Testing a Hypothesis of Non-REM Sleep Reinforcement and REM Sleep Refinement for the Benefits of Post-Learning Sleep on Memory Retrieval

Kevin John MacDonald

Psychology

Submitted in partial fulfillment
of the requirements for the degree of

Doctor of Philosophy

Faculty Social Sciences, Brock University
St. Catharines, Ontario

Abstract

It is well established that post-learning sleep benefits later memory retrieval, but there is still much to learn about the processes involved and the nature of these benefits. Sleep is composed of stages of non-REM (NREM) and REM sleep: NREM sleep, especially slow wave activity of NREM sleep, and REM sleep have been implicated in memory performance benefits, but the specific contributions of each state remain unclear. This thesis presents a hypothesis proposing that post-learning NREM sleep supports memory accessibility, benefitting the likelihood of successful memory retrieval, and that post-learning REM sleep supports memory fidelity, allowing for more accurate retrieval when retrieval is successful. This hypothesis was tested over studies examining the effects of an afternoon nap (Chapter 2), targeted memory reactivation during NREM slow wave sleep (Chapter 3), and both targeted memory reactivation during NREM slow wave sleep and selective deprivation of REM sleep (Chapter 4) on measures of memory accessibility and memory fidelity in visuospatial memory tasks. In each study, measures of sleep architecture and electroencephalographic power in sleep were examined as predictors of memory performance. Several identified associations and interactions further inform an understanding of how NREM sleep and REM sleep may benefit memory performance. Most notably, these studies consistently found greater slow wave activity of NREM sleep to be specifically associated with better maintenance of memory accessibility. These studies did not identify a clear effect of REM sleep. It is hoped that the hypothesis and findings presented stimulate additional inquires that will further our understanding of the individual and combined contributions of NREM and REM sleep.

Keywords: sleep, memory, NREM sleep, REM sleep, napping, targeted memory reactivation, REM sleep deprivation

Acknowledgements

Most importantly, I thank my friends and family for their understanding and encouragement through this process. I thank my advisor, Dr. Kimberly Cote, and my committee members, Dr. Karen Arnell and Dr. Stephen Emrich, for their guidance along the way. I am also thankful for all the support I have received from research assistants, the faculty and staff of Brock University, my fellow scientists at Brock University and beyond, and many helpful strangers of the internet.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

EEG (electroencephalography): electrophysiological measure of brain activity.

EMG (electromyography): electrophysiological measure of muscle tone.

EOG (electrooculography): electrophysiological measure of eye movements.

IQR (interquartile range): descriptive statistic referring to the numerical distance between the first quartile (25^{th} percentile) and the third quartile (75^{th} percentile). Here it is used as an indication of extremeness of a value for a variable in relation to the distribution of values for that variable; for example, a value 1.5 times the IQR above the mean is described as +1.5 IQR.

LTP (long-term potentiation): neurophysiological mechanism of memory formation characterized by a relatively long-lasting increase in synaptic strength.

Stage N1 (non-rapid eye movement sleep stage 1): one of three human sleep stages without rapid eye movements. Stage N1 sleep is the lightest of these three sleep stages.

Stage N2 (non-rapid eye movement sleep stage 2): one of three human sleep stages without rapid eye movements. Stage N2 sleep is of an intermediate depth relative to the two sleep stages without rapid eye movements.

Stage N3 (non-rapid eye movement sleep stage 3): one of three human sleep stages without rapid eye movements. Stage N3 sleep is the deepest of these three sleep stages.

Stage R (rapid eye movement sleep stage): stage of human sleep characterized by phasic bursts of rapid eye movements and low muscle tone. Stage R sleep is analogous to paradoxical sleep in other animals and is referred to more generally as rapid eye movement sleep.

NREMI (non-rapid eye movement sleep interruptions): an experimental condition in which participants are awakened during non-rapid eye movement sleep. Sleep interruptions in the NREMI condition are matched to those that occur during selective deprivation of rapid eye movement sleep.

NREM (non-rapid-eye-movement): a classification of sleep characterized by an absence of rapid eye movements. In humans, NREM sleep includes sleep stages N1, N2, and N3.

PSG (polysomnography): implementation of multiple electrophysiological measures of sleep, including EEG, EMG, and EOG.

REM (rapid eye movement): saccades observed in mammal and bird sleep. Phasic bursts of rapid eye movements are used in the classification of sleep stages, specifically the identification of REM/stage R sleep.

REMD (selective rapid eye movement sleep deprivation): a class of experimental protocols in which research subjects or participants are selectively deprived of REM sleep through experimental awakenings at the onset of REM sleep.

TMR (targeted memory reactivation): an experimental technique in which specific

memories are thought to be reactivated within a subject, typically during sleep, through re-exposure to a cue paired with those memories during encoding.

VA (visual angle): a unit of measurement of stimulus size referring to the estimated length of the arc on the retina subtended by the image at a given distance. Here, visual angle is reported in degrees.

Chapter 1

General Introduction

There is wide, longstanding support for the notion that sleep can benefit newly acquired memories (Rasch & Born, 2013). Memory performance benefits from sleep relative to wake have been observed across many paradigms and memory tasks. For example, sleep protects newly learned word or syllable pairs from interference caused by future learning (i.e., retroactive interference; Drosopoulos, et al., 2007; Ellenbogen et al., 2006) and promotes performance gains in newly learned skills such as finger tapping sequences (Fischer et al., 2002; Walker et al., 2003; Walker et al., 2002). However, the relationship between sleep and memory is more complex than sleep simply preserving or strengthening memories. Sleep is thought to reorganize memories, facilitating qualitative changes to memories (Landmann et al., 2014), and may even promote forgetting of some memories (Crick & Mitchison, 1983; Poe, 2017). This complexity in the relationship between sleep and memory may in part be attributed to the fact that sleep itself is complex in its composition. This introduction will outline sleep physiology and existing hypotheses regarding sleep and memory function before introducing a new hypothesis regarding the effect of sleep states on memory performance. Then, literature connecting sleep to memory will be discussed. Finally, this introduction will outline how this hypothesis will be tested over three studies reported in subsequent chapters.

Sleep Physiology

Sleep Architecture

In humans, sleep is composed of distinct stages physiologically defined through polysomnography (PSG) by brain activity measured via electroencephalography (EEG), eye movements measured via electrooculography (EOG), and muscle tone measured via

electromyography (EMG). The criteria used to score sleep stages and events were defined by Rechtschaffen and Kales (1968) and updated and redefined by the American Academy of Sleep Medicine (Berry et al., 2015). The stages defined by the American Academy of Sleep Medicine are described here. In different contexts and reports, the exact frequency boundaries for some features (e.g., spindles) often vary slightly. Rapid eye movement (REM) sleep is distinct from stages of non-rapid eye movement (NREM) sleep and characterized by low-amplitude and mixed-frequency EEG, low muscle tone, and periods of rapid eye movements. REM sleep is scored as stage R and is also known as paradoxical sleep, particularly in rodent research, based on the similarities between REM sleep EEG and waking EEG. In humans, NREM sleep is separated into three stages: N1, N2, and N3. Stage N1 is characterized by low-amplitude and mixed-frequency EEG and slow rolling eye movements. Stage N1 is differentiated from wakefulness by a reduction in alpha (8–13 Hz) EEG activity and an increase in low-amplitude, mixed-frequency EEG activity, particularly in the theta (4–8 Hz) range. Stage N2 characterized by mixed-frequency EEG that is generally slower with more delta (0–4 Hz) and theta EEG activity than wake and stage N1 and by the presence of sleep spindles and K-complexes. A spindle is defined as a burst of at least 0.5 s of EEG activity within the range of 11–16 Hz, and a K-complex consists of a sharp, high-amplitude (typically >100 μV) component followed by a slower, lower-amplitude component exceeding 0.5 s often followed by a spindle. Stage N3 is characterized by an abundance of high-amplitude (>75 μV peak-to-peak) slow waves (<4 Hz) and is defined by the presence of slow waves in at least 30% of a given epoch and, hence, is also referred to as slow wave sleep. Rechtschaffen and Kales (1968) defined stage 3 and stage 4 NREM sleep as sleep with, respectively, 30–50% and greater than 50% of an epoch containing slow waves, but these stages have often been

