RLS participants (2020, discussed in Experiment 4), the overall degree of synchrony between oscillations involved in coupling events could be thrown off track in a systematic fashion as a result of a lack of resilience against repeated interruption. The consequences

of this framework would be more global in nature, potentially affecting the qualitative nature of each event and with it each attempted integrative interaction with the frontal cortex, with yet unknown consequences.

Arising out of the work of Lui and colleagues (2019) and the studies across Experiments 3 and 4 is the idea that sleep fragmentation is a unique way of attacking and harming the infrastructure of memory formation and storage separable and independent of sleep duration. A natural follow-up question to this observation is that if the weapon is different, will the wound also be different? Setting aside momentarily the long-standing debate as to how appropriate it is to compare the human brain to a computer, the metaphor is useful in the current circumstances (even if wrong, as described in Chirimuuta et al., 2022). It is intuitive and the held experience of many that interrupting the saving of a file can corrupt both the file and its intended wider network to be stored within. Similarly, it is possible in the absence of some unknown safety feature that if sleep is abruptly interrupted at the exact time of acquisition and integration with the existing knowledge base that the representation will be stored in some partial or corrupt form and that the wider network will have inherited a weak link and weakened structure as a result. As intriguing as this possibility is to investigate, testing it would be very difficult. One potential way to explore questions like this one is to artificially create these conditions and subject an artificial intelligence to them (e.g., Golden et al., 2022).

A final implication of acknowledging that the influence of sleep fragmentation on declarative memory deficits is independent of that of sleep deprivation is that any neurotransmitter resources taxed as a result of sleep fragmentation at night will likely have a knock-on effect the following day for all behaviours which rely on the same resources. The underlying logic of this idea comes from Walker and Goldstein's (2014) work on emotion regulation, within which they argue that disrupting REM sleep causes a cascade of events in the brain which imbalances the daytime and nocturnal distribution of the neurotransmitter resources required for healthy emotion regulation. If Walker and Goldstein's logic is applied to Lui and colleagues' (2019) work on the involvement of Serotonin in a dedicated sleep fragmentation circuit in fruit flies, then this would suggest that severe sleep fragmentation might tax Serotonin stores during the night with the effect of impaired daytime functioning reliant on the same neurotransmitter resources. Human

behavioural paradigms are well-situated towards exploring this possibility further, with the most-specifically-reliant-on-Serotonin daytime behaviours to be prioritised for investigation.

However, the results observed in Experiments 3 and 4 need to also be considered in the context of several limitations. Firstly, the circumstances surrounding the pandemic meant that we had no recourse to objective measurement of sleep parameters. Future research might seek to use similar paradigms to Experiment 3 and Experiment 4 but alongside a PSG subset to measure sleep parameters related to slow-oscillation-spindle coupling events. Secondly, it is possible that the measure of awakenings associated with declarative memory in Experiment 3 was confounded by subjective sleep quality. As well as this, the associations between RLS and declarative memory observed in Experiment 4 may have been confounded by a characteristic of RLS other than awakenings.

Future research might also seek to adapt the designs of Experiments 3 and 4 to challenge a notable alternative to systems consolidation, the Contextual Binding account of consolidation (Yonelidas et al., 2019). Whilst not originally being the intention in the lead up to Experiments 3 and Experiment 4, it became clear that had very brief awakenings been associated with overnight forgetting, then this would be difficult to account for within the Contextual Binding account. This is because within that account, SWS is purported to behave as a contextual interference blocker. It is not intuitive that very brief awakenings, ones in which the individual does not even remember the event (Winser et al., 2013) are considered environments in which memory-altering contextual interference could occur. Whilst our results in the end cannot be used to evaluate this challenge to the Contextual Binding account, they set the scene for some further research on the topic. Microarousals, for example, have been shown the interrupt sleep parameters associated with consolidation (e.g., sleep spindles; Schabus et al., 2006), and could be considered as a type of sleep fragmentation. It would be especially hard for Contextual Binding to account for an association found between microarousals and consolidation in a tightly controlled polysomnography study.

## Contributions to models of sleep and affect

Previous research suggests that sleep is influential in the governance of positive and negative affect (Tomasso et al., 2020) and that slow wave sleep (SWS; Finan et al.,

2015) seems to be more important for next-day expressions of positive affect and REM sleep seems to be more important for next-day expressions of negative affect (Walker and Goldstein, 2014). Specifically, there is some research that has implicated dopamine levels in the expression of positive affect (see Krause et al., 2017, review) and noradrenaline levels in the expression of negative affect (Walker and Goldstein, 2014). Like the situation with declarative memory, the unique contribution of sleep fragmentation to the variance in positive and negative affect is relatively less understood than is the case for sleep deprivation. A secondary aim of this thesis was to investigate the association between sleep fragmentation and expressions of positive and negative affect whilst controlling for sleep duration.

Experiments 3 and 4 both used PANAS to calculate quantitative scores for positive and negative affect the day following a typical night of sleep for the participant, and these were entered into different mixed models as dependent variables. Associations were tested between affect and subjectively measured number of awakenings whilst controlling for sleep duration (Experiment 3) and between affect and new parenthood (Experiment 3) and RLS (Experiment 4), both of which were significantly higher in number of awakenings and sleep duration than controls. In Experiment 3, as awakenings increased, so too did the expression of negative affect, and this was the case regardless of whether awakenings were measured at the level of total awakenings the night before, total number of awakenings lasting 5 minutes or longer the night before, or the typical number of 5 minute or longer awakenings participants had across the night in the past month. As well as this, increased number of awakenings significantly predicted decreased expressions of positive affect, but only for awakenings measured at the level of 5-minute awakenings across the past month. An interesting pattern of results was observed when we tested associations between affect and new parents (Experiment 3) and RLS (Experiment 4) respectively. There was no association between having a child under the age of 6 and expressions of positive or negative affect the following day compared to healthy controls. And, for RLS, reporting symptoms of RLS significantly predicted decreased positive affect relative to healthy controls but did not predict the expression of negative affect.

Taken together across studies, this set of results were taken as support for sleep fragmentation's contribution to the variance in both positive and negative affect even when controlling for sleep duration. Specifically, they suggest that continuity of sleep is

important for the governance of both domains of affect even when controlling the amount of time slept. Future research directions and limitations will now be discussed.

Future research ought to consider the role of dopamine in both encoding and positive affect. The evidence implicating dopamine in both the regulation of positive affect and measures of declarative memory are considerable. Regarding positive affect, Yin (2019) reviews this topic and shares numerous examples of dopaminergic activity being positively associated with measures of positive affect. Krause et al., (2017) also argue that sleep deprivation lowers the activity of dopamine receptors. Further, Isotalus et al., (2020) found that nocturnally administered L-DOPA increased SWS by 11%. Finally, I also found in Experiment 3 that positive affect was associated with decreased encoding error when negative affect was not. Taken together, it is clear that dopamine plays a role in both positive affect and encoding and consolidation, but the exact mechanism for how is unclear. One possible explanation as to how dopamine can be involved in positive affect and encoding (synaptic downscaling) and consolidation (active systems) within SWS (and the more likely one) is that declarative memory depends on SWS which in turn depends on dopamine but that positive affect does not depend on SWS. The other possibility is that SWS is responsible for encoding, consolidation, the cleaning of brain cells with cerebrospinal fluid, and the regulation of positive affect via several distinct yet simultaneous mechanisms. The latter sounds less likely, and it is very unclear which properties of SWS would facilitate the regulation of positive affect if so. One way to explore these competing frameworks (and indeed other similar ones) would be to take the opposite approach to Isotalus and colleagues and administer dopamine antagonists during sleep among individuals with normal-high levels of natural dopaminergic functioning. If positive affect is reliant on SWS, then the temporary blocking of dopamine in sleep ought to impair next-day expression of positive affect; if not, only declarative memory will be impaired due to a reduction in SWS.

The evidence discussed has some important limitations, however. The main limitation is that I was not able to measure dopaminergic activity or sleep parameters relating to SWS directly and as a consequence of that therefore the results cannot speak towards these issues. As well as this, one of the main concerns leading up to data collection was to maximise participant retention and sample size as much as possible. As a result, I opted to use the short-form version of PANAS and did not have the

opportunity to expand my investigation to include more lengthy and rigorous measures of affect which might have increased confidence in the results observed.

## Reflections on the use of subjective sleep measurements

The main experimental chapters in this thesis (Experiment 3 and Experiment 4 in chapters 3 and 4 respectively) relied heavily on subjective measurements of sleep parameters. Previous research is unclear as to the degree of relatedness and agreement between subjective and more objective sleep measures but a recent meta-analysis on the same topic except between PSG and actigraphy suggested that relatedness and agreement would be poorer among those with chronic conditions (Conley et al., 2019). Therefore, A series of mega-analyses were carried out in Chapter 5 to aid interpretation of the findings in Experiment 3 and Experiment 4 and to try and establish the degree to which subjective and more objective sleep measures are related and the degree to which they can be said to measure the same thing (statistical agreement). We narrowed our search to include healthy adults and those with sleep disorders since sleep is directly affected in these groups and therefore sleep measures among them are likely to be the most acutely impacted. Relatedness and agreement were measured between subjective and objective measures of Total Sleep Time (TST), Number of Awakenings (NA), Sleep Onset Latency (SOL), Wake After Sleep Onset (WASO), and Sleep Efficiency (S,E) used a mixed-effects paradigm. It was observed that: all subjective-objective pairs were statistically related except for S.E; effect sizes were small, except for TST, which was large; all subjective-objective pairs were in statistical agreement, as measured by calculating Intraclass Correlation Coefficients (ICCs), except for NA; effect sizes were small, except for TST, which was moderate; relatedness and agreement was considerably poorer in all cases where relatedness or agreement could get materially poorer.

This series of mega-analyses aided interpretation of Experiment 3 and Experiment 4 in the following ways. Firstly, they offered some support for the most basic empirical assumption driving these two studies - that subjective sleep measures to some degree were related to and in agreement with more objective ones. As well as this, they justified the decision to include a broader range of measurements of sleep fragmentation in Experiment 3 and Experiment 4. Only measuring sleep fragmentation by having

participants self-report the total number of awakenings they had the night before would have been potentially problematic since subjective measures of awakenings and more objective ones only seem to be weakly statistically related and not significantly in agreement. Since the lack of agreement observed is most likely largely driven by the fact that human adults don't tend to remember awakenings that last less than around 5 minutes (Winser et al., 2013), complementing our measures of sleep fragmentation to also capture self-reports more likely to be driven by memory than feeling was arguably an important and justified safeguard towards false negatives. Finally, the widespread deterioration of relatedness and agreement observed in Chapter 5 among clinical subgroups justified the decision to include parenthood (Experiment 3) and RLS (Experiment 4) as predictors and to analyse separately the degree to which they differed to controls in the number of awakenings and that there were no relevant confounding variables. Taken together, Experiment 3 and Experiment 4 were taken to reflect the patterns observed in Chapter 5 in that: (1) self-reports of the total number of awakenings were largely insensitive towards associations that were argued to exist; (2) measures of sleep fragmentation driven by memory rather than feeling (5 minute awakenings) appeared to be more sensitive among healthy participants; and (3) insensitive among those with a chronic sleep disorder (RLS).

Overall, Chapters 3-5 suggest a clear path for those considering using subjective sleep measures in research. Specifically, the results suggest that: (1) objective measures should be used when convenient to do so, and the design doesn't inherently require very large sample sizes; (2) subjective measures are very useful tools when data needs to be collected fast and at a large scale; (3) when possible, PSG or actigraphy subsets should be considered to validate the subjective measures used; subjective sleep measures should not be relied upon in studies involving adults with sleep disorders; (4) when researching clinical populations, capturing the relevant phenomena, e.g sleep fragmentation, by using a group variable and being cognisant of and taken steps to avoid confounds is optimal.

Shifting from a practical to a theoretical level, the results in chapter 5 offer some support (some degree of statistical relatedness in the context of low statistical agreement) for the idea that subjective sleep measures are capturing something qualitatively different from more objective ones. This was a concern across Experiment 3 and Experiment 4,

where specifically I was concerned that the associations observed between subjectively measured awakenings and measures of declarative memory might have been confounded by subjective sleep quality. I maintain that this is less of a problem for awakenings than for subjectively measured sleep in general since research suggests that the relationship between subjective sleep quality and awakenings exists not because awakenings are just subjective sleep quality, but instead because subjective sleep quality is largely driven by awakenings (e.g., Harvey et al., 2008). In general, however, the mixed-effects mega-analytic paradigm used in Chapter 5 is very well-situated towards exploring this question in greater detail than was possible in this thesis. An obvious starting point is to collect new data or pool data from authors capturing the same variables as in Chapter 5 but also a measure of subjective sleep quality and including it in the same mixed-effects models as a control variable would help inform: the extent to which each of the most commonly measured sleep parameters are confounded by subjective sleep quality; and; whether or not subjective sleep quality is the only other thing subjective sleep measures of capturing outside of reflecting that captured by more objective measures.

## Practical implications and future directions for new parents and RLS

Chapters 2 and 5 were inward facing and made theoretical contributions to sleep and memory research already discussed above. However, Experiment 3 (parents) and Experiment 4 (RLS) also have practical implications for the individual difference groups who participated in them. I was very motivated to identify and contribute to open questions in the respective literature on each group, partially out of gratitude towards the many individuals who took part in my studies who may have been tired, under stress, or in general going through a particularly difficult time in their lives. My many interactions with the Restless Legs Syndrome community in particular left me with the impression that many with this condition are struggling, searching for answers, and eager to contribute to research that might help find them.

Concerning new parents first (Experiment 3), the first practical implication arising out of this thesis is that with regards to the formation of declarative memories at least, “baby brain” is not a myth. New parents do seem to be at a deficit in the encoding of declarative information relative to comparable non-parents. This was especially the case for parents within the first two years of the postpartum period, who reported the most

frequent awakenings and lowest sleep duration of the parents who took part. As well as this, cognitive deficits experienced by new parents are likely more serious than we were able to detect. This is both because parents with the youngest children were relatively under-represented in Experiment 3 and it is intuitive to suspect that the most sleep-disrupted parents with very young children were less likely to take part and more likely to drop out. What was more surprising, however, was that encoding deficits were observed in all categories of parents, even those with children aged 6 or over. This goes against the grain of previous literature suggesting that parents' cognitive functioning is fully restored by around age 6 (Richter et al., 2019). To test this idea fully, however, a longitudinal approach would be required. These tentative findings suggest that changes in the formation of declarative information among new parents are far more enduring than was previously thought. If so, the obvious question arising out of these results is whether brain changes with regards to memory in new parents are permanent or not.

It was also striking to observe that RLS was associated with declarative memory deficits, the expression of positive affect but not negative affect, and that medication use only seemed to be beneficial towards memory (Experiment 4). What is especially interesting about this set of results is that there were no significant differences between individuals with RLS who reported taking medication for RLS and controls This implies that medication use in RLS has some protective benefit for memory but not for affect, which is hard to understand. Of course, strong claims cannot be made about medication use in RLS from this thesis because the exact medications participants were taking and how often were not measured. Taken together, these findings suggest that: deficits in the formation and storage of declarative memories exist among RLS and that both sleep deprivation and sleep fragmentation make significant and meaningful contributions to them; positive affect appears to be specifically targeted whilst negative affect is to some degree spared; RLS medication use may help the restoration of memory but not regulation of positive affect.

The results of this thesis echo a wider and growing sentiment about the subtle, accumulative damage large pockets of society are enduring as a result of the demands of modern life. A larger and larger body of evidence is mounting, some of which was presented here, to suggest that each night with broken or short sleep carries with it a small penalty towards how we feel and how well we function. Many of the individuals who

were kind enough to take part in the studies across this thesis unfortunately have little recourse against the long-term consequences of unhealthy sleep. For many, however, the harm is preventable and the treatment highly accessible and effective. Perhaps more acknowledgement needs to be given to the fact that whilst our higher functions are clearly adaptive, they also clearly come at an evolutionary cost. The evolutionary plane has shifted more abruptly for humanity in the last 100 years than perhaps the last 10,000, and one of the things it has exposed is that the complex infrastructure required to maintain our mental health and cognitive functioning is delicate, vulnerable, and ought to be prioritised.

## Concluding remarks

This thesis attempted to isolate and quantify the association between sleep fragmentation and the encoding and consolidation of declarative memories, as well as the expression of positive and negative affect. The evidence across studies are a starting point towards suggesting that more frequent nocturnal awakenings are associated with increased encoding error, decreased expression of positive affect, and increased expression of negative affect. They also suggest that increased awakenings might be associated with poorer consolidation, although only when the level of encoding is controlled for. Importantly, these associations were observed whilst controlling for time slept. In sum, the evidence presented here suggests that maintaining healthy sleep has at least two key vulnerabilities, which have at least some degree of independence from one another, of which future research might seek to establish to what extent they carry their own unique consequences.

# Appendices

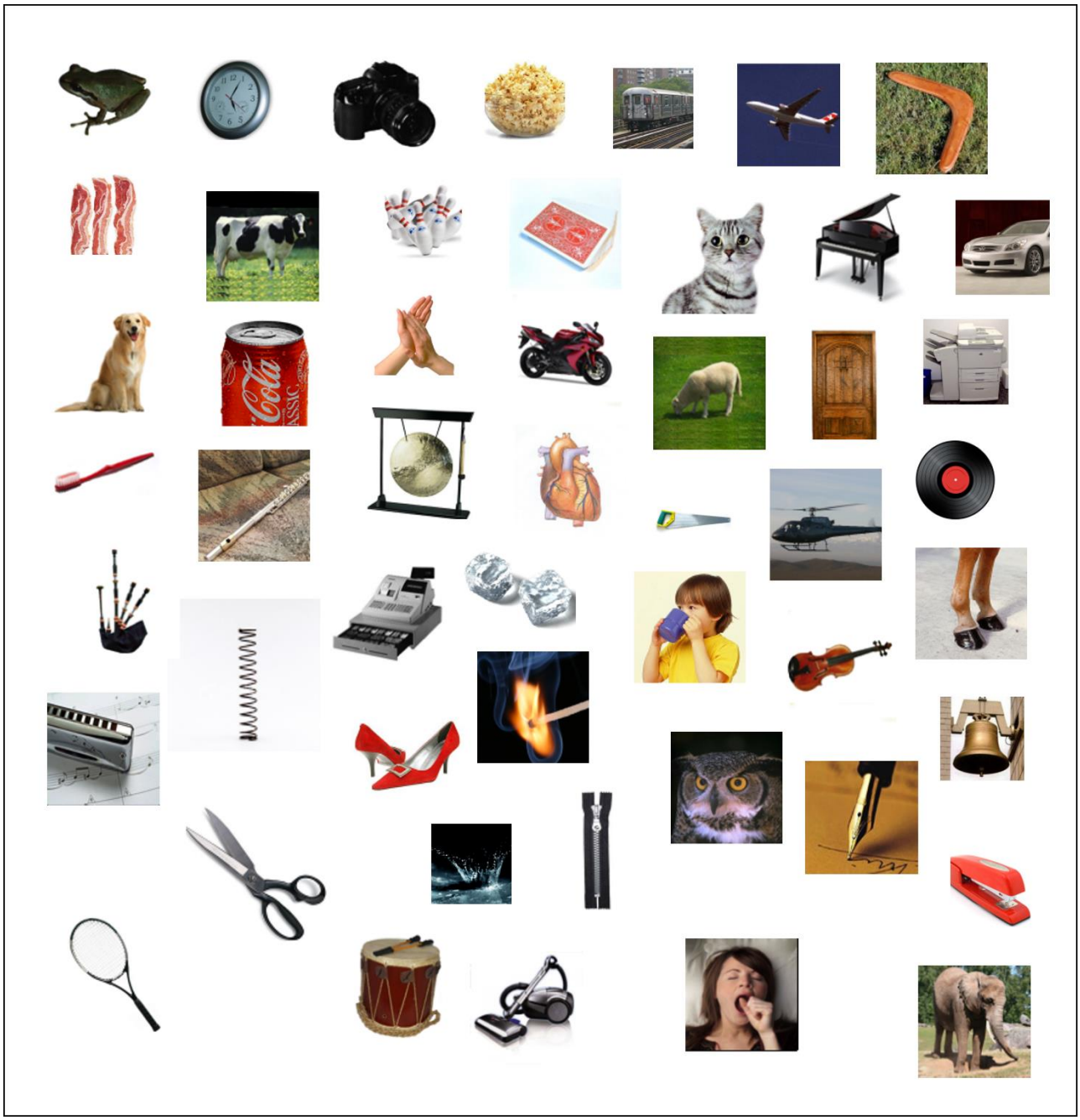
## Appendix A

Novel words used in Chapter 2.

| English word | Close Neighbours | Distant Neighbour |
|---|---|---|
| ravine | ravooce | bovooce |
| clarinet | clarinern | skorinern |
| pedestal | pedestoke | fadestoke |
| assassin | assassool | eggassool |
| apricot | aprickel | odrickel |
| anecdote | anecdel | olecdel |
| ornament | ornameast | atnameast |
| badminton | badmintel | sudmintel |
| mandarin | mandarook | gondarook |
| molecule | molekyen | vulekyen |
| partridge | partred | kortred |
| consensus | consensom | dinsensom |
| bayonet | bayoniss | feyoniss |
| specimen | specimal | trocimal |
| octopus | octopoth | artopoth |
| crocodile | crocodiss | flecodiss |
| skeleton | skeletobe | griletobe |
| pelican | pelikiyve | balikiyve |
| cathedral | cathedruke | yothedruke |
| pyramid | pyramon | biramon |

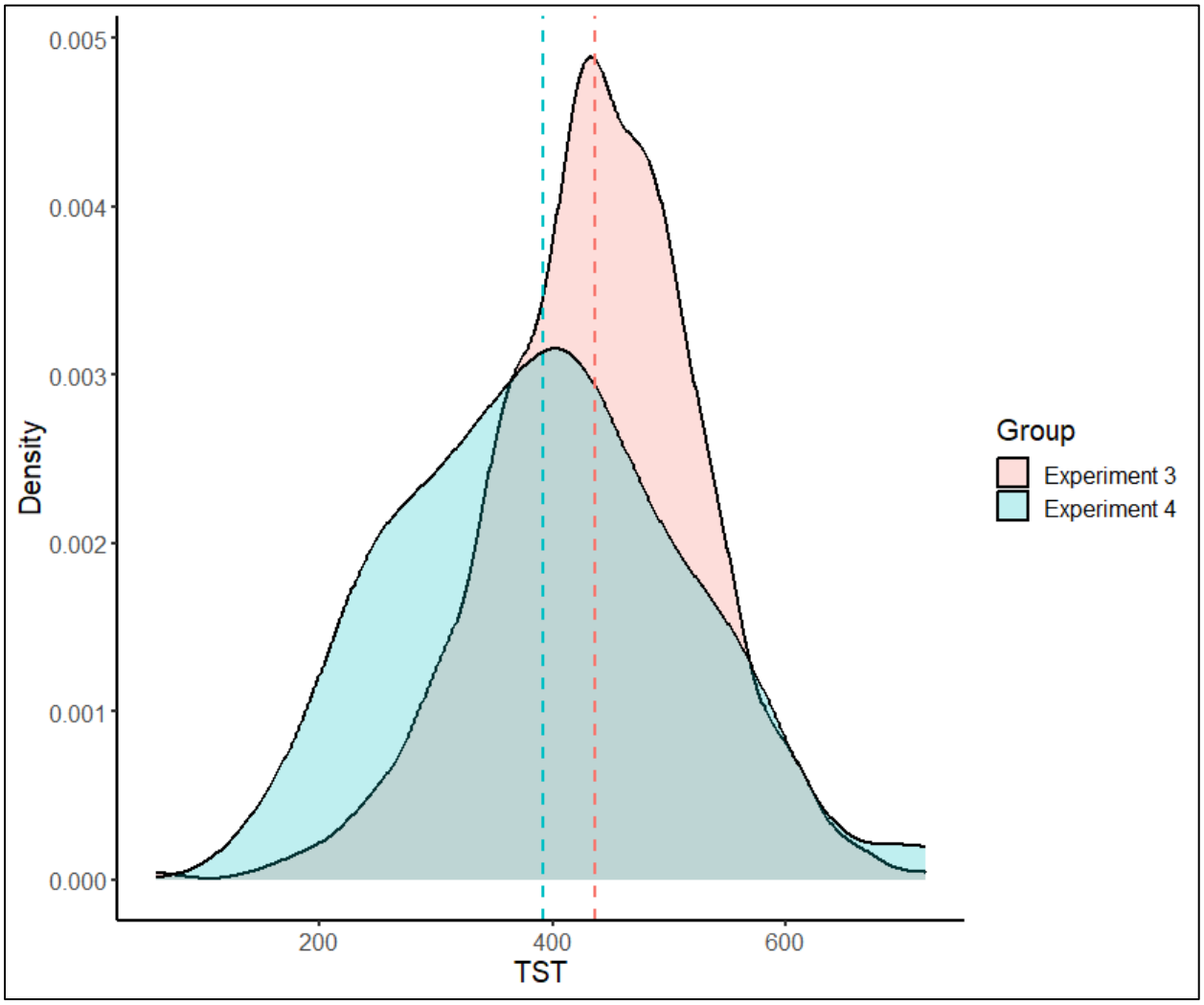
## Appendix B

The images used in the object location task in Experiment 2 (*n* =50).

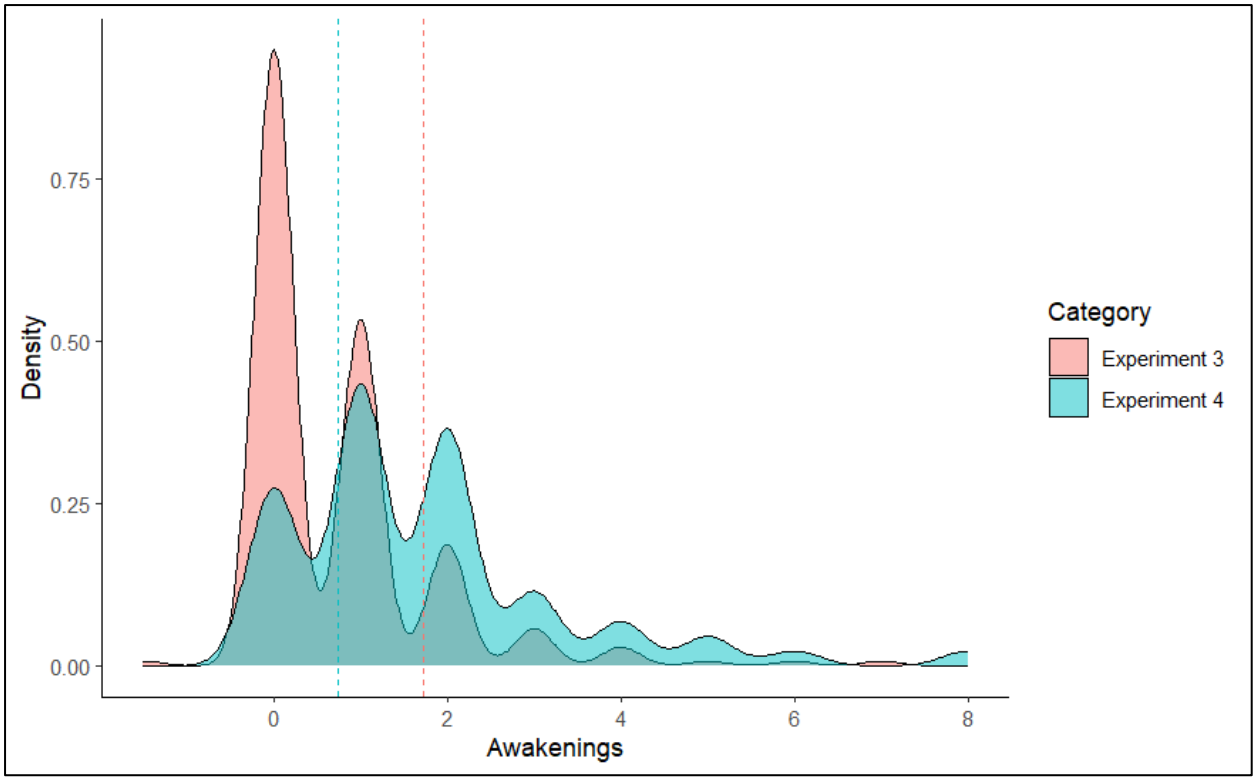


## Appendix C

Overlapping density plot with mean lines for sleep duration the evening before Session 2 for Experiments 3 and 4 samples. The plot shows greater representation of partial sleep deprivation (<6 hours) in Experiment 4.