considered together as slow wave sleep or, in the manual of sleep scoring by the American Academy of Sleep Medicine (Berry et al., 2015), stage N3. Although they may be less apparent among abundant slow waves, K-complexes and spindles are also present in N3 sleep. Research in rodents and other animals does not always distinguish between stages of NREM sleep. Finally, in addition to sleep stages, arousals within human sleep are scored at the appearance of an abrupt shift in EEG frequency that does not constitute a spindle (American Sleep Disorders Association, 1992; Berry et al., 2015)—abrupt shifts to relatively faster EEG activity in alpha or beta (16–30 Hz) frequency ranges are typical.

Sleep and wake states can also be distinguished by differences in their concentrations of neuromodulators, including acetylcholine and norepinephrine. In rats and cats, acetylcholine levels drop to one third of waking levels during slow wave sleep and reach near wake levels during REM sleep (Hasselmo & McGaughy, 2004; Jasper & Tessier 1971; Kametani & Kawamura, 1990; Marrosu et al., 1995). Cholinergic neuron activity in the pedunculopontine tegmentum and the laterodorsal tegmentum of the brainstem are involved in the initiation of REM sleep (McCarley, 1981; Sakai et al., 1979; Van Dort et al., 2015), and the synchronous activity of thalamocortical circuits in NREM sleep is possible, in part, through a reduction in activity in these nuclei (McCormick, 1989; Saper et al., 2005). The locus coeruleus, the main source of norepinephrine in the forebrain, shows high activity during wake, shows intermediate activity during NREM sleep, and is effectively silent during REM sleep (Aston-Jones & Bloom, 1981), reducing forebrain norepinephrine activity in REM sleep.

The stages of sleep follow a predictable pattern in healthy adult human sleep (Dement & Kleitman, 1957; Williams et al., 1964, 1966). Sleep consists of a repeated cycle from NREM sleep to REM sleep with a cycle lasting roughly 90 min although cycle

duration varies extremely both between and within individuals with a reported range of 58–176 min (Sterman & Hoppenbrouwers, 1971). A typical night of sleep starts with a brief period of N1 sleep then progresses into deeper NREM stages of N2 and then N3 before a brief period of REM sleep to complete the cycle. Occasional and brief returns to wakefulness and N1 sleep notwithstanding, this cycle continues throughout the night although the relative durations of N2, N3, and REM sleep within the cycle change over the night. In the first half of the night (i.e., the first 2–3 cycles), there is a relative abundance of N3 sleep over REM sleep. In the second half of the night, there is much more REM sleep and N3 sleep may not be observed at all. N2 sleep is present throughout the night and constitutes roughly 50% of sleep throughout the night.

Features of NREM Sleep

The major patterns of activity that exist within and fundamentally define NREM sleep are spindles and slow waves. Spindles appear nearly synchronous across the cortex, and this coherence relies on corticothalamic feedback to organize spindle activity within the thalamus where spindles are generated by reticular neurons targeting thalamocortical neurons with rhythmic bursts that result in rhythmic excitatory postsynaptic potentials and occasional action potentials in cortical neurons (Contreras et al., 1996, 1997; Steriade, 2003; Steriade et al., 1985; Steriade et al., 1987). The frequency of spindles in sleep EEG is known as the sigma band which, although variously defined, is approximately 10–16 Hz. Within individuals, there is evidence of both slow (≈10 Hz) and fast (≈13 Hz) spindles; however, there are large individual differences in the frequency of slow (9.3–12.0 Hz) and fast (12.5–15.4 Hz) spindles in young adults (Cox et al., 2017; Schabus et al., 2007). Although a functional distinction between slow and fast spindles is not yet certain, fast spindles have been more often linked to learning and memory

(Barakat et al., 2011; Cox et al., 2014; Fang et al., 2017; Mölle et al., 2011; Rihm et al., 2014; Tamaki et al., 2008; 2009). Indeed, it was within the timeframe of fast sigma power evoked by memory cues during sleep that memory content was recently decoded, providing evidence of memory reactivation (also referred to as memory replay) aligned with sleep spindles (Cairney et al., 2018).

Slow waves have been separated into slow oscillations (<1 Hz or 0.5–2 Hz; often 0.5–1 Hz) and delta waves (1–4 Hz) based on their frequency and believed source of generation: cortical circuits for the slow oscillation and, for delta waves, both cortical circuits on their own and a thalamocortical circuit like that of sleep spindles (Steriade, 2003). A slow oscillation is identified as a repeated sequence of depolarization of the cellular membrane that often occurs with high firing rates followed by a hyperpolarization of the cellular membrane with a very low firing rate, states that are known as up and down states, respectively. The functional significance of slow oscillations during NREM sleep likely lies in their ability to coordinate other brain rhythms with their up state inducing other rhythms such as spindles and delta waves (Mölle et al., 2002; Staresina et al., 2015; Steriade et al., 1993). While cortical circuits are sufficient for slow oscillations and some delta waves (Steriade et al., 1993; Timofeev et al., 2000), thalamocortical neurons and intrinsic thalamic oscillators likely play a role in the coherence of the slow waves across cortical regions that is characteristic of slow wave sleep (Crunelli & Hughes, 2010). Notably, the amount slow wave activity observed during NREM sleep is linked to homeostatic regulation of sleep and is considered a physiological marker of sleep need (Borbély & Achermann, 1999). Furthermore, slow wave activity appears to be determined in a local, use-dependent manner such that, for example, reduced slow wave activity in the contralateral sensorimotor cortex is observed after immobilization of an

arm during the day (Huber et al., 2006).

NREM sleep also includes well-studied neural events known as hippocampal sharp wave ripples (Buzsáki, 2015). The sharp wave ripple complex is observed in mammals and consists of a large 40–100-ms deflection (sharp wave) followed by shorter fast oscillations (ripple) occurring in CA1 of the hippocampus during conditions of reduced extrahippocampal input such as NREM sleep (Bragin et al., 1999; Buzsáki et al., 1983). Timing of sharp wave ripples in NREM sleep are highly linked to delta waves, slow oscillations and spindles, forming an extensive hippocampal-thalamic-neocortical circuit (Clemens et al., 2007; Sirota et al., 2003; Staresina et al., 2015). Excitatory bursts from sharp wave ripples can result in large scale effects on cortical and subcortical structures, including increased sigma activity in the prefrontal cortex (Wierzynski et al., 2009). Notably, there are structural similarities between hippocampal sharp wave ripples and neuronal events in other brain regions, including sharp waves and ripples in the olfactory cortex and amygdala and the neocortical K-complex with spindle (Buzsáki, 2015). Of importance to the discussion of sleep and memory, sharp wave ripples have been identified as a source of memory reactivation during NREM sleep as both neocortical and hippocampal firing patterns surrounding these events resemble a temporally compressed version of the patterns observed in recent learning in rats (Ji & Wilson, 2007; Nádasdy et al., 1999; O'Neill et al., 2008; Wilson & McNaughton, 1994). In addition, Zhang et al. (2018) recently identified similar spontaneous memory reactivation in the medial temporal lobe using intracranial EEG.