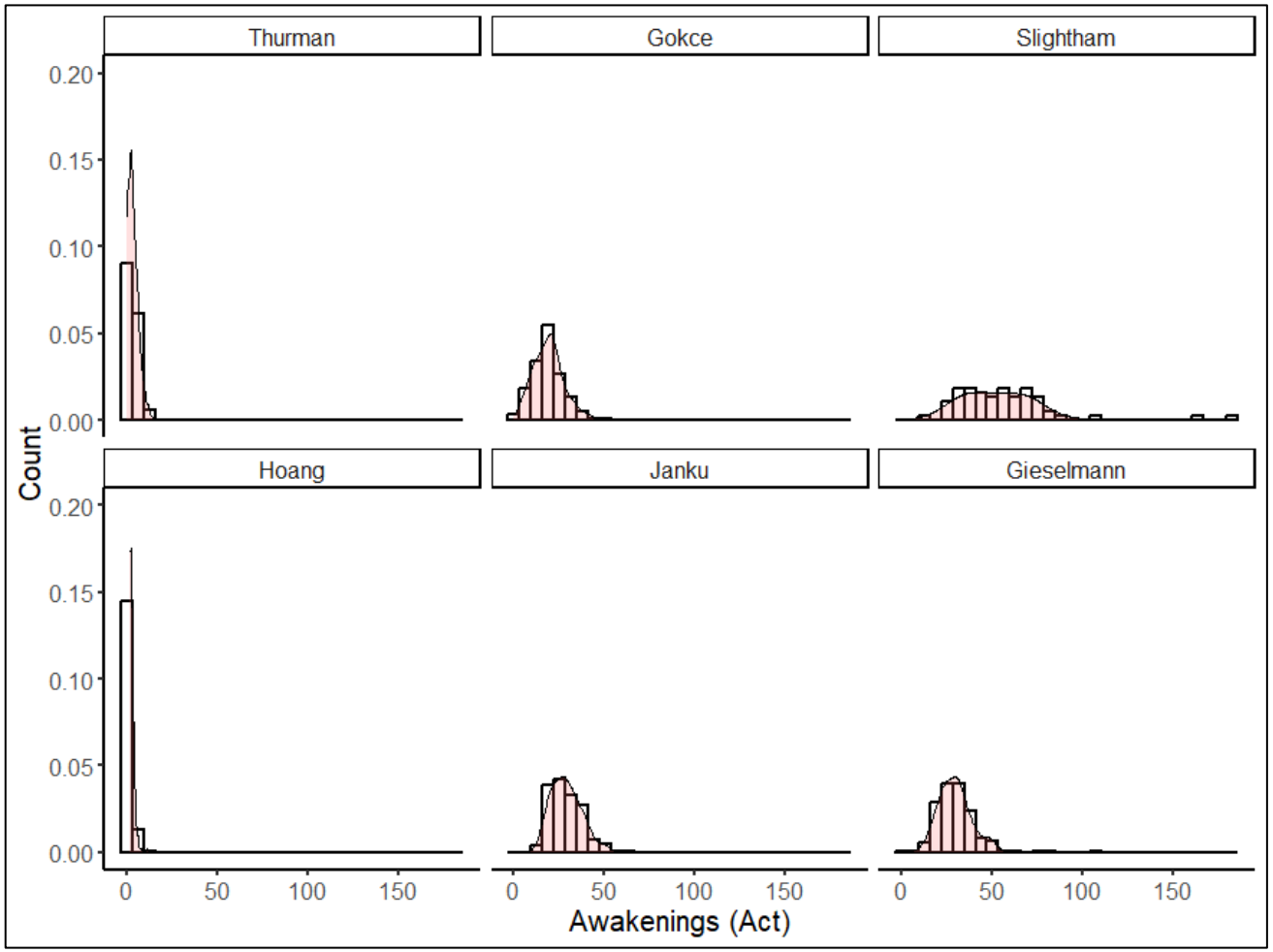


Similarly, the plot below shows a greater probability of a participant in Experiment 4 having a higher frequency of nocturnal awakenings in a typical night across the last month, despite the overall average being slightly lower.



## Appendix D

Distributions of number of awakenings for studies in Chapter 5 M3. Thurman and Hoang have distributions that suggest use of a different actigraphy algorithm than the rest.



# References


Abrishami, A., Khajehdehi, A., Chung, F. (2010). A systematic review of screening questionnaires for obstructive sleep apnoea. *Canadian Journal of Anesthetics*, 57, 423-428.

Ageborg Morsing, J., Smith, M., Ögren, M., Thorsson, P., Pedersen, E., Forssén, J., & Persson Waye, K. (2018). Wind Turbine Noise and Sleep: Pilot Studies on the Influence of Noise Characteristics. *International Journal of Environmental Research and Public Health*, *15*(11), 2573. https://doi.org/10.3390/ijerph15112573

Ahuja, A., Chen, R.K., Korey Kam., Pettibone, W.D., Osorio, R.S., and Varga, A.W. (2018). Role of normal sleep and sleep apnea in human memory processing. *Nature and Science of Sleep*, 10, 255-269.

Åkerstedt, T. (2003). Shift work and disturbed sleep/wakefulness. *Occupational Medicine*, 53(2), 89-94.

Allen, R. P. (2015). Restless Leg Syndrome/Willis-Ekbom Disease Pathophysiology. *Sleep Medicine Clinics*, *10*(3), 207–214. https://doi.org/10.1016/j.jsmc.2015.05.022

Allen, R.P., Picchietti, D., Hening, W.A., Trenkwalder, C., Walters, A.S., Montplaisi J, *et al*. (2004). Restless legs syndrome: Diagnostic criteria, special considerations, and epidemiology. A report from the restless legs syndrome diagnosis and epidemiology workshop at the National Institutes of Health*. Sleep Med*, 4, 101-19.

Alves, M. V. C., &and Bueno, O. F. A. (2017). Retroactive interference: forgetting as an interruption of memory consolidation. Trends Psych. /Temas em Psicol., 25, 1055–1067.

Alzoubaidi, M., and Mokhlesi, B. (2017). Obstructive sleep apnea during REM sleep: Clinical relevance and therapeutic implications. *Curr Opin Pulm Med*, 22(6), 545-554.

Ancoli-Israel, S., Cole, R., Alessi, C., Chambers, M., Moorcroft, W., & Pollak, C. P. (2003). The role of actigraphy in the study of sleep and circadian rhythms. *Sleep*, 26(3), 342-392.

Ansara, D., Cohen, M.M., Gallop, R., Kung, R., and Schei, B, (2005). Predictors of women's physical health problems after childbirth. *J. Psychosom. Obstet. Gynecol.,* 26, 115–125.

Antic, N.A, Malow, B.A, Lange, N., McEvoy, R.D, Olson, A.L, Turkington, P., Windisch, W., Samuels, M., Stevens, C.A, Berry-Kravis, E.M., Weese-Mayer, D.E. (2006). PHOX2B mutation-confirmed congenital central hypoventilation syndrome: presentation in adulthood. *Am J Respir Crit Care Med* 174: 923–927.

Antunes, B.M., Campos, E.Z., Parmezzani, S.S., Santos, R.V., Franchini, E., and Lira, F.S. (2017). Sleep quality and duration are associated with performance in maximal incremental test. *Physiol Behav.,* 177, 252-256.

Arora T, Hosseini-Araghi M, Bishop J, Yao GL, Thomas GN, Taheri S. The complexity of obesity in U.K. adolescents: relationships with quantity and type of technology, sleep duration and quality, academic performance and aspiration. Pediatr Obes. 2013 Oct;8(5):358-66. doi: 10.1111/j.2047-6310.2012.00119.x. Epub 2012 Dec 13. PMID: 23239604.

Arzt, M., Young, T., Finn, L., Skatrud, J.B., and Bradley, T.D. (2005). Association of sleep-disordered breathing and the occurrence of stroke. *Am J Respir Crit Care Med* 172, 1447–1451.

Astbury, L., Bennett, C., Pinnington, D. M., & Bei, B. (2022). Does breastfeeding influence sleep? A longitudinal study across the first two postpartum years. *Birth*.

Aubert-Tulkens et al., (1987). Ambulatory Evaluation of Sleep Disturbance and Therapeautic Effects in Sleep Apnea Syndrome by Wrist Actigraphy Monitoring. *AM REV RESPIR DIS*, 136(4), 851-6.

Axmacher, N., Helmstaedter, C., Elger, C.E, and Fell, J. (2008). Enhancement of neocortical-medial temporal EEG correlations during non-REM sleep. *Neural Plast*, 563028.

Ayalon, L., Ancoli-Israel, S., Drummond, S.P.A. (2010). Obstructive Sleep Apnea and Age. A Double Insult to Brain Function? *American Journal of Respiratory and Critical Care Medicine,* 182, 411-419.

Baillet, A., Gossec, L., Carmona, L., Wit, M. de, van Eijk-Hustings, Y., Bertheussen, H., Alison, K., Toft, M., Kouloumas, M., Ferreira, R. J. O., Oliver, S., Rubbert-Roth, A., van Assen, S., Dixon, W. G., Finckh, A., Zink, A., Kremer, J., Kvien, T. K., Nurmohamed, M., & van der Heijde, D. (2016). Points to consider for reporting, screening for and preventing selected comorbidities in chronic inflammatory rheumatic diseases in daily practice: a EULAR initiative. *Annals of the Rheumatic Diseases*, *75*(6), 965–973. https://doi.org/10.1136/annrheumdis-2016-209233

Bakker, I., Takashima, A., vanHell, J. G., Janzen, G., and McQueen, J. M. (2014). Competition from unseen or unheard novel words: Lexical consolidation across modalities. *Journal of Memory and Language*, 73, 116–130.

Bakker-Marshall, I., Takashima, A., Schoffelen, J.-M., vanHell, J. G., Janzen, G., and McQueen, J. M. (2018). Theta-band oscillations in the middle temporal gyrus reflect novel word consolidation. *Journal of Cognitive Neuroscience*, 30, 621–633.

Barsky, M.M., Tucker, M.A., and Stickgold, R. (2015). REM sleep enhancement of probabilistic classification learning is sensitive to subsequent interference. *Neurobiology of Learning and Memory,* 122, 63-68.

Basner, M., Mollicone, D., & Dinges, D. F. (2011). Validity and Sensitivity of a Brief Psychomotor Vigilance Test (PVT-B) to Total and Partial Sleep Deprivation. *Acta astronautica*, *69*(11-12), 949–959.

Becker, P. M., & Sharon, D. (2014). Mood Disorders in Restless Legs Syndrome (Willis-Ekbom Disease). *The Journal of Clinical Psychiatry*, *75*(07), e679–e694.

Benca, R. M., Okawa, M., Uchiyama, M., Ozaki, S., Nakajima, T., Shibui, K., & Obermeyer, W. H. (1997). Sleep and mood disorders. *Sleep Medicine Reviews*, *1*(1), 45–56. https://doi.org/10.1016/s1087-0792(97)90005-8

Benjafield, A.V., Ayas, N.T., Eastwood, P.R., Heinzer, R., Mary, S.M.I, Morrell, M.J., Nunez, C.M., Patel, R., Penzel, T., Pepin, J.L.D., Peppard, P.E., Sinha, S., Tufil, S., Valentine, K., and Malhorta, A. (2019). Estimation of global prevalence and burden of obstructive sleep apnoea: a literature-based analysis. *Lancet Respir Med,* 7(8), 687-698.

Berndt, C., Diekelmann, S., Alexander, N., Pustal, A., and Kirschbaum, C. (2014). Sleep fragmentation and false memories during pregnancy and motherhood. *Behavioural Brain Research,* 266, 52-57.

Berner, I, Schabus, M., Wienerroither, T., and Klimesch, W. (2006). The significance of sigma neurofeedback training on sleep spindles and aspects of declarative memory. *Appl Psychophysiol Biofeedback* 31: 97–114.

Bertrand, S.J., Zhang, Z., Patel, R., O'Ferrell, C., Punjabi, N.M., Kudchadkar, S.R., and Kannan, S. (2020). Transient neonatal sleep fragmentation results in long-term neuroinflammation and cognitive impairment in a rabbit model. *Experimental Neurology,* 327, 113212.

Betta, M., Handjaras, G., Ricciardi, E., Pietrini, P., Haba-Rubio, J., Siclari, F., Heinzer, R., and Bernardi, G. (2020). Quantifying peripheral sympathetic activations during sleep by means of an automatic method for pulse wave amplitude drop detection. *Sleep Med*, 69, 220-232.

Bianchi, M. T., Williams, K. L., Mckinney, S., & Ellenbogen, J. M. (2013). The subjective-objective mismatch in sleep perception among those with insomnia and sleep apnea. *Journal of Sleep Research*, *22*(5), 557–568. https://doi.org/10.1111/jsr.12046

Biddle, D. J., Robillard, R., Hermens, D. F., Hickie, I. B., & Glozier, N. (2015). Accuracy of self-reported sleep parameters compared with actigraphy in young people with mental ill-health. *Sleep Health*, *1*(3), 214–220. https://doi.org/10.1016/j.sleh.2015.07.006

Blandini, F., Nappi, G., Tassorelli, C., & Martignoni, E. (2000). Functional changes of the basal ganglia circuitry in Parkinson's disease. *Progress in Neurobiology*, *62*(1), 63–88. https://doi.org/10.1016/s0301-0082(99)00067-2

Bogan, R. K. (2006). Effects of restless legs syndrome (RLS) on sleep. *Neuropsychiatric Disease and Treatment*, *2*(4), 513–519. https://doi.org/10.2147/nedt.2006.2.4.513

Bonnar, J., Franklin, M., Nott, P., and Mcneilly, A. (1975). Effect of breast-feeding after childbirth. *Br. Med. J.,* 4, 82–84.

Bonnet, M.H., and Arand, D.L. (2003). Clinical effects of sleep fragmentation versus sleep deprivation. Sleep Med. Rev. 7, 297–310.

Borbély, A.A, and Achermann, P. (1999). Sleep homeostasis and models of sleep regulation. *J Biol Rhythms* 14, 557–568.

Born, J., & Wilhelm, I. (2011). System consolidation of memory during sleep. *Psychological Research*, *76*(2), 192–203. https://doi.org/10.1007/s00426-011-0335-6

Botvinick, M.M., et al. (2001). Conflict monitoring and cognitive control. *Psychol Rev.,* 108(3), 624–52.

Bradburn, N. M. (1969). *Bradburn Affect Balance Scale (ABS)* [Database record]. APA PsycTests.

Brashers-Krug, T., Shadmehr, R., and Bizzi, E. (1996). Consolidation in human motor memory. *Nature*, 382, 252.

Braver, T. S. (2012). The variable nature of cognitive control: a dual-mechanisms framework. *Trends Cogn. Sci*., 16, 106–113.

Braver, T.S, Cohen, J.D. (1999). Dopamine, cognitive control, and schizophrenia: The gating model. *Progress in Brain Research,* 121, 327–349.

Braver, T.S., and Cohen, J.D. (2000). On the control of control: The role of dopamine in regulating prefrontal function and working memory. In: Monsell, S., Driver, J., editors. (2000). *Attention and Performance XVIII*. Cambridge, MA: MIT Press, p. 713-738.

Brawn, T. P., Nusbaum, H. C., and Margoliash, D. (2013). Sleep consolidation of interfering auditory memories in starlings. *Psychol. Sci*., 24, 439–447.

Brawn, T. P., Nusbaum, H. C., and Margoliash, D. (2018). Differential development of retroactive and proactive interference during post-learning wakefulness. *Learn. Mem*., 25, 325–329.

Brewer, J., & Hunter, A. (2006). Foundations of multimethod research: Synthesizing styles. Sage Publications.

Brower, K. J., Aldrich, M. S., Robinson, E. A. R., Zucker, R. A., & Greden, J. F. (2001). Insomnia, Self-Medication, and Relapse to Alcoholism. *American Journal of Psychiatry*, *158*(3), 399–404. https://doi.org/10.1176/appi.ajp.158.3.399

Brown, E., and Schaffir, J. (2019). "Pregnancy Brain": A Review of Cognitive Changes in Pregnancy and Postpartum. *Obstetrical and Gynecological Survey,* 74(3), 178-185.

Brown, R., and Robertson, E.W. (2007). Off-Line Processing: Reciprocal Interactions between Declarative and Procedural Memories. *The Journal of Neuroscience,* 27(39): 10468 – 10475.

Brown, R.M., and Robertson, E.M. (2007). Off-line processing: Reciprocal interactions between declarative and procedural memories. *The Journal of Neuroscience*, 27(39), 10468-10475.

Bruni, O., Baumgartner, E., Sette, S., Ancona, M., Caso, G., Di Cosimo, M. E., Mannini, A., Ometto, M., Pasquini, A., Ulliana, A., & Ferri, R. (2014). Longitudinal Study of Sleep

Behavior in Normal Infants during the First Year of Life. *Journal of Clinical Sleep Medicine*. https://doi.org/10.5664/jcsm.4114

Bucks, R.S, Olaithe, M., Rosenzweig, I., and Morrell, M.J. (2017). Reviewing the relationship between OSA and cognition: where do we go from here? *Respirology*, 22(7), 1253–1261.

Bucks, R.S., Olaithe, M., and Eastwood, P. (2013). Neurocognitive function in obstructive sleep apnea: A meta-review. *Respirology*, 18, 61-70.

Buckwalter, J.G., Buckwalter, D.K., Bluestein, B.W., et al. (2001). Pregnancy and postpartum: changes in cognition and mood. *Prog Brain Res.,* 133, 303–319.

Burgess, G.C., and Braver, T.S. (2010). Neural Mechanisms of Interference Control in Working Memory: Effects of Interference Expectancy and Fluid Intelligence. *PLoS One.*, 5(9), e12861.

Buysse, D. J. (2014). Sleep health: can we define it? Does it matter? *Sleep*, 37(1), 9-17.

Buzsaki, G. (1998). Memory consolidation during sleep: a neurophysiological perspective. *J Sleep Res* 7, 17–23.

Caddick, Z.A., Gregory, K., Arsintescu, L., and Flynn-Evans, E.E. (2018). A review of the environmental parameters necessary for an optimal sleep environment. *Building and Environment,* 132, 11-20.

Cairney, S. A., Lindsay, S., Sobczak, J. M., Paller, K. A., & Gaskell, M. G. (2016). The Benefits of Targeted Memory Reactivation for Consolidation in Sleep are Contingent on Memory Accuracy and Direct Cue-Memory Associations. *Sleep*, *39*(5), 1139–1150. https://doi.org/10.5665/sleep.5772

Caligiore, D., Pezzulo, G., Baldassarre, G., Bostan, A. C., Strick, P. L., Doya, K., Helmich, R. C., Dirkx, M., Houk, J., Jörntell, H., Lago-Rodriguez, A., Galea, J. M., Miall, R. C., Popa, T., Kishore, A., Verschure, P. F. M. J., Zucca, R., & Herreros, I. (2017). Consensus Paper: Towards a Systems-Level View of Cerebellar Function: the Interplay Between

Cerebellum, Basal Ganglia, and Cortex. *The Cerebellum*, *16*(1), 203–229. https://doi.org/10.1007/s12311-016-0763-3

Camina, E., & Güell, F. (2017). The neuroanatomical, neurophysiological and psychological basis of memory: Current models and their origins. *Frontiers in Pharmacology*, *8*(1). https://doi.org/10.3389/fphar.2017.00438

Campanini, M. Z., Lopez-Garcia, E., Rodríguez-Artalejo, F., González, A. D., Andrade, S. M., & Mesas, A. E. (2017). Agreement between sleep diary and actigraphy in a highly educated Brazilian population. *Sleep Medicine*, *35*, 27–34. https://doi.org/10.1016/j.sleep.2017.04.004

Campos-Rodriguez, F., M. A., Martinez-Garcia, Martinez, M., Duran-Cantolla, de la Pena, M., Masdeu, Gonzalez, M., Campo, Inmaculada, G., Marin, J.M., Barbe, F., Montserrat, J.M., and Farre, R. (2012). Association between Obstructive Sleep Apnoea and cancer incidence in a large multicenter Spanish cohort. *American Journal of Respiratory and Critical Care Medicine*, 187(1), 99-105.

Canessa, N., Castronovo, V., Cappa, S.F., Aloia, M.S., Marelli, S., Falini, A., Alemanno, F., and Ferini-Strambi, L. (2011). Obstructive sleep apnea: Brain structural changes and neurocognitive function before and after treatment. *Am J Resp Crit Care Med*, 183, 1419-1426.

Carpenter, G. A., and Grossberg, S. (1988). The ART of adaptive pattern recognition by a self-organizing neural network. *Computer*, 21, 77–88.

Casey, P., Huntsdale, C., Angus, G., and Janes, C. (1999). Memory in pregnancy. II: Implicit, incidental, explicit, semantic, short-term, working and prospective mem- ory in primigravid, multigravid and postpartum women. *J Psychosom Obstet Gynaecol,* 158–64.

Castronovo, V., Castellano, A., Aloia, M.S., Iadanza, A., Marelli, S., Cappa, S.F., Strambi, L.F, and Falini. (2014). White matter integrity in obstructive sleep apnea before and after treatment. *Sleep*, 37(9), 1465-75.

Cha, K. S., Kim, T.-J., Jun, J.-S., Byun, J.-I., Sunwoo, J.-S., Shin, J.-W., Kim, K. H., Lee, S. K., & Jung, K.-Y. (2020). Impaired slow oscillation, sleep spindle, and slow oscillation−spindle coordination in patients with idiopathic restless legs syndrome. *Sleep Medicine*, *66*, 139–147. https://doi.org/10.1016/j.sleep.2019.09.021

Cha, K.S., Kim, T, Jun, J., Byun, J., Sunwoo, J., Shin, J., Kim, K.H., Lee, S.K., and Jung, K. (2015). Impaired slow oscillation, sleep spindle, and slow oscillation_spindle coordination in patients with idiopathic restless legs syndrome. *Sleep Medicine*, 66, 139-147.

Cespedes, E. M., Hu, F. B., Redline, S., Rosner, B., Alcantara, C., Cai, J., Hall, M. H., Loredo, J. S., Mossavar-Rahmani, Y., Ramos, A. R., Reid, K. J., Shah, N. A., Sotres-Alvarez, D., Zee, P. C., Wang, R., & Patel, S. R. (2016). Comparison of self-reported sleep duration with actigraphy: results from the hispanic community health study/study of latinos sueno ancillary study. American journal of epidemiology, 183(6), 561–573.

Chai, Y., Fang, Z., Yang, F.N., Xu, S., Deng, Y., Raine, A., Wang, J., Yu, M., Basner, M., Goel, N., Kim, J.J., Wolk, D.A., Detre, J.A., Dinges, D.F., and Rao, H. (2020). Two nights of recovery sleep restores hippocampal connectivity but not episodic memory after total sleep deprivation. *Nature,* 10, 8774.

Chen, X., Wang, R., Zee, P., Lutsey, P.L., Jahaveri, S., Alcantara, C., Jackson, C.L., Williams., M.A., and Redline, S. (2014). Racial/Ethnic Differences in Sleep Disturbances: The Multi-Ethnic Study of Atherosclerosis (MESA). *SLEEP*, 38 (6), 877-888.

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, *6*(4), 284–290. https://doi.org/10.1037/1040-3590.6.4.284

Collado-Seidel V, Kazenwadel J, Wetter TC, et al. A controlled study of additional sr-L-dopa in L-dopa-responsive restless legs syndrome with late-night symptoms. *Neurology.* 1999;52:285–90

Colzato, L. S., Zhang, W., Brandt, M. D., Stock, A.-K., & Beste, C. (2021). Cognitive profile in Restless Legs Syndrome: A signal-to-noise ratio account. *Current Research in Neurobiology*, *2*, 100021. https://doi.org/10.1016/j.crneur.2021.100021

Conley, S., Knies, A., Batten, J., Ash, G., Miner, B., Hwang, Y., Jeon, S., & Redeker, N. S. (2019). Agreement between actigraphic and polysomnographic measures of sleep in adults with and without chronic conditions: A systematic review and meta-analysis. *Sleep Medicine Reviews*, *46*, 151–160. https://doi.org/10.1016/j.smrv.2019.05.001

Conte, F., Cerasuolo, M., Fusco, G., Giganti, F., Inserra, I., Malloggi, S., Di Iorio, I., & Ficca, G. (2021). Sleep continuity, stability and organization in good and bad sleepers. *Journal of Health Psychology*, *26*(12), 2131–2142. https://doi.org/10.1177/1359105320903098

Cordi, M. J., & Rasch, B. (2021). How robust are sleep-mediated memory benefits? *Current Opinion in Neurobiology*, *67*, 1–7. https://doi.org/10.1016/j.conb.2020.06.002

Côté, C., Baril, I., Morency, C.-È., Montminy, S., Couture, M., Leblond, J., Roos, M., & Roy, J.-S. (2019). Long-Term Effects of a Multimodal Physiotherapy Program on the Severity of Somatosensory Tinnitus and Identification of Clinical Indicators Predicting Favorable Outcomes of the Program. *Journal of the American Academy of Audiology*, *30*(08), 720–730. https://doi.org/10.3766/jaaa.17147

Cousins, J. N., & Fernández, G. (2019). The impact of sleep deprivation on declarative memory. *Progress in Brain Research*, *246*, 27–53.

Cousins, J. N., Sasmita, K., & Chee, M. W. L. (2017). Memory encoding is impaired after multiple nights of partial sleep restriction. *Journal of Sleep Research*, *27*(1), 138–145. https://doi.org/10.1111/jsr.12578

Creti, I., Rizzo, D., Fichten, C., Bailes, S., Zelkowitz, and P., Libman, E. (2013). Effects of sleep disturbance in the postpartum: are new mothers an exception to the rule? *Sleep Medicine,* 14S, e93-e164.

Creti, L., Libman, E., Rizzo, D., Fichten, C. S., Bailes, S., Tran, D.-L., & Zelkowitz, P. (2017). Sleep in the Postpartum: Characteristics of First-Time, Healthy Mothers. *Sleep Disorders*, *2017*. https://doi.org/10.1155/2017/8520358

Crossley, M., Lorenzetti, F.D., Naskar, S., O'Shea, M., Kemenes, G., Benjamin, P.R., and Kemenes, I. (2019). Proactive and retroactive interference with associative memory consolidation in the snail Lymnaea is time and circuit dependent. *Communications Biology,* 2, 242, https://doi.org/10.1038/s42003-019-0470-y.

Croy, I., Smith, M. G., Gidlöf-Gunnarsson, A., & Persson-Waye, K. (2016). Optimal Questions for Sleep in Epidemiological Studies: Comparisons of Subjective and Objective Measures in Laboratory and Field Studies. *Behavioral Sleep Medicine*, *15*(6), 466–482. https://doi.org/10.1080/15402002.2016.1163700

Csabi, E., Varszegi-Schulz, M., Janacsek, K., Malecek., N., and Nemeth, D. (2014). The consolidation of implicit sequence memory in Obstructive Sleep Apnea. *PLOS ONE*, 9(10), e109010.

Cuttler, C., Graf, P., Pawluski, J.L, et al. (2011). Everyday life memory deficits in pregnant women. *Can J Exp Psychol Can Psychol expérimentale.,* 65, 27–37.

Czeisler, C.A., and Dijk, D. (1995). Use of bright light to treat maladaptation to night shift

Czeisler, C.A., and Gooley, G. (2007). Sleep and circadian rhythms in humans. *Cold Spring* Harbor Symp. *Quant. Biol*., 72, 579–597.

Danböck, & Werner. (2019). Cardiac Vagal Control, Regulatory Processes and Depressive Symptoms: Re-Investigating the Moderating Role of Sleep Quality. *International Journal of Environmental Research and Public Health*, *16*(21), 4067.

Dang-Vu, T.T., Bonjean, M., Schabus, M., et al. (2011). Interplay between spontaneous and induced brain activity during human non-rapid eye movement sleep*. Proc Natl Acad Sci*., 108, 15438e43.

Darnaudry, M., Perez-Martin, M., Del Favero, F., Gomez-Roldan, C., Segura, L.M., and Stefania, M. (2007). Early motherhood in rats is associated with a modification of hippocampal function. *Psychoneuroendocrinology*, 32(7), 803–812.

Daurat, A., Huet, N., and Tiberge, M. (2014). Memory monitoring and memory control in patients suffering from obstructive sleep apnoea syndrome. *Memory*, 22(6), 710-721.

Davis, M.H., and Gaskell, M.G. (2009). A complementary systems account of word learning: neural and behavioural evidence. *Philosophical Transactions of the Royal Society of London B: Biological Sciences,* 364, 3773-3800.

Davis, M.H., Di Betta, A.M., Macdonald, M.J., and Gaskell, M.G. (2009). Learning and consolidation of novel spoken words. *Journal of Cognition Neuroscience,* 21, 803-820.

Davis, M.H., Di Betta, A.M., MacDonald, M.J.E., and Gaskell, M.G. (2009). Learning and consolidation of novel spoken words. *Journal of Cognitive Neuroscience*, 21(4), 803-820.

De Francesco, D., Sabin, C. A., Winston, A., Rueschman, M. N., Doyle, N. D., Anderson, J., Vera, J. H., Boffito, M., Sachikonye, M., Mallon, P. W. G., Haddow, L., Post, F. A., Redline, S., & Kunisaki, K. M. (2021). Sleep health and cognitive function among people with and without HIV: the use of different machine learning approaches. *Sleep*, *44*(8). https://doi.org/10.1093/sleep/zsab035

de Vivo, L., Bellesi, M., Marshall, W., Bushong, E. A., Ellisman, M. H., Tononi, G., & Cirelli, C. (2017). Ultrastructural evidence for synaptic scaling across the wake/sleep cycle. *Science*, *355*(6324), 507–510. https://doi.org/10.1126/science.aah5982

Dempsey, J.A., Veasley, B.J., and O'Donnell, C.P. (2010). *Physiological Reviews*, 90, 47-112.

Diekelmann, S., & Born, J. (2010). The memory function of sleep. *Nature Reviews Neuroscience*, *11*(2), 114–126. https://doi.org/10.1038/nrn2762

Djonlagic, I., Guo, M., Matteis, P., Carusona, A., Stickgold, R., and Malhotra, A. (2015). First night of CPAP: impact on memory consolidation, attention and subjective experience. *Sleep Med*, 16(6), 697–702.

Djonlagic, I., Sabiosky, J., Carusona, A., Stickgold, R., and Malhorta, A. (2012). Increased sleep fragmentation leads to impaired off-line consolidation of motor memories in humans. *PLoS One*, 7 (3), e34106.

Doornbos, B., Fokkema, D.S., Molhoek, M., Tanke, M.A.C., Postema, F., and Korf, J. (2009). Abrupt rather than gradual hormonal changes induce postpartum blues-like behavior in rats. *Life Sci.,* 84, 69–74.