Features of REM Sleep

REM sleep electrophysiology includes biphasic potentials termed ponto-geniculo-occipital (PGO) waves and theta activity. PGO waves have been identified in REM sleep

of mammals, including cats, rats, and non-human primates, typically as a waveform generated in the pons that propagates through the lateral geniculate nucleus of the thalamus and potentially to multiple cortical regions (Callaway et al., 1987; Datta, 1997; Gott et al., 2017; Jouvet, 1962). The term P wave is also used, especially in cases for which only the pontine element is identified. Bursts of PGO waves in cats are highly linked to eye movements during REM sleep (Nelson et al., 1983; Vanni-Mercier & Debilly, 1998). In rats, it has been shown that P waves are typically phase-locked with the hippocampal theta rhythm (Karashima et al., 2007; Karashima et al., 2001). PGO waves likely support synaptic change and development as blocking PGO waves activity in kittens impairs developmentally beneficial reductions in plasticity of the lateral geniculate nucleus (Shaffery et al., 1999). Although not as well characterized, evidence of P waves in humans has been identified in scalp (McCarley et al., 1983), cortical (Salzarulo et al., 1975), pontine (Lim et al., 2007), and subthalamic (Fernández-Mendoza et al., 2009) EEG recordings.

The hippocampal theta rhythm has been studied extensively in rats as it is evident during active exploration, prominent during REM sleep, and implicated in learning and memory (see Buzsáki, 2002; Hutchison & Rathore, 2015; Pignatelli et al., 2012). The rat hippocampal theta rhythm is generated in brainstem nuclei including the pedunculopontine tegmentum and laterodorsal tegmentum (Datta & Siwek, 1997; Nowacka et al., 2002) and relies on other extrahippocampal structures including cholinergic (Lee et al., 1994) and GABAergic (Yoder & Pang, 2005) neurons in the medial septum. Theta activity coherent with the hippocampal theta rhythm has also been observed in the rat amygdala and prefrontal cortex (Durán et al., 2018; Lesting et al., 2011. Popa et al., 2010; Siapas et al., 2005). There is evidence of a role for the

hippocampal theta rhythm in bidirectional synaptic plasticity. In anesthetized rats and in rat hippocampal slices, stimulation of hippocampal cells induces LTP when it occurs during theta rhythm peaks and induces depotentiation when it occurs during theta rhythm troughs (Hölscher et al., 1997; Huerta & Lisman, 1995; Pavlides et al., 1988). It has been proposed that the human analog of the hippocampal theta rhythm observed in smaller mammals is a ≈3 Hz delta rhythm that is observed in both wake and REM sleep, is associated with encoding, and exhibits hippocampal-neocortical coherence (Lega et al., 2012; Moroni et al., 2007). Traditional scalp EEG methods are insufficient in detecting this hippocampal rhythm, but, due to the noted coherence between it and neocortical activity, REM sleep theta or ≈3 Hz activity measured at the scalp may be indicative of the underlying hippocampal rhythm.

Hypotheses of Sleep and Memory

Memory Terminology

Memory function consists of experience-driven neural changes (encoding) that persist for some period (storage) and are expressed through thoughts or behaviour (retrieval). Multiple terms have been used to indicate the neural changes representing a specific experience or expression, including memory trace, memory representation, and engram. At the cellular level, memories are thought to be formed through long-term potentiation (LTP) and long-term depression, which respectively refer to long-lasting increases and decreases in synaptic strength typically induced by stimulation (Abraham & Williams, 2008; Baltaci et al., 2019; Collingridge et al., 2010). LTP notably consists of an early stage independent of protein synthesis and a late stage dependent on protein synthesis, and characteristics of LTP are consistent with Hebbian learning rules (Hebb,

1949). Long-term depression can occur through various mechanisms, can involve protein synthesis, and can occur de novo or after LTP, in which case it is referred to as depotentiation. At the systems level, memory is often divided into declarative and non-declarative memory systems. Declarative memories are those for facts and events which rely on the hippocampus, at least initially. There are multiple types of non-declarative memory, but the most relevant here is procedural memory, which includes acquired skills and habits (i.e., how to do things) and relies on a neocortical-striatal system (Squire et al., 1993).

Another key term in the discussion of sleep and memory is memory consolidation. Most typically, consolidation refers to the process by which memories are transformed from a labile, temporary state into a more stable, long-lasting state. The process was proposed to account for retroactive interference in declarative memories (Müller and Pilzecker, 1900; Lechner et al., 1999), and is supported by evidence of temporally-graded retrograde amnesia (i.e., forgetting recent memories while remote memories remain intact) after damage or removal of the medial temporal lobe, most notably through Scoville and Milner's (1957) study with patient H.M. The term memory consolidation has since been used to cover multiple offline (i.e., not during learning) alterations to declarative memories and procedural memories that result in enhancement, stabilization, or integration measured physiologically or behaviourally (Robertson et al., 2004; Squire et al., 2015). A distinction has been made between synaptic or cellular consolidation, referring to changes that substantiate memory at the cellular level, and system consolidation, referring to larger-scale reorganization of memory or its integration with other long-term memories. The structures and connections supporting memory retrieval indeed change over time, supporting the notion of memory consolidation; however, the

time course of memory formation and the properties which would distinguish consolidated from “non-consolidated” memories are not yet clear and vary by mode (Moscovitch et al., 2005; Nadel et al., 2000; Runyan et al., 2019).

Existing Hypotheses

One approach to understanding sleep’s role in memory has been to associate different sleep states to different types of memory in variations of a “dual process” hypothesis. Investigation of dissociable NREM and REM sleep effects has included use of the early-night/late-night paradigm in which researchers compare the effects of an early-night sleep retention period naturally rich in deep NREM sleep to the effects of a late-night sleep retention period naturally rich in REM sleep (e.g., Yaroush et al., 1971). To this end, researchers have also used selective deprivation of REM sleep (REMD) both in rodents—by placing subjects in an apparatus in which REM sleep onset muscle atonia causes awakenings after falling into water (e.g., Pearlman, 1969)—and in humans—by waking participants upon PSG signs of REM sleep (e.g., Empson & Clarke, 1970). Although details varied by account and evidence considered, the emergent notion from research with these paradigms was that new declarative memories preferentially benefit from NREM sleep, typically slow wave sleep, and new non-declarative, especially procedural, memories preferentially benefit from REM sleep (Gais & Born, 2004; Peigneux et al., 2001; Plihal & Born, 1997, 1999; Smith, 1995).