Driver, H. S., & Shapiro, C. M. (1992). A Longitudinal Study of Sleep Stages in Young Women During Pregnancy and Postpartum. *Sleep*, *15*(5), 449–453.

Driver, H.S., and Shapiro, C.M. (1992). A longitudinal study of sleep stages in young women during pregnancy and postpartum, *Sleep,* 15, 449–453.

Driver-Dunckley, E., Connor, D., Hentz, J., et al. (2009). No evidence for cognitive dysfunction or depression in patients with mild restless legs syndrome. *Mov Disord.,* 24, 1840-1842.

Dudai Y. (2012). The restless engram: consolidations never end. *Annu. Rev. Neurosci*., 35, 227-247

Dudai, Y., and Morris, R.G.M. (2000). To consolidate or not to consolidate: What are the questions? *Brain, Perception, Memory. Advances in Cognitive Sciences,* 149-162.

Dudai, Y., Karni, A., and Born, J. (2015). The Consolidation and Transfer of Memory. *Neuron*, 88(1), 20-32.

Dumay, N., & Gaskell, M. G. (2012). Overnight lexical consolidation revealed by speech segmentation. *Cognition*, 123, 119–132.

Dumay, N., and Gaskell, M.G. (2007). Sleep-associated changes in the mental representation of spoken words. *Psychological Science,* 18, 35-39.

Durrant, S.J., Cairney, S.A., McDermott, C., and Lewis, P.A. (2014). Schema-Conformant Memories are Preferentially Consolidated During REM Sleep. *Neurobiology of Learning and Memory*, doi: http:// dx.doi.org/10.1016/j.nlm.2015.02.011.

E. Montgomery-Downs, H., Stremler, R., & P. Insana, S. (2013). Postpartum Sleep in New Mothers and Fathers. *The Open Sleep Journal*, *6*(1), 87–97. https://doi.org/10.2174/1874620901306010087

Eckert, D. (2018). Phenotypic approaches to obstructive sleep apnoea – New pathways for targeted therapy. *Sleep Medicine Reviews*, 37, 45-59.

Ellenbogen, J. M., Hu, P.T., Payne, J.D., and Titone D., and Walker, M.P. (2007). Human relational memory requires time and sleep. *Proc. Natl Acad. Sci. USA*, 104, 7723–7728.

Engleman, H.M, Martin, S.E, Douglas, N.J. (1994). Compliance with CPAP therapy in patients with the sleep apnoea/hypopnoea syndrome. Thorax, 49, 263–6.

Eschenko, O., and Sara S.J. (2008). Learning-dependent, transient increase of activity in noradrenergic neurons of locus coeruleus during slow wave sleep in the rat: brain stem-cortex interplay for memory consolidation? *Cereb Cortex* 18: 2596–2603.

Feld, G.B., and Born, J. (2017). Sculpting memories during sleep: concurrent consolidation and forgetting. *Current Opinion in Neurobiology*, 44, 20-27.

Ficca, G., Lombardo, P., Rossi, L., & Salzarulo, P. (2000). Morning recall of verbal material depends on prior sleep organization. *Behavioural Brain Research, 112*(1-2), 159–163.

Ficca, G., Lombardo, P., Rossi, L., and Salzarulo, P. (2000). Morning recall of verbal material depends on prior sleep organization. *Behav Brain Res.,* 112, 159–63.

Finan, P. H., Quartana, P. J., & Smith, M. T. (2015). The Effects of Sleep Continuity Disruption on Positive Mood and Sleep Architecture in Healthy Adults. *Sleep*, *38*(11), 1735–1742. https://doi.org/10.5665/sleep.5154

Finan, P. H., Quartana, P. J., Remeniuk, B., Garland, E. L., Rhudy, J. L., Hand, M., Irwin, M. R., & Smith, M. T. (2016). Partial Sleep Deprivation Attenuates the Positive Affective System: Effects Across Multiple Measurement Modalities. *Sleep*, *40*(1).

Fraigne, J.J., Torontali, Z.A., Snow, M.B., and Peever, J.H. (2015). REM sleep at its core – circuits, neurotransmitters, and pathophysiology. *Frontiers in Neurology,* 6, Article 123.

Franklin, K.A., and Lindberg, E. (2018). Obstructive Sleep Apnoea is a common disorder in the population: a review on the epidemiology of sleep apnoea. *Journal of Thoracic Disease*, 7(8), 1311-1322.

Fulda, S., and Schulz, H. (2001). Cognitive dysfunction in sleep disorders. *Sleep Medicine Reviews*, 5(6), 423-445.

Gagné, M, and Cohen, H. (2016). Interference effects between memory systems in the acquisition of a skill. *Exp Brain Res,* 234, 2883-2891.

Gais, S., Mölle, M., Helms, K., and Born, J. (2002). Learning-dependent increases in sleep spindle density. *J Neurosci* 22: 6830–6834.

Galbiati, A., Marelli, S., Giora, E., Zucconi, M., Oldani, A., and Ferini- Strambi, L. (2015). Neurocognitive function in patients with idiopathic Restless Legs Syndrome before and after treatment with dopamine-agonist. *Int. J. Psychophysiol.* 95, 304–309.

Gamaldo, C.E., Benbrook, A.R, Allen, R.P., Oguntimein, O., and Earley, C.J. (2008). A further evaluation of the cognitive deficits associated with restless legs syndrome (RLS). *Sleep Med.,* 9, 500-505.

Gaskell, M.G., and Dumay, N. (2003). Lexical competition and the acquisition of novel words. *Cognition,* 89, 105–132.

Gay, C.L, *et al.* (2004). Sleep patterns and fatigue in new mothers and fathers. *Biol Res Nurs.,* 5(4), 311–318.

Gennaro, de L., Marzano, C., Fratello, F., Moroni, F., Pellicciari, M.C., Ferlazzo, F., Costa, S., Couyoumdjian, A., Curcio, G., Sforza, E., Malafosse, A., Finelli, L.A., Pasqualetti, P., Ferrara, M., Bertini, M., and Rossini, P.M. (2008). The electroencephalographic fingerprint of sleep is genetically determined: a twin study. *Ann Neurol* 64: 455–460.

Gieselmann, A., & Pietrowsky, R. (2019). The effects of brief chat-based and face-to-face psychotherapy for insomnia: a randomized waiting list controlled trial. *Sleep Medicine*, *61*, 63–72. https://doi.org/10.1016/j.sleep.2019.03.024

Gildeh, N., Drakatos, P., Higgins, S., Rosenzweig, I., Kent, B.D. (2016). Emerging co-morbidities of obstructive sleep apnea: cognition, kidney disease, and cancer. *J. Thorac. Dis*., 8: E901– 17.

Girschik, J., Fritschi, L., Heyworth, J., & Waters, F. (2012). Validation of self-reported sleep against actigraphy. Journal of Epidemiology, 22(5), 462-468.

Glynn, L.M. (2010). Giving birth to a new brain: hormone exposures of pregnancy influence human memory. *Psychoneuroendocrinology*., 35, 1148–1155.

Goedert, K., and Willingham, D. (2002). Patterns of interference in sequence learning and prism adaptation inconsistent with the consolidation hypothesis. Learn Mem, 9, 279 –292.

Goelema, M. S., Regis, M., Haakma, R., van den Heuvel, E. R., Markopoulos, P., & Overeem, S. (2017). Determinants of perceived sleep quality in normal sleepers. *Behavioral Sleep Medicine*, *17*(4), 388–397. https://doi.org/10.1080/15402002.2017.1376205

Gökce, F., Ehring, T., Werner, G. G., & Takano, K. (2020). Misperception of sleep is associated with intrinsic motivation toward thinking about sleep. *Journal of Behavior Therapy and Experimental Psychiatry*, *69*, 101591. https://doi.org/10.1016/j.jbtep.2020.101591

Goldstein, A. N., & Walker, M. P. (2014). The Role of Sleep in Emotional Brain Function. *Annual Review of Clinical Psychology*, *10*(1), 679–708. https://doi.org/10.1146/annurev-clinpsy-032813-153716

Goldstein, A., and Walker, M.P. (2014). The Role of Sleep in Emotional Brain Function. *Annul Rev Clin Psychol,* 19, 679-708.

Gottlieb, D.J., Punjabi, N.M., Mehra, R., Patel, S.R., Quan, S.F., Babineau, D.C., Tracy, R.P., Rueschman, M., Blumenthal, R.S., Lewis, E.F., Bhatt, D.L., Redline, S. (2014). CPAP versus oxygen in obstructive sleep apnea. *The New England Journal of Medicine*, 270(24), 2276-2285.

Gozal, D., Daniel, J.M., Dohanich, G.P. (2001). Behavioral and anatomical correlates of chronic episodic hypoxia during sleep in the rat. *J Neurosci* 21: 2442–2450.

Groch, S., McMakin, D., Guggenbühl, P., Rasch, B., Huber, R., and Wilhelm, I. (2016). Memory cueing during sleep modifies the interpretation of ambiguous scenes in adolescents and adults. *Developmental Cognitive Neuroscience*, 17, 10-18.

Guo, S., Huang, J., Jiang, H., Han, C., Li, J., Xu, X., Zhang, G., Lin, Z., Xiong, N., & Wang, T. (2017). Restless Legs Syndrome: From Pathophysiology to Clinical Diagnosis and Management. *Frontiers in Aging Neuroscience*, *9*. https://doi.org/10.3389/fnagi.2017.00171

Guo, S., Huang, J., Jiang, H., Han, Li, J., Xu, X., Zhang, G., Lin, Z., Ziong, N., and Wang, T. (2017). Restless Legs Syndrome: From Pathophysiology to Clinical Diagnosis and Management. *Frontiers in Aging Neuroscience,* 9, Article 171. doi:10.3389/fnagi.2017.00171.

Harvey, A. G., & Tang, N. K. (2012). (Mis)perception of sleep in insomnia: A puzzle and a resolution. *Psychological Bulletin*, 138(1), 77-101.

Hendricks, J.C., Finn, S.M., Panckeri, K.A., Chavkin, J., Williams, J.A., Sehgal, A., and Pack, A.I. (2000). Rest in Drosophila is a sleep-like state. *Neuron,* 25, 129–138.

Hening W, Allen RP, Tenzer P, Winkelman JW. Restless legs syndrome: demographics, presentation, and differential diagnosis. *Geriatrics*. 2007 Sep;62(9):26-9. PMID: 17824723.

Hening, W., Walters, A.S., Allen, R.P., Montplaisir, J., Myers, A., and Ferini-Strambi, L. (2004). Impact, diagnosis and treatment of restless legs syndrome (RLS) in a primary care population: The REST (RLS epidemiology, symptoms, and treatment) primary care study. *Sleep Med.,* 5, 237-46.

Henry, J.D., and Rendell, P.G. (2007). A review of the impact of pregnancy on memory function. *J Clin Exp Neuropsychol.,* 29, 793–803.

Hermans, L., van Gilst, M., Langen, H., van Mierlo, P., Leufkens, T., & Overeem, S. (2019). Modelling sleep onset misperception in insomnia. *Sleep Medicine*, *64*, S151–S152. https://doi.org/10.1016/j.sleep.2019.11.416

Hoang, H. T. X., Molassiotis, A., Chan, C. W., Nguyen, T. H., & Liep Nguyen, V. (2019). New-onset insomnia among cancer patients undergoing chemotherapy: prevalence, risk factors, and its correlation with other symptoms. *Sleep and Breathing*, *24*(1), 241–251. https://doi.org/10.1007/s11325-019-01839-x

Hornyak, M., Feige, B., Voderholzer, U., Philipsen, A., & Riemann, D. (2007). Polysomnography Findings in Patients with Restless Legs Syndrome and in Healthy Controls: A Comparative Observational Study. *Sleep*, *30*(7), 861–865. https://doi.org/10.1093/sleep/30.7.861

Hrubos-Strøm, H., Nordhus, I. H., Einvik, G., Randby, A., Omland, T., Sundet, K., Moum, T., & Dammen, T. (2011). Obstructive sleep apnea, verbal memory, and executive function in a community-based high-risk population identified by the Berlin Questionnaire Akershus Sleep Apnea Project. *Sleep and Breathing*, *16*(1), 223–231. https://doi.org/10.1007/s11325-011-0493-1

Huber, R., Maki, H., Rosanova, M., Casarotto, S., Canali, P., Casali, A. G., … Massimini, M. (2013). Human cortical excitability increases with time awake. *Cerebral Cortex*, *23*, 332–338.

Hunter, L.P., Rychnovsky, J.D., and Yount, S.M. (2009). A selective review of maternal sleep characteristics in the postpartum period, *J. Obstet. Gynecol. Neonatal. Nurs.,* 38, 60–68.

Iber, C., Ancoli-Israel, S., Chesson, A., & Quan, S. F. (2007). The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications. *American Academy of Sleep Medicine*.

Insana, S. P., & Montgomery-Downs, H. E. (2012). Sleep and sleepiness among first-time postpartum parents: A field- and laboratory-based multimethod assessment. *Developmental Psychobiology*, *55*(4), 361–372.

Insana, S. P., Costello, C. R., & Montgomery-Downs, H. E. (2011). Perception of Partner Sleep and Mood: Postpartum Couples' Relationship Satisfaction. *Journal of Sex & Marital Therapy*, *37*(5), 428–440. https://doi.org/10.1080/0092623x.2011.607053

Insana, S. P., Williams, K. B., & Montgomery-Downs, H. E. (2013). Sleep Disturbance and Neurobehavioral Performance among Postpartum Women. *Sleep*, *36*(1), 73–81. https://doi.org/10.5665/sleep.2304

Insana, S.P, *et al.*(2013). Sleep and sleepiness among first-time postpartum parents: a field- and laboratory-based multimethod assessment. *Dev Psychobiol.,* 55(4):361–372.

Jackson, M.L., McEvoy, D., Banks, S., Barnes, M. (2018). Neurobehavioural impairment and CPAP treatment response in mild-moderate obstructive sleep apnea. *Journal of Clinical Sleep Medicine*, 14(1), 47-56.

Jacoby, L.L., Kelley, C.M., and McElree, B.D. (1999). The role of cognitive control: Early selection versus late correction. In: Chaiken, S., Trope, E., editors. (1999). *Dual Process Theories in Social Psychology*. NY: Guildford, p. 383-400.

James, E., Gaskell, G.M., Weighall, A., and Henderson, L. (2017). Consolidation of vocabulary during sleep: the rich get richer? *Neuroscience & Biobehavioural Reviews*, 77, ISSN 0149-7634.

James, E., Gaskell, M.G., and Henderson, L.M. (2018). Offline consolidation supersedes prior knowledge benefits in children's (but not adults') word learning. *Developmental Science*, 22(3).

James, E., Gaskell, M.G., Weighall, A., and Henderson, L.M. (2017). Consolidation of vocabulary during sleep: the rich get richer? *Neuroscience & Biobehavioural Reviews,* 77(1).

Janes, C., Casey, P., Huntsdale, C., et al. (1999). Memory in pregnancy. I: Subjective experiences and objective assessment of implicit, explicit and working memory in primigravid and primiparous women. *J Psychosom Obstet Gynaecol.*, 20, 80–87.

Janků, K., Šmotek, M., Fárková, E., & Kopřivová, J. (2020). Subjective–objective sleep discrepancy in patients with insomnia during and after cognitive behavioural therapy: An actigraphy study. *Journal of Sleep Research*, *29*(4). https://doi.org/10.1111/jsr.13064

Javaheri, S., Barbe, F., Campos-Rodgriguez, F., Dempsey, J.A., Khayat, R., Javaheri, S., Malhorta, A., Martniez-Garcia, M.A., Mehra, R., Pack, A.I., Polotsly, V.Y., Redline, S., and Somers, V.K. (2017). Sleep Apnea: Types, mechanisms and clinical cardiovascular consequences. *Journal of the American College of Cardiology*, 7, 841-858.

Ji, D., Wilson, M.A. (2007). Coordinated memory replay in the visual cortex and hippocampus during sleep. *Nat Neurosci*, 10: 100–107.

Joyeux-Faure, M., Naegele, B., Pepin, J.L., Tamisier, R., Levy, P., and Launois, S.H. (2016). Continuous positive airway pressure treatment impact on memory processes in obstructive sleep apnea patients: a randomized sham-controlled trial. *Sleep Medicine*, 24, 44-50.

Jung, K.-Y. (2015). Cognition in Restless Legs Syndrome. *Journal of Sleep Medicine*, *12*(1), 1–6. https://doi.org/10.13078/jsm.15001

Kattler, Dijk, and Borbély. (1994). Effect of unilateral somatosensory stimulation prior to sleep on the sleep EEG in humans. *J Sleep Res* 3: 159–164.

Kilgore, D.S. (2010). Effects of sleep deprivation on cognition. *Prog Brain Res*, 185, 105-29.

Kawada, T., and Suzuki, S. (1995). Instantaneous change in sleep stage with noise of a passing truck, *Percept. Mot. Skills* 80(3),1031–1040.

Kempermann, G., Gast, D., Gage, F.H. (2002). Neuroplasticity in old age: sustained fivefold induction of hippocampal neurogenesis by long-term environmental enrichment. *Ann Neurol*. 52(2), 135-143.

Kenny, S, Burdayron, R, Lannes, ÉEM, Dubois-Comtois, K, Béliveau, M-J, Pennestri, M-H. Mothers' and fathers' sleep: Is there a difference between first-time and experienced parents of 6-month-olds?. J Sleep Res. 2021; 30:e13238. https://doi.org/10.1111/jsr.13238

Kheirandish-Gozal, L., De Jong, M.R., Spruyt, K., Chamuleau, S.A, and Gozal D. (2010). Obstructive sleep apnoea is associated with impaired pictorial memory task acquisition and retention in children. *Eur Respir J*, 36(1):164–169.

Kim SM, Choi JW, Lee C, et al. Working memory deficit in patients with restless legs syndrome: an event-related potential study. *Sleep Med* 2014;15:808-815.

Kim, H., Joo, E.Y., Suh, S., Kim, J.H., Kim, S.T., and Hong, S.B. (2016). Effects of long-term treatment on brain volume in patients with obstructive sleep apnea. *Human Brain Mapping*, 37, 395-409.

Kim, P. (2016). Human Maternal Brain Plasticity: Adaptation to Parenting. *New Dir Child Adolesc Dev.,* 153, 47-58.

Kim, P., and Bianco, H. (2014). How motherhood and poverty change the brain. *Zero to Three*., 34(4), 29–36.

Kim, S. (2020). Bidirectional competitive interactions between motor memory and declarative memory during interleaved learning. *Nature,* 10, 6916, https://doi.org/10.1038/s41598-020-64039-8.

Kinsley, C.H., Blair, J.C., Karp, N.E., Hester, N.W., McNamara, I.M., Orthmeyer, A.L., McSweeney, M.C., Bardi, M.M., Karelina, K., Christon, L.M., Sirkin, M.R., Victoria, L.W., Skurka, D.J., Fyfe, C.R., Hudepohl, M.B., Felicio, L.F., Franssen, R.A., Meyer, E.E., da Silva, I.S., Lambert, K.G., 2014. The mother as hunter: significant reduction in foraging costs through enhancements of predation in maternal rats. *Horm. Behav.* 66, 649–654.

Klappenbach, M., Nally, A., and Locatelli, F. F. (2017). Parallel memory traces are built after an experience containing aversive and appetitive components in the crab Neohelice. *Proc. Natl Acad. Sci*., 114, 4666–4675.

Klinzing, J. G., Niethard, N., & Born, J. (2019). Mechanisms of systems memory consolidation during sleep. *Nature Neuroscience*, *22*(10), 1598–1610. https://doi.org/10.1038/s41593-019-0467-3

Kloepfer, C., Riemann, D., Nofzinger, E.A., Feige, B., Unterrainer, J., O'Hara, Sorichter, S., Nissen, C. Memory before and after sleep in patients with moderate obstructive sleep apnea. *Journal of Clinical Sleep Medicine*, 5(6), 540-548.

Kobayashi, I., Huntley, E., Lavela, J., & Mellman, T. A. (2012). Subjectively and Objectively Measured Sleep With and Without Posttraumatic Stress Disorder and Trauma Exposure. *SLEEP*. https://doi.org/10.5665/sleep.1960

Kölling S, Wiewelhove T, Raeder C, Endler S, Ferrauti A, Meyer T, Kellmann M. Sleep monitoring of a six-day microcycle in strength and high-intensity training. Eur J Sport Sci. 2016

Aug;16(5):507-15. doi: 10.1080/17461391.2015.1041062. Epub 2015 Jun 11. PMID: 26062597.

Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Krauchi, K., and Wirz-Justice, A. (1994). Circadian rhythm of heat production, heart rate, and skin and core temperature under unmasking conditions in men. *Am. J. Physiol.*

Krause, A. J., Simon, E. B., Mander, B. A., Greer, S. M., Saletin, J. M., Goldstein-Piekarski, A. N., & Walker, M. P. (2017). The sleep-deprived human brain. *Nature Reviews Neuroscience*, *18*(7), 404–418. https://doi.org/10.1038/nrn.2017.55

Kreutzmann, J. C., Havekes, R., Abel, T., & Meerlo, P. (2015). Sleep deprivation and hippocampal vulnerability: Changes in neuronal plasticity, neurogenesis and cognitive function. *Neuroscience*, *309*, 173–190.

Kribbs N.B, Pack, A.I, Kline, L.R, Smith, P.L, Schwartz, A.R, Schubert, N.M, et al. (1993). Objective measurement of patterns of nasal CPAP use by patients with obstructive sleep apnea. *American Review of Respiratory Diseases*, 147, 887–95.

Kushida, C.A., Nichols, D.A., Holmes, T.H., et al. (2012). Effects of continuous positive airway pressure on neurocognitive function in obstructive sleep apnea patients: The Apnea Positive Pressure Long-term Efficacy Study (APPLES) [published correction appears in Sleep. 2016 Jul 1;39(7):1483], *Sleep*, 35(12), 1593-1602.

Laharna, N., Fatek, J., Zemann, M., Glos, M., Lederer, K., Suvorov, A.V., Demin, A.V., Penzel, T., Fietze, I. (2020). A sleep intervention study comparing effects of sleep restriction and fragmentation on sleep and vigilance and the need for recovery. *Physiology & Behaviour,* 215, 112794.

Landmann, N., Kuhn, M., Maier, J.G., Spiegelhalder, K., Baglioni, C., Frase, L., Riemann, D., Sterr, A., and Nissen, C. (2015). REM sleep and memory reorganization: Potential relevance for psychiatry and psychotherapy. *Neurobiol. Learn. Mem.,* 122, 28-40.

Landry, S.H., Anderson, C., Andrewartha, P., Sasse, A., Conduit, R. (2014). The impact of sleep apnea on motor skill acquisition and consolidation. *Journal of Clinical and Sleep Medicine*, 10(5), 491-496.

Leach, L., and Samuel, A. G. (2007). Lexical configuration and lexical engagement: When adults learn new words. *Cognitive Psychology*, 55, 306-353.

Lee HB, Hening WA, Allen RP, et al. Restless legs syndrome is associated with DSM-IV major depressive disorder and panic disorder in the community. *J Neuropsychiatry Clin Neurosci* 2008;20:101-105.

Lee, H.B., Ramsey, C.M., Spira, A.P., Vachon, J., Allen, R., and Munro, C.A. (2014). Comparison of cognitive functioning among individuals with treated restless legs syndrome (RLS), untreated RLS, and no RLS. *J Neuropsychiatry Clin Neurosci.,* 26, 87-91.

Lee, Y.F., Geraschenko, D., Timofeev, I., Bacskai, B.J., and Kastenenka, K.V. (2020). Slow Wave Sleep Is a Promising Intervention Target for Alzheimer's Disease. Front. Neurosci., https://doi.org/10.3389/fnins.2020.00705

Lewis, P. A., & Durrant, S. J. (2011). Overlapping memory replay during sleep builds cognitive schemata. *Trends In Cognitive Sciences*, 15(8), 343–51. doi:10.1016/j.tics.2011.06.004.

Liguori, C., Holzknecht, E., Placidi, F., Izzi, F., Mercuri, N.B., Hogl, B., and Stefani, A. (2020). Seasonality of restless legs syndrome: symptom variability in winter and summer times. *Sleep Medicine,* 66, 10-14.

Lim, A. S., Kowgier, M., Yu, L., Buchman, A. S., & Bennett, D. A. (2013). Sleep Fragmentation and the Risk of Incident Alzheimer's Disease and Cognitive Decline in Older Persons. *Sleep*, 36(7), 1027–1032.

Liu J, Tang W, Chen G, Lu Y, Feng C, Tu XM. Correlation and agreement: overview and clarification of competing concepts and measures. Shanghai Arch Psychiatry. 2016 Apr 25;28(2):115-20. doi: 10.11919/j.issn.1002-0829.216045. PMID: 27605869; PMCID: PMC5004097.

Liu, C., Meng, Z., Wiggin, T. D., Yu, J., Reed, M. L., Guo, F., Zhang, Y., Rosbash, M., & Griffith, L. C. (2019). A Serotonin-Modulated Circuit Controls Sleep Architecture to Regulate Cognitive Function Independent of Total Sleep in Drosophila. *Current Biology*, *29*(21), 3635-3646.e5. https://doi.org/10.1016/j.cub.2019.08.079

Liu, C., Meng, Z., Wiggin, T., Yu, J., Reed, M.L., Guo, F., Zhang, Y., Rosbash, M., and Griffith, L.C. (2019). A Serotonin-Modulated Circuit Controls Sleep Architecture to Regulate Cognitive Function Independent of Total Sleep in Drosophila. *Current Biology*, 29, 3635-3646.

Liu, Y., and van Hell, J.G. (2020). Learning Novel Word Meanings: An ERP Study on Lexical Consolidation in Monolingual, Inexperienced Foreign Language Learners. *Language Learning,* 70(2), 45-74.

Logan, D.M., Hill, K.R., Jones, R., Holt-Lunstad, J., and Larson, M.J. (2014). How do memory and attention change with pregnancy and childbirth? A controlled longitudinal examination of neuropsychological functioning in pregnant and postpartum women. *J Clin Exp Neuropsychol.,* 36(5), 528-39.

Lubas, M. M., Szklo-Coxe, M., Mandrell, B. N., Howell, C. R., Ness, K. K., Srivastava, D. K., Hudson, M. M., Robison, L. L., Krull, K. R., & Brinkman, T. M. (2021). Concordance between self-reported sleep and actigraphy-assessed sleep in adult survivors of childhood cancer: the impact of psychological and neurocognitive late effects. *Supportive Care in Cancer*, *30*(2), 1159–1168. https://doi.org/10.1007/s00520-021-06498-x

Luders, E., Kurth, F., Gingell, M., Engman, J., Yong, E., Poromaa, I.S., and Gaser, C. (2020). From baby brain to mommy brain: Widespread gray matter gain after giving birth. *Cortex,* 126, 334-342.

Lutz, N. D., Wolf, I., Hübner, S., Born, J., & Rauss, K. (2018). Sleep Strengthens Predictive Sequence Coding. *The Journal of Neuroscience*, *38*(42), 8989–9000.

Lyons, O.D., and Ryan, C.M. (2015). Sleep apnea and stroke. *Canadian Journal of Cardiology*, 31(7), 918-927.

Marino, M., Li, Y., Rueschman, M. N., Winkelman, J. W., Ellenbogen, J. M., Solet, J. M., ... & Buxton, O. M. (2013). Measuring sleep: accuracy, sensitivity, and specificity of wrist actigraphy compared to polysomnography. *Sleep*, 36(11), 1747-1755.

Marshall, L., & Born, J. (2007). The contribution of sleep to hippocampus-dependent memory consolidation. *Trends in Cognitive Sciences*, *11*(10), 442–450.

Marshall, L., Molle, M., Hallschmid, M., and Born, J. (2004). Transcranial direct current stimulation during sleep improves declarative memory. *J Neurosci* 24: 9985–9992.

Martin, J. L., & Hakim, A. D. (2011). Wrist Actigraphy. *Chest*, *139*(6), 1514–1527. https://doi.org/10.1378/chest.10-1872

Martin, M. (1990). On the induction of mood. Clinical Psychology Review, 10(6), 669–697. https://doi.org/10.1016/0272-7358(90)90075-L

Mattys, S. L., and Clark, J. H. (2002). Lexical activity in speech processing: Evidence from pause detection. *Journal of Memory and Language*, 47, 343-359.

McBean, A. L., & Montgomery-Downs, H. E. (2013). Timing and variability of postpartum sleep in relation to daytime performance. *Physiology & Behavior*, *122*, 134–139. https://doi.org/10.1016/j.physbeh.2013.09.003

McBean, A. L., Kinsey, S. G., & Montgomery-Downs, H. E. (2016). Effects of a single night of postpartum sleep on childless women's daytime functioning. *Physiology & Behavior*, *156*, 137–147. https://doi.org/10.1016/j.physbeh.2016.01.014

McBean, A.L., and Montgomery-Downs, H.E. (2014). What are postpartum women doing while the rest of the world is asleep? *J. Sleep Res.,* 24, 270-278.

McBean, A.L., Kinsey, S.G., Montgomery- Downs, H.E. (2016). Effects of a single night of postpartum sleep on childless women's daytime functioning. *Physiology & Behaviour,* 156, 137-147.

MCCALL, C., & MCCALL, W. V. (2011). Comparison of actigraphy with polysomnography and sleep logs in depressed insomniacs. *Journal of Sleep Research*, *21*(1), 122–127. https://doi.org/10.1111/j.1365-2869.2011.00917.x

McClelland, J. L., McNaughton, B. L., and O'Reilly, R. C. (1995). Why there are complementary learning-systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychol. Rev.,* 102, 419–457.