However, a complete dissociation of NREM and REM sleep serving different memory systems is incongruent with many findings. For example, simple procedural memory tasks were seen to be more dependent on stage 2 NREM sleep than REM sleep (Smith et al., 2004; Smith & MacNeill, 1994). Conversely, REM sleep was implicated in memory for prose (Empson & Clarke, 1970; Tilley & Empson, 1978) and learning a

second language (De Koninck et al., 1989). As such, it may be that both sleep states support declarative and procedural memory, and REM sleep has specific benefits for new memories in complex tasks, although the defining features of a "complex" task are undetermined (Ackermann & Rasch, 2014; Smith et al., 2004; Stickgold, 1998; Tilley et al., 1992). Some have been critical of the early research, contending that it does not indicate there are memory functions unique to REM sleep or sleep in general (Siegel, 2001; Vertes & Eastman, 2000; Vertes & Siegal, 2005). Genzel et al. (2015) argued that REM sleep specifically supports emotional or "amygdala-related" memory processing whereas NREM sleep is important for cortically based memories. Indeed, REM sleep has frequently been implicated in emotionally charged memory while REM sleep connections with other material is less conclusive (Ackermann & Rasch, 2014; Genzel et al., 2015; Tempesta et al., 2018), but this hypothesis has received little direct testing thus far.

Consideration of the sequential nature of sleep cycles led to the development of the sequential hypothesis of the sleep function which emphasizes the importance of slow wave sleep to REM sleep sequences in memory processing. The sequential hypothesis proposes that NREM sleep contains selective processes that weaken non-adaptive memories before REM sleep stores the surviving memories and integrates them with preexisting memories (Giuditta, 1985, 2014; Giuditta et al., 1995). This hypothesis is supported by studies relating overnight retention of words to the integrity of NREM-REM sleep cycles (Ficca et al., 2000; Mazzoni et al., 1999). There is convincing evidence for adaptive selectivity in memory processing over sleep and for sleep involvement in the integration of new memories with existing knowledge (Stickgold & Walker, 2013); however, it is inconclusive whether NREM sleep and REM sleep provide these benefits specifically and respectively.

The active system consolidation hypothesis also proposes that NREM and REM sleep work together in memory processing, and it is especially focused on the neural mechanisms mediating sleep's beneficial effect on memory consolidation (Diekelmann & Born, 2010; Rasch & Born, 2013). This hypothesis assumes that events of wake are encoded across cortical networks and bound together by the medial temporal lobe; then, during sleep, hippocampal reactivation of these events reactivates cortical components, strengthening cortico-cortical connections and reorganizing memory representations. NREM sleep, particularly slow wave sleep, mediates the transformation of memories from temporary to long-term states and the integration of new memories with pre-existing ones (i.e., system consolidation) through reactivations driven by slow oscillations and occurring within hippocampal sharp wave ripples and thalamocortical spindles. REM sleep is thought to subsequently stabilize changes acquired in NREM sleep through LTP or long-term depression (i.e., synaptic consolidation). This active system consolidation hypothesis is well-supported by many findings in the to-be-discussed research, including memory benefits from externally-induced reactivations during NREM sleep (Belal et al., 2018; Oudiette & Paller, 2013) and increased expression of genes associated with synaptic plasticity in the cortex during REM sleep (Ribeiro et al., 2007).

Alternative distinctions for the roles of NREM and REM sleep have also been proposed. Poe et al. (2010) proposed that slow wave activity of NREM sleep is important for converting early LTP into long-lasting LTP, in part because late LTP requires protein synthesis (Frey et al., 1988; Krug et al., 1984; Otani et al., 1989) and protein synthesis is increased during slow wave sleep (Nakanishi et al., 1997; Ramm and Smith, 1990). Poe et al. (2010) proposed that REM sleep is an opportune time for bidirectional synaptic change including both new LTP and depotentiation (at least in the hippocampus) given that

heightened cholinergic activity favours induction of LTP and that low norepinephrinergic activity is essential for depotentiation (Katsuki et al., 1997; O'Dell et al., 2015; Thomas et al., 1996; Yang et al., 2002). Seibt and Frank (2019) proposed a similar model stating that waking experience induces transient neuronal changes and primes memory circuits for stabilization during sleep through, for example, processes of synaptic "tagging" (Frey & Frey, 2008; Redondo & Morris, 2011) or changes in neuronal excitability (Benito & Barco, 2010). They propose that REM sleep selectively strengthens and weakens memory traces that have been marked for these actions by NREM sleep through memory reactivation and oscillatory activity within primed circuits.

Synaptic weakening, depotentiation, or forgetting during sleep has been considered important in preventing saturation of memory systems or reducing energy demands. Whereas Giuditta (1985) proposed that selective weakening of non-adaptive memories occurs during NREM sleep, Crick and Mitchison (1983) argued that REM sleep served this function. The influential synaptic homeostasis hypothesis (Tononi & Cirelli, 2003, 2006) argues that slow waves of NREM sleep prevent system saturation by driving a non-specific, global downscaling of synaptic strength. This process is proposed to benefit memory performance indirectly by nullifying the weakest connections effectively constituting neuronal background noise relative to stronger memory traces. Poe (2017) expanded upon her earlier model, proposing that targeted depotentiation of synapses during sleep would contribute to forgetting, a reduction of noise in perceptual and memory systems, schema development, and synaptic pruning during development. Poe (2017) proposed that the inactivity of the norepinephrine-providing locus coeruleus during REM sleep and the second preceding spindles (Aston-Jones & Bloom, 1981) makes these periods ideal for targeted forgetting of somatosensory and hippocampal

memories (i.e., memories dependant on forebrain regions in reach of the locus coeruleus) and that reduced extracellular dopamine during slow wave sleep (Léna et al., 2005) provides a similar opportunity for dorsal striatal-dependent memories, including motor and procedural memories. Thus, it is noted that while much attention has been given to the notion that sleep strengthens memories, an additional role for sleep in forgetting should also be considered.

Hypotheses proposing complementary functions for NREM and REM sleep are supported by a few studies showing that improvements in visual texture discrimination, a learned perceptual skill, are best promoted by sleep containing both slow wave sleep and REM sleep (Mednick et al., 2003; Stickgold, Whidbee et al., 2000; Gais et al., 2000). However, studies of sleep and memory often implicate either NREM sleep or REM sleep and rarely both for the same memory task, contributing to development of hypotheses of separate sleep states for separate memory systems. Identification of a precise role for REM sleep in memory processing has been particularly elusive (Ackermann & Rasch, 2014; Rasch & Born, 2015), and it remains undetermined whether sleep promotes active and selective forgetting. Multiple existing models account for research into the memory processing during sleep and the structure of memory representations; however, additional questions remain regarding the effects sleep on memory performance: when will performance associate with properties of NREM sleep? when will performance associate with properties of REM sleep? and what specific effect on memory performance is granted by each of these sleep states?

Sleep Reinforcement and Refinement Hypothesis

Here I detail a new hypothesis addressing the effect of sleep on memory performance in which sleep is proposed to support later retrieval of newly acquired

memories through processes of reinforcement and refinement. In this sleep reinforcement and refinement (SR2) hypothesis, memory reinforcement is primarily attributed to NREM sleep, and memory refinement is primarily attributed to REM sleep. NREM sleep memory reinforcement is proposed to support memory strength, maintaining or potentially increasing the accessibility of memories (i.e., their likelihood of retrieval) and increasing their resistance to interference from future learning. REM sleep memory refinement is proposed to support memory fidelity, maintaining or potentially increasing the accuracy of memory retrieval. NREM and REM sleep are considered primary for reinforcement and refinement, respectively; however, it should be acknowledged that a perfect dissociation is unlikely and that wake states and features of sleep physiology not directly tied to specific states may also contribute to these memory benefits in ways beyond this proposed dissociation.