McGaugh, J.L. (2000). Memory–a century of consolidation. (2003). Science 287, 248–251. McDermott, C.M., LaHoste, G.J., Chen, C., Musto, A., Bazan, N.G., and Magee, J.C. Sleep deprivation causes behavioral, synaptic, and membrane excitability alterations in hippocampal neurons. *J Neurosci,* 23, 9687–9695.

McGuire, S., Muller, U., Elmenhorst, E.M., Basner, M. (2016). Inter-individual differences in the effects of aircraft noise on sleep fragmentation. Sleep, 39(5).

Medeiros, C.A.M., Carvalhedo de Bruin, P.F., Ponte e Silva, L., Coutinho, W.M., and Sales de Bruin, V.M. (2012). Sleep abnormalities and memory alterations in obstructive sleep apnea. *Sleep Science*, 5(1), 14-18.

Medina, D., Ebben, M., Milrad, S., Atkinson, B., & Krieger, A. C. (2015). Adverse Effects of Daylight Saving Time on Adolescents' Sleep and Vigilance. *Journal of Clinical Sleep Medicine*, *11*(08), 879–884. https://doi.org/10.5664/jcsm.4938

Michaelides, M., Thanos, P. K., Kim, R., Cho, J., Ananth, M., Wang, G.-J., & Volkow, N. D. (2012). PET imaging predicts future body weight and cocaine preference. *NeuroImage*, *59*(2), 1508–1513. https://doi.org/10.1016/j.neuroimage.2011.08.028

Milgrom, J., Gemmill, A.W., Bilszta, J.L., Hayes, B., Barnett, B., Brooks, J., et al., (2008). Antenatal risk factors for postnatal depression: a large prospective study. *J. Affect. Disord.,* 108, 147–157.

Miller, E.K., and Cohen, J.D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience.,* 21, 167–202.

Mishra, I., Pullum, K.B., Thayer, D.C., Morris, A.J., O'Hara, B.F., Conkright, B.W., Demas, G.E., and Ashley, N.T. (2020). Chemical sympathectomy reduces peripheral inflammatory responses to acute and chronic sleep fragmentation. *American Journal of Physiology-Regulatory Integrative and Comparative Physiology,* 318(4), 781-789.

Montgomery-Downs, H. E., Insana, S. P., Clegg-Kraynok, M. M., & Mancini, L. M. (2010). Normative longitudinal maternal sleep: the first 4 postpartum months. *American Journal of Obstetrics and Gynecology*, *203*(5), 465.e1–465.e7.

Montgomery-Downs, H.E, *et al.* (2010). Normative longitudinal maternal sleep: the first 4 postpartum months. *Am J Obstet Gynecol.,* 203, 465–e1.

Montplaisir, J., Boucher, S., Poirier, G., Lavigne, G., Lapierre, O., & Lespérance, P. (1997). Clinical, polysomnographic, and genetic characteristics of restless legs syndrome: A study of 133 patients diagnosed with new standard criteria. *Movement Disorders*, *12*(1), 61–65. https://doi.org/10.1002/mds.870120111

Moon, H. J., Chang, Y., Lee, Y. S., Song, H. J., Chang, H. W., Ku, J., et al. (2014). T2 relaxometry using 3.0-tesla magnetic resonance imaging of the brain in early- and late-onset restless legs syndrome. *J. Clin. Neurol.,*10, 197–202.

Morin, C. M., LeBlanc, M., Daley, M., Gregoire, J. P., & Mérette, C. (2006). Epidemiology of insomnia: prevalence, self-help treatments, consultations, and determinants of help-seeking behaviors. *Sleep Medicine*, 7(2), 123-130.

Morris, R. G. (2007). *Schemas and memory consolidation. Science*, 316, 76-82

Muzet, A. (2007). Environmental noise, sleep and health. *Sleep Med. Rev.* (11), 135-142.

Natarajan, R. (2010). Review of periodic limb movement and restless leg syndrome. *J Postgrad Med*, 56(2), 157-162.

Natarajan, R. (2010). Review of periodic limb movement and restless leg syndrome. *Journal of Postgraduate Medicine*, *56*(2), 157. https://doi.org/10.4103/0022-3859.65284

Navarette-Opazo, A., and Mitchell, D.S. (2014). Therapeutic potential of intermittent hypoxia: a matter of dose. *American Journal of Physiology,* 307(10), R1181–R1197.

Newbury, C. R., Crowley, R., Rastle, K., & Tamminen, J. (2021). Sleep deprivation and memory: Meta-analytic reviews of studies on sleep deprivation before and after learning. *Psychological Bulletin*, *147*(11), 1215–1240.

Nishihara, K., Horiuchi, S., Eto, H., Uchida, S., Honda, M. (2004). Delta and theta power spectra of night sleep EEG are higher in breast-feeding mothers than in non-pregnant women. *Neurosci. Lett.,* 368, 216–220.

O'Brien, E. M., Fava, J., Subak, L. L., Stone, K., Hart, C. N., Demos, K., & Wing, R. (2012). Sleep duration and weight loss among overweight/obese women enrolled in a behavioral weight loss program. *Nutrition & Diabetes*, *2*(9), e43–e43. https://doi.org/10.1038/nutd.2012.17

O’Regan, D., & Anderson, K. N. (2020). Restless legs syndrome and periodic limb movements of sleep. *British Journal of Hospital Medicine*, *81*(1), 1–8.

O’Regan, D., and Anderson, K.N. (2020). Restless legs syndrome and periodic limb movements of sleep. *British Journal of Hospital Medicine,* https://doi.org/10.12968/hmed.2019.0319

O’Sullivan, C.E., Issa, F.G., Berthon-Jones, M., and Eves, L. (1981). Reversal of obstructive sleep apnoea by continuous positive airway pressure applied through the nares. *Lancet, 1(8225)*, 862-865.

Ohayon, M. M., Carskadon, M. A., Guilleminault, C., & Vitiello, M. V. (2004). Meta-Analysis of Quantitative Sleep Parameters From Childhood to Old Age in Healthy Individuals: Developing Normative Sleep Values Across the Human Lifespan. *Sleep*, *27*(7), 1255–1273. https://doi.org/10.1093/sleep/27.7.1255

Ohayon, M. M., O’hara, R., and Vitiello, M. V. (2012). Epidemiology of restless legs syndrome: a synthesis of the literature. *Sleep Med. Rev.,* 16, 283–295.

Ohayon, M., Wickwire, E. M., Hirshkowitz, M., Albert, S. M., Avidan, A., Daly, F. J., ... & Vitiello, M. V. (2017). National Sleep Foundation's sleep quality recommendations: first report. *Sleep Health*, 3(1), 6-19.

Ondo, W. G., & Verma, A. (2001). Physiological assessment of paroxysmal dystonia secondary to subacute sclerosing panencephalitis. *Movement Disorders*, *17*(1), 154–157. https://doi.org/10.1002/mds.10005

Parra, O., Sánchez-Armengol, Á., Capote, F., et al. (2015). Efficacy of continuous positive airway pressure treatment on 5-year survival in patients with ischaemic stroke and obstructive sleep apnea: a randomized controlled trial [published correction appears in J Sleep Res. 2015 Aug;24(4):474].

Pawluski, J.L., Lambert, K.G., and Kinsley, C.H. (2016). Neuroplasticity in the maternal hippocampus: relation to cognition and effects of repeated stress. *Horm Behav.,* 77, 86–97.

Pearson, V.E., Allen, R.P., Dean, T., Gamaldo, C.E., Lesage, S.R., and Earley, C.J. (2006). Cognitive deficits associated with restless legs syndrome (RLS). *Sleep Med.,* 7, 25-30.

Peppard, P. E., Young, T., Barnet, J. H., Palta, M., Hagen, E. W., & Hla, K. M. (2013). Increased prevalence of sleep-disordered breathing in adults. *American Journal of Epidemiology*, 177(9), 1006-1014.

Pereda, A.E. (2014). Electrical Synapses and Their Functional Interactions with Chemical Synapses. Nature Reviews Neuroscience 15, 250-263.

Peter-Derex, L., Subtil, F., Lemaitre, G., Ricordeau, F., Bastuji, H., Bridoux, A., Onen, F., & Onen, S-Hakki. (2020). Observation and Interview-based Diurnal Sleepiness Inventory for measurement of sleepiness in patients referred for narcolepsy or idiopathic hypersomnia. *Journal of Clinical Sleep Medicine*, *16*(9), 1507–1515.

Plihal, W., & Born, J. (1999). Effects of early and late nocturnal sleep on priming and spatial memory. *Psychophysiology*, *36*(5), 571–582. https://doi.org/10.1111/1469-8986.3650571

Plihal, W., and Born, J. (1997). Effects of early and late nocturnal sleep on declarative and procedural memory. *J Cogn Neurosci*, 534–547.

Plihal, W., and Born, J. (1999). Effects of early and late nocturnal sleep on priming and spatial memory. *Psychophysiology* 36, 571–582.

Quin, N., Tikotzky, L., Astbury, L., Stafford, L., Fisher, J., Wiley, J., & Bei, B. (2022). 0474 Three-arm randomised controlled trial of Cognitive Behavioural Therapy for Insomnia, a responsive bassinet, and sleep hygiene for preventing postpartum insomnia: Preliminary findings on maternal insomnia and sleep outcomes (Study for Mother-Infant Sleep). *Sleep*, *45*(Supplement_1), A210–A210.

Rakowska, M., Abdellahi, M. E. A., Bagrowska, P., Navarrete, M., & Lewis, P. A. (2021). Long term effects of cueing procedural memory reactivation during NREM sleep. *NeuroImage*, *244*, 118573. https://doi.org/10.1016/j.neuroimage.2021.118573

Rasch, B., and Born, J. (2013). About Sleep's Role in Memory. *Physiological Review,* 93, 681-766.

Raven, F., Heckman, P.R.A., Havekes, R, and Meerlo, P. (2019). Sleep deprivation-induced impairment of memory consolidation is not mediated by glucocorticoid stress hormones. *Journal of Sleep Research,* DOI: 10.1111/jsr.12972.

Ray, L. B., Sergeeva, V., Viczko, J., Owen, A. M., & Fogel, S. M. (2017). Sleep-dependent motor sequence memory consolidation in individuals with periodic limb movements. *Sleep Medicine*, *40*, e276. https://doi.org/10.1016/j.sleep.2017.11.810

Reeves-Hoche, M.K, Meckm R., and Zwillich, C.W. (1994). Nasal CPAP: An objective evaluation of patient compliance. *American Journal of Respiratory & Critical Care Medicine*, 149, 149–54. *Regul. Integr. Comp. Physiol.* 267(3), 819–829.

Reid, M. J., Espie, C. A., Omlin, X., & Kyle, S. D. (2019). One night of sleep continuity disruption engenders impairment in overnight memory consolidation and vigilance but not emotion perception or regulation: results from a randomised, lab-based experimental study. *Sleep Medicine*, *64*, S315. https://doi.org/10.1016/j.sleep.2019.11.883

Rendell, P.G., and Henry, J.D. (2008). Prospective-memory functioning is affected during pregnancy and postpartum. *J Clin Exp Neuropsychol.,* 30, 913–919.

Reutrakul, S., and Van Cauter, E. (2014). Interactions between sleep, circadian function, and glucose metabolism: implications for risk and severity of diabetes. *Annals of the New York Academy of Sciences,* 1311, 151-173.

Ribeiro, S. (2012). Sleep and plasticity. *Pflugers Arch.,* 463, 111-120.

Ribeiro, S., Shi, X., Engelhard, M., Zhou, Y., Zhang, H., Gervasoni, D., Lin, S.C., Wada, K., Lemos, N.A., and Nicolelis, M.A. (2007). Novel experience induces persistent sleep-dependent plasticity in the cortex but not in the hippocampus. *Front. Neurosci.,* 1, 43-55.

Richter, D.R., Kramer, M.D., Tang, N.K.Y., Montgomery-Downs, H.E., and Lemola, S. (2019). Long-term effects of pregnancy and childbirth on sleep satisfaction and duration of first-time and experienced mothers and fathers. *SleepJ,* 1-10.

Rist, P.M., Elbaz, A., Dufouil, C., Tzourio, C., and Kurth, T. (2015). Restless Legs Syndrome and Cognitive Function: A Population-based Cross-Sectional Study. *Am J Med* 2015, http://dx.doi.org/10.1016/j.amjmed.2015.04.011.

Roberts, D. M., Schade, M. M., Mathew, G. M., Gartenberg, D., & Buxton, O. M. (2020). Detecting Sleep Using Heart Rate and Motion Data from Multisensor Consumer-Grade Wearables, Relative to Wrist Actigraphy and Polysomnography. *Sleep*.

Robin, E.D., Whaley, R.D., Crump, C.H., and Travis, D.M. (1958). Alveolar gas tensions, pulmonary ventilation and blood pH during physiologic sleep in normal subjects, *J.Clin. Invest*., 37(7), 981.

Rosenberg, E. L. (1998). Levels of Analysis and the Organization of Affect. *Review of General Psychology*, *2*(3), 247–270. https://doi.org/10.1037/1089-2680.2.3.247

Rotenberg, B. W., Vicini, C., Pang, E. B., and Pang, K. P. (2016). Reconsidering first-line treatment for obstructive sleep apnea: a systematic review of the literature, *Journal of Otolaryngology - Head and Neck Surgery*, 45(3), 23-67.

Rozenweig, I., Glasser, M., Crum, W.R., Kempton, M.J., Milosevic, M., McMilan, A., Leschziner, G.D., Kumari, V., Goadsby, P., Simonds, A.K., Williams, S.C.R., and Morrell, M.J. (2016). Changes in the neurocognitive architecture in patients with obstructive sleep apnea treated with continuous positive airways pressure. *EBioMedicine*, 7, 221-229.

Rudoy, J. D., Voss, J. L., Westerberg, C. E., & Paller, K. A. (2009). Strengthening Individual Memories by Reactivating Them During Sleep. *Science*, *326*(5956), 1079–1079. https://doi.org/10.1126/science.1179013

Rundo, J. V., & Downey, R. (2019). Polysomnography. *Handbook of Clinical Neurology*, *160*, 381–392. https://doi.org/10.1016/B978-0-444-64032-1.00025-4

Sadeh, A. (2011). The role and validity of actigraphy in sleepmedicine: An update. *Sleep Medicine Reviews*, 15(4), 259-267.

Saletu M, Anderer P, Saletu-Zyhlarz G, et al. Acute placebo-controlled sleep laboratory studies and clinical follow-up with pramipexole in restless legs syndrome. *Eur Arch Psychiatry Clin Neurosci.* 2002b;252:185–94.

Sawyer, A. M., Gooneratne, N., Marcus, C. L., Ofer, D., Richards, K. C., and Weaver, T. E. (2011). A Systematic Review of CPAP Adherence Across Age Groups: Clinical and Empiric Insights for Developing CPAP Adherence Interventions. *Sleep Medicine Review*, 15(6), 343-356.