Details of this hypothesis may be best expressed using terms of signal detection theory. It is assumed that a dominant (often the most accurate) memory trace (signal) exists within noise from highly related and competing memory traces (interference) and from random or external sources. It is proposed that NREM sleep reinforces memories by further raising signal and interference levels above random and external noise. This effect may occur through application of gain to new memory traces, an attenuation of random and external noise, or both. It is proposed that REM sleep refines memories by applying signal gain and attenuating interference, increasing both the signal-to-noise and signal-to-interference ratios. It may be that NREM sleep gain or attenuation is multiplicative such that levels of stronger and weaker or interfering memory traces are differentially affected, but selective processing in REM sleep is thought to be more critical in separating signal and interference traces and contributing to memory fidelity.

The SR2 hypothesis is built upon the notion that the fidelity of memory representations can vary even among equally accessible memories. In other words, an individual may remember the occurrence of two events with equal ease, but one of those memories may be remembered more clearly, with more details, or with more confidence. The distinction between the presence and fidelity of memory representations is prominent in research of visual short-term memory (e.g., Zhang & Luck, 2008), but the distinction also has relevance to memories held over longer durations (e.g., Brady et al., 2013). Much of the research in sleep and memory has not used tasks in which memory strength and fidelity are separately examined. However, it seems likely that performance in simple memory tasks would be largely or entirely influenced my memory strength, whereas performance in complex memory tasks would also depend on memory fidelity. Thus, SR2 hypothesis, proposing memory strength to be largely influenced by NREM sleep and memory fidelity to be largely influenced by REM sleep, is consistent with the previously articulated notion that REM sleep has specific benefits for memories of complex material (Ackermann & Rasch, 2014; Smith et al., 2004; Stickgold, 1998; Tilley et al., 1992).

While the SR2 hypothesis is primarily focused on memory performance, potential mechanisms should be considered. Memory reinforcement from NREM sleep is thought to occur through repeated experience-dependent offline memory reactivations coordinated by slow wave activity, sharp wave ripples, and spindles. These features are thought to maintain the accessibility of a newly acquired memories by converting the early LTP of transient memory traces into long-lasting LTP via protein synthesis, perhaps while also downscaling synapses external to the reactivated memories. REM sleep refinement is thought to occur through bidirectional action on these memory traces, including additional potentiation within dominant memory traces and depotentiation of weaker

memory traces, perhaps through phase-dependent firing on low frequency REM sleep oscillations.

The SR2 hypothesis is informed by and compatible with previous hypotheses of sleep and memory as its intended purpose is to bind elements of these hypotheses to predict memory change over sleep as it is expressed through performance. The reinforcement of newly acquired memories during NREM sleep may be considered a product of system consolidation of NREM sleep the active system consolidation hypothesis (Diekelmann & Born, 2010; Rasch & Born, 2013), and global synaptic downscaling by slow wave activity in the synaptic homeostasis hypothesis (Tononi & Cirelli, 2003, 2006) may contribute to this effect. Memory refinement through REM sleep is consistent with proposals of bidirectional synaptic change during sleep (Poe et al., 2000; Poe, 2017; Rasch & Born. 2013; Seibt & Frank, 2019). Hypotheses claiming that selectivity in memory consolidation or weakening of non-adaptive memories occurs through NREM sleep, Giuditta's (2014) sequential hypothesis for example, are not necessarily inconsistent with the notion of REM sleep refinement. As suggested by Seibt and Frank (2019), it is possible that bidirectional selection occurs during NREM sleep and is enacted during REM sleep.

Sleep and Memory Relationships

Memory Change Over Sleep

Sleep-dependent memory consolidation has often been studied by comparing the effect of retention periods with sleep to retention periods of only wake, sometimes involving sleep deprivation with or without recovery sleep before memory testing. Sleep results in greater memory performance for syllable or word pairs or word lists (Barret & Ekstrand, 1972; Benson & Feinberg, 1977; Empson & Clarke, 1970; Gorfine et al., 2007;

Grosvenor & Lack, 1984; Jenkins & Dallenbach, 1924; Lahl et al., 2008; Schabus et al., 2005; Yaroush et al., 1971), an effect most prominent when interfering material is learned after the sleep or wake period (Drosopoulos et al., 2007; Ellenbogen et al., 2006), and for visuospatial knowledge of a map (Plihal & Born, 1999) or locations of objects (Wilhelm et al., 2008) or faces (Talaini et al., 2008). For procedural memories, sleep results in greater visual texture discrimination (Karni et al., 1994; Stickgold et al., 2000) and better performance of a learned finger tapping sequence (Fischer et al., 2002; Korman et al., 2007; Walker et al., 2002), but these benefits may not extend to all procedural memory tasks (Robertson et al., 2004). An important study from Yang et al. (2014) examined memory consolidation at the cellular level by measuring postsynaptic dendritic spine formation in the motor cortex of mice after training to run on a rotating rod. Trained mice showed progressive increases in spine formation relative to untrained mice from 6 to 48 hr after rotarod training. Spine formation was both branch-specific in that it was driven specifically by neuronal branches with a relatively high spine formation and task-specific in that later training on backward rotarod running induced spine formation on the branches with previously low spine formation. Critically, 7-hr sleep deprivation after training reduced spine formation on high formation branches 8 hr after training and new spine survival on these branches 24 hr after training whilst having no effect on spine formation on low formation branches or spine elimination. Non-sleep-deprived mice also showed greater performance improvements at one day and five days post training.

Memory processing over sleep seems to involve a qualitative transformation and reorganization of memories (Landmann et al., 2014; Rasch & Born, 2013). Sleep after learning promotes later generation of insight into hidden patterns contained within learned material (Wagner et al., 2004; Wilhelm et al., 2012; Yordanova et al., 2012),

abstraction of probabilistic information (Durrant et al., 2011), and relational memory (i.e., the ability to infer relationships between distinct events; Ellenbogen et al., 2007; Lau et al., 2010). Finger tapping task performance can improve over intervening wake when the same hand is used both during learning and test, but intervening sleep allows task improvements from both same and opposite hand learning (Cohen et al., 2005; Witt et al., 2010), from motor imagery (Debarnot et al., 2009), and from observation (Van Der Werf et al., 2009). Sleep after learning lists of related words also promotes false recall (but not false recognition; Fenn et al., 2009) of words that were not presented but are highly related to the presented words, suggesting sleep may serve a gist-extraction function (Diekelmann et al., 2010; Payne et al., 2009). In addition, neuroimaging research has implicated sleep in changes to the patterns of neural activity associated with memory retrieval (Durrant et al., 2013; Fischer et al., 2005; Gais et al., 2007; Orban et al., 2006; Sterpenich et al., 2007; Takashima et al., 2006; Walker et al., 2005).

Researchers have used the early-night/late-night paradigm to compare the relative effects of NREM slow wave sleep and REM sleep. Early-night sleep rich in slow wave sleep has been found to benefit memory of word pairs (Barret & Ekstrand, 1972; Yaroush et al., 1971, Plihal & Born, 1997), object locations (Plihal & Born. 1999), picture-colour associations (Groch et al., 2015), and visual texture discrimination (Gais et al., 2000) whereas late-night sleep rich in REM sleep enhances mirror tracing skill (Plihal & Born, 1997), memory for prose (especially emotional prose; Wagner et al., 2001), and the recognition preference for emotional over neutral pictures (Groch et al., 2015). Although informative, the early-night/late-night method is unable to firmly dissociate the contributions of NREM and REM sleep because both states occur in both halves of the night and the design may be confounded by other circadian differences.

NREM Sleep and Memory

Beyond the early-night/late-night paradigm, a role for NREM sleep in memory is supported by evidence that time spent in post-learning slow wave sleep or slow wave activity has been positively correlated with cued recall of word pairs (Schabus et al., 2005), visuospatial memory (Diekelmann et al., 2012), and recognition of learned faces and houses (Schönauer et al., 2017), and more stage 2 sleep in nap was associated with offline improvements in a finger tapping motor sequence task (Nishida & Walker, 2007). Furthermore, suppression of slow wave sleep by acoustic stimulation reduced offline benefits in visual texture discrimination (Aeschbach et al., 2008) and a visuomotor task (Landsness et al., 2009). Short naps containing only NREM sleep were beneficial for memory of word pairs and word lists (Lahl et al., 2008; Tucker & Fishbein, 2008).

More has been learned about the role of NREM sleep in memory processing by examining how learning and memory relate to the previously described concert of spindles, neocortical slow waves, hippocampal sharp wave ripples, and memory reactivation. NREM sleep is affected by previous learning experience as evidenced by the reoccurrence during post-learning NREM sleep of specific neuronal activity patterns present during learning in the rodent hippocampus (Ji & Wilson, 2007; Nádasdy et al., 1999; O'Neill et al., 2008; Peigneux et al., 2004; Qin et al., 1997; Wilson & McNaughton, 1994), neocortex (Hoffman & McNaughton, 2002; Ji & Wilson, 2007; Qin et al., 1990; Yang et al., 2014), and ventral striatum (Lansink et al., 2008), and in human EEG both at the scalp (Schönauer et al., 2017) and intracranially in the medial temporal lobe (Zhang et al., 2018). Furthermore, there is evidence of increased spindle activity after learning declarative material (Clemens et al., 2005; Gais et al., 2002; Meier-Koll et al., 1999; Schabus et al., 2004) and procedural material (Fogel & Smith, 2006; Fogel,

Smith, & Cote, 2007; Morin et al., 2008; Peters et al., 2008; Tamaki et al., 2013), increased slow wave activity after learning procedural material (Huber et al. 2004; Tamaki et al., 2013), increased sharp wave ripples in the rat hippocampus after learning declarative material (Eschenko et al., 2008; Mölle et al., 2009), and greater coherence between slow wave and sigma activity after learning word pairs (Mölle et al., 2004). Greater post-learning spindle activity in task-related regions has been associated with greater retention of words learned before sleep (Clemens et al., 2005; Holz et al., 2012), improved mirror tracing skill (Holz et al., 2012), and improved motor sequence performance (Nishida & Walker, 2007). Greater post learning slow wave activity has been associated with greater word list retention and mirror tracing (Holz et al., 2012). Finally, learning-related increases in spindle activity (Schabus et al., 2004; Tamaki et al., 2013), slow wave activity (Huber et al., 2004; Landsness et al., 2009; Ruch et al., 2012; Tamaki et al., 2013), regional blood flow to the hippocampus during sleep (Peigneux et al., 2004), and evidence of spontaneous memory reactivation (Schönauer et al., 2017) have all been associated with better memory performance. Notably, these learning-related sleep alterations and correlations with performance occur in task-related regions of the brain. Thus, it appears that periods of learning induce changes to NREM sleep physiology that, in turn, relate to better performance outcomes.

Causal evidence for memory reactivation in NREM sleep benefitting performance comes from an experimental technique known as targeted memory reactivation (TMR). Rasch et al. (2007) had participants learn a procedural finger tapping task and a series of paired two-dimensional object locations (i.e., card matching) while exposed to a rose odour. Re-exposure to the odour during nocturnal slow wave sleep improved performance on the location memory task compared to those not re-exposed, and there was no effect of

odour exposure only during sleep or re-exposure during REM sleep. Odour re-exposure during slow wave sleep results in a significant blood oxygenation level-dependent response in the left anterior hippocampus and protects visuospatial memories from retroactive interference (Diekelmann et al., 2011; Rasch et al., 2007). TMR has also benefitted visuospatial memory performance in a within-subjects design in which distinct audio cues were used to cue some memories during the slow wave sleep of a nap (Rudoy et al., 2009). Although Rasch et al. (2007) found no procedural memory benefit of cue re-exposure, at least three studies have reported a benefit: Antony et al. (2012) did using a design that had participants tap patterns to different melodies and then had one of those melodies played covertly during slow wave sleep; Schönauer et al. (2014) did using auditory cues associated with presses of a finger tapping sequences presented during the first 2 hr of a sleep period; and Laventure et al. (2016) did using a design similar to that of Rasch et al. (2007) but with odour re-exposure during stage 2 NREM sleep rather than slow wave sleep. Bendor and Wilson (2012) examined the reactivation during sleep of neuronal firing patterns that were present during learning (i.e., memory replay events) and revealed that re-exposure to sound cues did not increase the absolute number of reactivation events but rather biased these events toward reactivating the patterns associated with the sound cues. Belal et al. (2018) recently found that, after pairing separate motor sequences with separate audio cues, EEG pattern classification could reliably identify the cue presented during NREM sleep based on patterns in sleep EEG following the cue, indicating that memories were indeed reactivated in TMR.

Manipulations of slow oscillations, spindles, and sharp wave ripples have also been shown to affect memory performance. Online identification and blocking of hippocampal sharp wave ripples in rats were found to reduce daily memory

improvements in a radial maze (Girardeau et al., 2009). Ngo et al. (2013) found auditory stimulation during the up state of slow oscillations in human slow wave sleep to induce a prolonged train of slow oscillations, increase amplitude of the slow oscillations, increase 12–15 Hz fast sigma power during the up state, increase 9–12 Hz slow sigma power during up-to-down state transitions, and, critically, increase retention of word pairs learned before sleep relative to sham control. In rats, electrical stimulation designed to reinforce coordination between slow waves, spindles, and sharp wave ripples led to increased memory measured via object discrimination (Maingret et al., 2016). These findings support the abundance of correlational data and further demonstrate a causal link between NREM sleep physiology and memory performance.

The effect of NREM sleep on memory can also be studied at the cellular and molecular level. The low cholinergic activity during slow wave sleep (Jasper & Tessier 1971; Kametani & Kawamura, 1990; Lapierre et al., 2007; Marrosu et al., 1995) limits induction of LTP during slow wave sleep (Bramham & Srebro, 1989; Leonard et al., 1987) as evidenced by an absence expression of plasticity-related immediate early gene EGR-1 during NREM sleep in rats even after exposure to an enriched environment or induction of hippocampal LTP during wake (Ribeiro et al., 1999; Ribeiro et al., 2002). However, slow wave sleep is associated with increased synthesis of proteins (Ramm & Smith, 1990; Nakanishi et al., 1997), including actin (Vazquez et al., 2008), which is involved in maintenance of LTP and the modulation of dendritic spines (Bramham, 2008). Yang et al. (2014) who found sleep deprivation reduced branch- and task-specific dendritic spine formation in the mouse motor cortex also investigated the role of memory reactivation. Reducing post-running sleep activity of neurons associated with forward-running (i.e., memory reactivation) via either *N*-methyl-D-aspartate receptor blocker

MK801 or backward treadmill running halfway through the sleep period also reduced branch-specific dendritic spine formation. Notably, REMD did not disrupt branch-specific spine formation after rotarod training. Thus, there is support for the notion that NREM sleep memory reactivation reinforces memory traces, in part through a stabilizing early LTP into late-LTP, as suggested by Poe et al. (2010) and endorsed by the SR2 hypothesis.