Schabus, M., Dang-Vu, T.T., Heib, D.P.J., et al. (2012). The fate of incoming stimuli during NREM sleep is determined by spindles and the phase of the slow oscillation. *Front Neurol.,* 3(40).

Sharkey, K.M., Pearlstein, T.B., and Carskadon, M.A. (2013). Circadian phase shifts and mood across the perinatal period in women with a history of major depressive disorder: a preliminary communication, *J. Affect. Disord.,* 150, 1103–1108.

Shaw, P.J., Cirelli, C., Greenspan, R.J., and Tononi, G. (2000). Correlates of sleep and waking in Drosophila melanogaster. *Science,* 287, 1834–1837.

Shen, L., van Schie, J., Ditchburn, G., Brook, L., & Bei, B. (2018). Positive and Negative Emotions: Differential Associations with Sleep Duration and Quality in Adolescents. *Journal of Youth and Adolescence*, *47*(12), 2584–2595. https://doi.org/10.1007/s10964-018-0899-1

Sivertsen, B., *et al.* (2015). Trajectories of maternal sleep problems before and after childbirth: a longitudinal population-based study. *BMC Pregnancy Childbirth.,* 15, 129.

Sleep Apnoea Trust (2017). *Patient Information Sheet*. [Online]. Available at: http://www.sleep-apnoea-trust.org/wp-content/uploads/2016/01/patient-information-sheet.pdf [Accessed 19 October 2017].

Slightam, C., Petrowski, K., Jamison, A. L., Keller, M., Bertram, F., Kim, S., & Roth, W. T. (2017). Assessing sleep quality using self-report and actigraphy in PTSD. *Journal of Sleep Research*, *27*(3), e12632. https://doi.org/10.1111/jsr.12632

Smith, M.G., Croy, I., Orgen, M., Waye, K.P. (2013). On the Influence of Freight Trains on Humans: A Laboratory Investigation of the Impact of Nocturnal Low Frequency Vibration and Noise on Sleep and Heart Rate. *PLOSone,* 8(2), e55829.

Squire, L., R. (1992). Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans. *Psychol. Rev*., 99, 195–231.

Srivanitchapoom, P., Pandey, S., and Hallett, M. (2014). Restless legs syndrome and pregnancy: a review. *Parkinsonism Relat. Disord*. 20, 716–722.

Staresina, B. P., Bergmann, T. O., Bonnefond, M., van der Meij, R., Jensen, O., Deuker, L., Elger, C. E., Axmacher, N., & Fell, J. (2015). Hierarchical nesting of slow oscillations, spindles and ripples in the human hippocampus during sleep. *Nature Neuroscience*, *18*(11), 1679–1686. https://doi.org/10.1038/nn.4119

Stickgold, R., and Walker, M.P. (2013). Sleep-dependent memory triage: Evolving generalization through selective processing. *Nat Neuroscience,* 16(2), 139-145.

Takashima, A., Bakker, I., vanHell, J. G., Janzen, G., and McQueen, J. M. (2014). Richness of information about novel words influences how episodic and semantic memory networks interact during lexicalization. *NeuroImage*, 84, 265–278.

Takashima, A., Bakker, I., vanHell, J. G., Janzen, G., and McQueen, J. M. (2017). Interaction between episodic and semantic memory networks in the acquisition and consolidation of novel spoken words. *Brain and Language*, 167, 44–60.

Tamminen, J., Payne, J. D., Stickgold, R., Wamsley, E. J., and Gaskell, M. G. (2010). Sleep spindle activity is associated with the integration of new memories and existing knowledge. *Journal of Neuroscience, 30*, 14356–14360.

Tartar, J. L., Ward, C. P., Cordeira, J. W., Legare, S. L., Blanchette, A. J., McCarley, R. W., & Strecker, R. E. (2009). Experimental sleep fragmentation and sleep deprivation in rats increases exploration in an open field test of anxiety while increasing plasma corticosterone levels. *Behavioural Brain Research*, *197*(2), 450–453.

Tassi, P., Rohmer, O., Bonnefond, A., Margiocchio, F., Poisson, F., and Schimchowitsch, S. (2013). Long term exposure to nocturnal railway noise produces chronic signs of cognitive deficits and diurnal sleepiness. *Journal of Environmental Psychology,* 33, 45-52.

Thomas, K.A., and Burr, R.L. (2006). Melatonin level and pattern in postpartum versus non-pregnant nulliparous women. *J. Obstet. Gynecol. Neonatal. Nurs.,* 35, 608–615.

Thurman, S. M., Wasylyshyn, N., Roy, H., Lieberman, G., Garcia, J. O., Asturias, A., Okafor, G. N., Elliott, J. C., Giesbrecht, B., Grafton, S. T., Mednick, S. C., & Vettel, J. M. (2018). Individual differences in compliance and agreement for sleep logs and wrist actigraphy: A longitudinal study of naturalistic sleep in healthy adults. *PLOS ONE*, *13*(1), e0191883. https://doi.org/10.1371/journal.pone.0191883

Tikkinen, K. A. O., Johnson, T. M., Tammela, T. L. J., Sintonen, H., Haukka, J., Huhtala, H., ... & Auvinen, A. (2010). Nocturia frequency, bother, and quality of life: how often is too often? A population-based study in Finland. *European Urology*, 57(3), 488-496.

Tikotzky L. Postpartum maternal sleep, maternal depressive symptoms and self-perceived mother-infant emotional relationship. *Behavioral Sleep Medicine.* 2016;14(1):5–22. doi: 10.1080/15402002.2014.940111.

Tison, F., Crochard, A., Leger, D., Bouee, S., Lainey, E., and El Hasnaoui, A. (2005). Epidemiology of restless legs syndrome in French adults: a nationwide survey: the INSTANT Study. *Neurology* 65, 239–246. doi: 10.1212/01.wnl.0000168910. 48309.4a

Tomaso, C. C., Johnson, A. B., & Nelson, T. D. (2020). The Effect of Sleep Deprivation and Restriction on Mood, Emotion, and Emotion Regulation: Three Meta-Analyses in One. *Sleep*, *44*(6). https://doi.org/10.1093/sleep/zsaa289

Tononi, G., & Cirelli, C. (2006). Sleep function and synaptic homeostasis. *Sleep Medicine Reviews*, *10*(1), 49–62.

Tononi, G., & Cirelli, C. (2014). Sleep and the Price of Plasticity: From Synaptic and Cellular Homeostasis to Memory Consolidation and Integration. *Neuron*, *81*(1), 12–34. https://doi.org/10.1016/j.neuron.2013.12.025

Tononi, G., Cirelli, C. (2018). Sleep and synaptic down-selection. *Eur J Neurosci,* 51, 413-421.

Trenkwalder C, Stiasny K, Pollmacher T, et al. L-dopa therapy of uremic and idiopathic restless legs syndrome:a double-blind, crossover trial. *Sleep.* 1995;18:681–8.

Trenkwalder, C., and Paulus, W. (2010). Restless legs syndrome: pathophysiology, clinical presentation and management. *Nat. Rev. Neurol.,* 6, 337–346.

Trenkwalder, C., Tinelli, M., Sakkas, G. K., Dauvilliers, Y., Ferri, R., Rijsman, R., Oertel, W., & Jaarsma, J. (2020). Socioeconomic impact of restless legs syndrome and inadequate restless legs syndrome management across European settings. *European Journal of Neurology*, *28*(2), 691–706. https://doi.org/10.1111/ene.14582

Tse, D., Langston, R. F., Kakeyama, M., Bethus, I., Spooner, P. A., Wood, E. R., ... & Morris, R. G. (2007). Schemas and memory consolidation. Science, 316(5821), 76-82.Tulchinsky,

D., Hobel, C.J., Yeager, E., and Marshall, J.R. (1975). Plasma estrone, estradiol, estriol, progesterone, adn 17-hydroxyprogesterone in human pregnancy. *Am. J. Obstet. Gynecol.*, 112, 1095–1100.

Van Someren, E.J.W., Cirelli, C., Dijk, D, Van Cauter, E., Schwartz, S., Chee, M.W.L. (2015). Disrupted Sleep: From Molecules to Cognition. *J Neuroscience,* 35(41), 13889-95.

Vitiello, M. V., Larsen, L. H., & Moe, K. E. (2004). Age-related sleep change: Gender and estrogen effects on the subjective-objective sleep quality relationships of healthy, noncomplaining older men and women. *Journal of Psychosomatic Research*, 56(5), 503-510.

Volk, D.W., and Lewis, D,A, (2005). GABA Targets for the Treatment of Cognitive Dysfunction in Schizophrenia. *Curr Neuropharacol,* 3(1), 45-62.

Vyazovskiy, V., Borbély, A.A., and Tobler, I. (2000). Unilateral vibrissae stimulation during waking induces interhemispheric EEG asymmetry during subsequent sleep in the rat. *J Sleep Res,* 9, 367–371.

Walker, M.P., Brakefield, T., Hobson, J.A, and Stickgold, R. (2003). Dissociable stages of human memory consolidation and reconsolidation. *Nature*, 425, 616–620.

Wallace, A., and Bucks, R. S. (2013). Memory and Obstructive Sleep Apnea: A Meta-Analysis. *SLEEP*, 36(2), 203-220.

Wallace, E., Kim, D. Y., Kim, K.-M., Chen, S., Blair Braden, B., Williams, J., Jasso, K., Garcia, A., Rho, J. M., Bimonte-Nelson, H., & Maganti, R. (2015). Differential effects of duration of sleep fragmentation on spatial learning and synaptic plasticity in pubertal mice. *Brain Research*, *1615*, 116–128. https://doi.org/10.1016/j.brainres.2015.04.037

Wallace, E.W., Kim, D.Y., Chen, S., Braden, B.B., Williams, J., Jasso, K., Garcia, A., Rho, J.M., Bimonte-Nelson, Maganti, R. (2015). Differential effects of duration of sleep fragmentation on spatial learning and synaptic plasticity in pubertal mice. *Brain Research,* 1615, 116-128.

Walters, A. S., Hickey, K., Maltzman, J., Verrico, T., Joseph, D., Hening, W., et al. (1996). A questionnaire study of 138 patients with restless legs syndrome: the 'Night-Walkers' survey. *Neurology*, 46, 92–95.

Wang, S., and Morris, R.G.M. (2010). Hippocampal-neocortical interactions in memory formation, consolidation, and reconsolidation. *Annu Rev Psychol*, 61, 49–79.

Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin*, *98*(2), 219–235. https://doi.org/10.1037/0033-2909.98.2.219

Weaver, T.E., Laizner, A.M., Evans, L.K., et al. (1997). An instrument to measure functional status outcomes for disorders of excessive sleepiness. *Sleep*, 20, 835-43.

West, G., Vadillo, M.A., Shanks, D.R., and Hulme, C. (2017). The procedural learning deficit hypothesis of language learning disorders: we see some problems. *Developmental Science,* 2017, e12552.

Wichniak, A., Wierzbicka, A., Walecka, M., and Jernajczyk, W. (2017). Effect of Antidepressants on Sleep. *Curr Psychiatry Rep,* 19(63), 62-69.

Wiggin, T.D., Goodwin, P.R., Donelson, N.C., Liu, C., Trinh, K., Sanyal, S., and Griffith, L.C. (2019). Sleep structure and quantity are determined by behavioral transition probability in Drosophila melanogaster. *bioRxiv.* https://doi.org/10.1101/263301

Wilhelm, I., Diekelmann, S., and Born, J. (2008). Sleep in children improves memory performance on declarative but not procedural tasks. *Learning & Memory*, 15, 373-377.

Wilhelm, I., Diekelmann, S., Molzow, I., Ayoub, A., Molle, M., & Born, J. (2011). Sleep Selectively Enhances Memory Expected to Be of Future Relevance. *Journal of Neuroscience*, *31*(5), 1563–1569. https://doi.org/10.1523/jneurosci.3575-10.2011

Wilhelm, I., Prehn-Kristensen, A., and Born, J. (2012). Sleep-dependent memory consolidation–What can be learnt from children? *Neuroscience & Biobehavioral Reviews*, 36, 1718-1728.

Wimmer, R.D., Astori, S., Bond, C.T., et al. (2012). Sustaining sleep spindles through enhanced SK2-channel activity consolidates sleep and elevates arousal threshold. *J Neurosci.,* 32, 13917e28.

Winkelman J, Prager M, Lieb R, et al. “Anxietas Tibiarum” Depression and anxiety disorders in patients with restless legs syndrome. *J Neurol*, 2005;25:67 71

Winkelman, J.W., Redline, S., Baldwin, C.M., Resnick, H.E., Newman, A.B., and Gottlieb, D.J. (2009). Polysomnographic and health-related quality of life correlates of restless legs syndrome in the Sleep Heart Health Study. *Sleep,* 32, 772-778.

Wixted, J. T. (2004). The psychology and neuroscience of forgetting. *Ann. Rev. Psychol.*, 55, 235–269.

Wong, M. M., Brower, K. J., Fitzgerald, H. E., & Zucker, R. A. (2004). Sleep Problems in Early Childhood and Early Onset of Alcohol and Other Drug Use in Adolescence. *Alcoholism: Clinical & Experimental Research*, *28*(4), 578–587.

Xie, T., Wang, T., Wei, T, Ye, E. (2019). Declarative memory affects procedural memory: The role of semantic association and sequence matching. *Psychology of Sport & Exercise,* 43, 253-260.

Xu, W., Chi, L., Row, B.W., Xu, R., Ke, Y., Xu, B., Luo, C., Kheirandish, L., Gozal, D., and Liu, R. (2014). Increased oxidative stress is associated with chronic intermittent hypoxia-mediated brain cortical neuronal cell apoptosis in a mouse model of sleep apnea. *Neuroscience*, 126: 313– 23.

Yin, C., and Wei, K. (2014). Interference from mere thinking: mental rehearsal temporarily disrupts recall of motor memory. *J Neurophysiol.,* 112, 594-602.

Yonelinas, A. P., Ranganath, C., Ekstrom, A. D., & Wiltgen, B. J. (2019). A contextual binding theory of episodic memory: Systems consolidation reconsidered. Nature Reviews Neuroscience, 20(6), 364–375. https://doi.org/10.1038/s41583-019-0150-4

Yoo, S.S., Hu, P.T., Gujar, N., Jolesz, F.A., and Walker, M.P. (2007). A deficit in the ability to form new human memories without sleep. *Nature,* 10(3), doi:10.1038/nn1851.

Zimmerman, J.E., Chan, M.T., Lenz, O.T., Keenan, B.T., Maislin, G., and Pack, A.I. (2017). Glutamate Is a Wake-Active Neurotransmitter in Drosophila melanogaster. *SLEEP,* 40(2), zsw046.

Zimmerman, M.E, Arnedt, J.T, Stanchina, M., Millman, R.P, Aloia, M.S. (2006). Normalization of memory performance and positive airway pressure adherence in memory-impaired patients with obstructive sleep apnea. *Chest*, 130(6), 1772–8