REM Sleep and Memory

In animal models, there are often increases in REM sleep after periods of learning (Destrade et al., 1978; Smith, 1995; Smith et al., 1980), but this effect is not as robust among human studies (De Koninck et al., 1990; Fogel & Smith, 2006; Fogel et al., 2007; Mandai et al., 1989; Smith & Lapp, 1991; Verschoor & Holdstock, 1984) and may depend on the novelty, complexity, or emotional tone of the learning period. The amount of post-learning REM sleep has been positively associated with performance outcomes in learning a second language (De Koninck et al., 1990), Morse code (Mandai et al., 1988), and a finger tapping sequence (Fischer et al., 2002). However, greater REM sleep duration has also been associated with overnight forgetting of low-value items in a visuospatial memory task (Oudiette et al., 2013) and overnight decrements in learning how to ride a bicycle with reversed handlebars (Hoedlmoser, 2015). These recent findings highlight a complex relationship between REM sleep and memory and are consistent with proposals that sleep, perhaps REM sleep especially, is for the removal of undesirable or cluttering memories in synaptic networks (Crick & Mitchison, 1983; Poe, 2017).

Much of the evidence for REM sleep involvement in memory comes from studying the effect of REMD after learning. Early work in animal models, primarily avoidance learning in rodents, showed that REMD impairs performance particularly for

more complex, two-way avoidance tasks and when REMD is applied in time windows in which there would otherwise be a post-learning increase in REM sleep (Smith, 1985). Human REMD studies have found mixed results, which may reflect their typically small sample sizes or important differences in their memory tasks. For overnight retention of verbal declarative memory, REMD impairments, compared to concurrent or matched NREM awakenings or slow wave sleep deprivation, have been observed for memory of prose (Empson & Clarke, 1970; Tilley & Empson, 1978) and words evoking ego threat (Grieser et al., 1972) but not neutral word lists (Empson & Clarke, 1970; Lewin & Glaubman, 1975) and word pairs (Chernik, 1972; Ekstrand et al., 1971). For procedural memories, Karni et al. (1994) found overnight improvements in visual discrimination to be blocked by REMD, and Smith and colleagues identified overnight REMD impairments in tapping learned sequences, tracing figures viewed through a mirror, the Tower of Hanoi puzzle, but not in many declarative memory tasks, direct tracing, or the pursuit rotor task (reviewed in Smith et al., 2004). Thus, REMD sometimes impairs memory performance, particularly when task material is complex. This claim is consistent with late-night sleep benefitting memory for mirror tracing, but not word pairs or object locations (Barret & Ekstrand, 1972; Plihal & Born, 1997, 1999; Yaroush et al., 1971), and the SR2 hypothesis notion that REM sleep offers the benefit of memory refinement. It has been noted that long-term antidepressant suppression of REM sleep does not result in overt declines in cognitive ability (Siegel, 2001; Vertes & Eastman, 2000). However, there is some evidence that some antidepressants may impair overnight changes in memory (Genzel et al., 2011; Goerke et al., 2014; Dresler et al., 2010; Watts et al., 2012), and subtle differences in memory fidelity may not be apparent in everyday living.

Neural oscillations of REM sleep have been linked to memory processing. Post-

training increases in P wave density or P wave generator activity are observed following two-way avoidance training (Data, 2000; Ulloor & Datta, 2005) and fear extinction training (Datta & O'Malley, 2013), and post-learning increases in P wave density predict retention of avoidance (Data, 2000) and extinction (Datta & O'Malley, 2013) training. Hippocampal theta power in REM sleep has been correlated with a decreased rate and increased synchrony of neuronal firing in the rat hippocampus over REM sleep (Grosmark et al., 2012). Rat hippocampal cells active during exploration have been observed to fire on theta rhythm peaks during post-exploration REM sleep and then on theta rhythm troughs after the explored setting became familiar (Poe et al., 2000). Louie and Wilson (2001) showed that the firing patterns of rat hippocampal neurons that occurred while running a path were reproduced in subsequent REM sleep with equivalent theta rhythm modulation such that neurons tended to fire during the same phase of the rhythm during REM sleep as they did during path running. Given the previously noted phase-dependent bidirectional synaptic plasticity associated with the hippocampal theta rhythm in rats (Hölscher et al., 1997; Huerta & Lisman, 1995; Pavlides et al., 1988), Poe et al. (2000) noted that such phase modulation might drive a beneficial selective strengthening and weakening of memories in REM sleep. Boyce et al. (2016) provided more causal evidence of theta activity involvement in memory processing, showing that post-learning optogenetic inhibition of the hippocampal theta rhythm in mice impaired subsequent expression of object place recognition and conditioned fear. In humans, REM sleep theta power has been associated with recognition of schema-consistent melodies learned before sleep (Durrant et al., 2015) and selectivity of memory for paired associates (MacDonald & Cote, 2016).

REM sleep has been proposed to be an opportune time for synaptic change

(Diekelmann & Born, 2010; Poe et al., 2010; Ribeiro et al., 2007), a proposal in part due to cholinergic tone being at near-waking levels during REM sleep and substantially higher than during NREM sleep (Hasselmo & McGaughy, 2004; Jasper & Tessier 1971; Kametani & Kawamura, 1990; Marrosu et al., 1995). High cholinergic activity has been shown to support late LTP in the medial prefrontal cortex of anesthetized rats (Lopes-Aguiar et al., 2008), support long-term depression in slices of rat visual cortex (Kirkwood et al., 1999), and activate plasticity-related immediate early genes, ARC in rats (Teber et al., 2004), and EGR-1 in human cell cultures (von der Kammer et al., 1998). Indeed, there is increased ARC and EGR-1 expression during REM sleep in rats following exposure to novel stimuli (Ribeiro et al., 1999; Ribeiro et al., 2007), induced hippocampal LTP (Ribeiro et al., 2002), or shock avoidance learning (Ulloor & Datta, 2005) and in response cholinergic activation of P waves (Datta et al., 2008). Post-learning increases in ARC are associated with post-learning increases in P wave density (Ulloor & Datta, 2005) and are abolished by elimination of P wave generating cells (Datta et al., 2008). In addition, blocking cholinergic activity during post-training REM sleep impairs later performance in a habit-based version of a radial arm maze in rats (Legault et al., 2004, 2006). In humans, blocking cholinergic activity during post-learning late-night sleep, but not wake, reduced offline gains on a newly learned finger tapping task (Rasch et al., 2009), and increasing acetylcholine availability increased offline gains in mirror tracing ability (Hornung et al., 2007).

An understanding of the role of REM sleep in synaptic plasticity and memory processing is greatly informed by a study from Li et al. (2017) measuring dendritic spine formation and elimination in the mouse motor cortex. Groups of mice subjected to REMD by gentle handling, compared to no deprivation and similar NREM sleep interruptions,

showed reduced elimination of spines that were newly formed after rotarod training whilst there were no differences observed between groups for elimination of existing spines. Newer spines formed after subsequent training on a reversed direction rotarod tended to form near where REM sleep spine elimination occurred, and REMD mice had lower spine formation after reversed rotarod training and impaired performance on the reversed rotarod compared to non-deprived mice, suggesting REM sleep pruning of spines is important for subsequent learning. Of the persistent spines that survived initial pruning, a greater number showed continued survival four days post training in groups with REM sleep compared to REMD mice, an effect attributable to a post-training REM sleep-dependent increase in persistent spine size. Thus, Li et al. found REM sleep to not only prune some newly formed spines but also strengthen other newly formed spines, and both actions were dependent on calcium spikes on apical dendrites during REM sleep. Critically, REMD mice showed less performance improvement over time than mice with REM sleep even 12 hr after recovery from REMD, indicating that optimal memory performance relies on a REM sleep-dependent selective strengthening and weakening of newly formed spines. These findings and others support the notion that REM sleep is a time of bidirectional synaptic change (Poe, 2017; Rasch & Born. 2013; Seibt & Frank, 2019), and this effect of REM sleep may result in refined memory traces and greater accuracy in memory performance, as proposed by the SR2 hypothesis.

Conclusion

Over a century of research on sleep and memory has made it clear that sleep provides benefits for newly acquired memories. However, the effect of sleep or a particular sleep state on memory performance is not always apparent and the clarity of an effect likely depends on properties of the material and testing procedure. A surge of

research since the turn of the century has also made it clear that active processes restructure and transform newly acquired memories during sleep. NREM sleep likely acts on memories in part through reactivations of recent experiences within a coordinated concert of neural oscillations and an associated increase in protein synthesis may support late LTP within reactivated memory circuits. The mechanisms of REM sleep memory processing are less clear, but REM sleep appears to be a time of bidirectional synaptic change with impacts on memory function. The SR2 hypothesis proposes that NREM sleep benefits memory accessibility through processes of reinforcement and REM sleep benefits memory fidelity through processes of refinement. These proposals are consistent with the pattern that an effect of REM sleep is most evident for tasks demanding retrieval of complex material or a fine-tuned performance. Plausible mechanisms of NREM sleep reinforcement and REM sleep refinement have been provided by research on synaptic plasticity, particularly work linking NREM sleep to task-specific dendritic spine formation (Yang et al., 2014) and work linking REM sleep to selective strengthening and weakening of newly formed dendritic spines (Li et al., 2017). However, direct tests of the SR2 hypothesis are lacking.

Testing the Sleep Reinforcement and Refinement Hypothesis

The SR2 hypothesis proposes that sleep supports later retrieval of newly acquired memories through reinforcement and refinement; thus, direct testing of the hypothesis requires memory tasks that can adequately separate the contributions of reinforcement and refinement at retrieval. NREM sleep reinforcement is thought to benefit the accessibility of memories, making them more likely to be retrieved. REM sleep refinement is thought to benefit the fidelity of a memory representations, allowing them to be retrieved more precisely and more accurately. The studies conducted for this

investigation utilized continuous report memory tasks in which participants were shown items varying on a continuous feature and then reported on that feature on a continuous scale after a retention period. The frequency at which participants showed retrieval of the critical feature from memory by reporting near the target value was used to measure memory accessibility and thus the extent of memory reinforcement. Exactly how close reports tended to be to their corresponding targets when there was evidence of memory retrieval was used to measure memory fidelity and thus the extent of memory refinement.

The SR2 hypothesis was examined over a series of three studies. Study 1 was designed to test whether sleep reinforces and refines memories by comparing performance between a condition in which the retention period of the memory task included a nap and a condition in which the retention period did not include a nap. It was predicted that both memory accessibility and memory fidelity would be greater in the nap condition. Study 2 was designed to test whether NREM sleep reinforces memories by utilizing TMR during N3 sleep to boost NREM sleep processing of memories. It was predicted that there would be memory accessibility benefits for items cued during NREM sleep relative to items not cued during sleep. Study 3 was designed to test whether REM sleep refined memories by utilizing REMD in effort to disrupt REM sleep processing of memories. It was predicted that memory fidelity would be impaired for participants who received REMD relative to participants who received a comparable number and pattern of awakenings during light NREM sleep. In addition to the manipulations of sleep, it was predicted that, throughout these three studies, measures of NREM sleep such as slow wave activity would correlate more with memory accessibility than memory fidelity and that measures of REM sleep such as the amount of stage R sleep would correlate more with memory fidelity than memory accessibility.

Study 1, Study 2, and Study 3 are reported in Chapter 2, Chapter 3, and Chapter 4, respectively. Each chapter includes a brief introduction highlighting pertinent past research of similar design to the study reported in that chapter. Each chapter also includes a discussion of the results of the study reported within and a brief discussion of limitations that are largely particular to that study. Chapter 5 concludes the doctoral dissertation with a general discussion of what was learned from the results of these three studies regarding the SR2 hypothesis and a discussion of the strengths and limitations of the methods shared among these studies.

Chapter 2

Study 1: The Effect of an Afternoon Nap

The proposal that sleep aids memory accessibility and fidelity through reinforcement and refinement was first examined in an experiment comparing the effect of an afternoon retention period with a 90-min nap opportunity to an afternoon retention period without a nap. Similar designs have found nap-related benefits in memory performance. Gorfine et al. (2007) found improved cued-recall of unrelated word pairs after a 120-min nap opportunity but no such improvement after only wakefulness. A 60-min nap opportunity without REM sleep has been shown to result in greater free recall of word lists learned before retention (Lahl et al., 2008) and greater performance improvement from training for cued-recall on unrelated word pairs, maze completion, and complex figure recall for those over median performance at training (Tucker & Fishbein, 2008). Lahl et al. (2008) even found free recall performance in a 6-min nap group to be greater than the wake group, although not as great as longer nap group with six times more sleep on average. For non-declarative memory tasks, Mednick et al. (2003) found visual texture discrimination improvement from both 60- and 90-min nap opportunities if the nap contained both slow wave sleep and REM sleep, and Korman et al. (2007) found the number of correct sequences tapped to be greater 2 hr after training if a 90-min nap opportunity occurred between training and test sessions.

The declarative memory task used in this investigation, referred to as the item-colour task, asked participants to remember the colour of many household objects presented to them in learning trials (Figure 2.1a) and then report back that colour from a continuous colour wheel approximately two hours later in test trials (Figure 2.1b). This task was programmed by K. M. but directly modeled after the task used by Brady et al.

(2013) to examine the fidelity of memory representations held over several minutes.

Memory accessibility was measured to address the notion that sleep reinforces newly acquired memories, and a separate measure of memory fidelity was acquired to address the notion that sleep refines newly acquired memories. The study was conducted using a crossover design such that all participants completed the task twice, once while assigned to the 90-min nap condition and once while assigned to the wake control condition. The order of conditions was counterbalanced and predetermined by a randomized list constructed before the study. It was predicted that test performance would be greater following nap retention than wake retention for both performance measures. It was further predicted that memory-associated properties of NREM sleep such as the

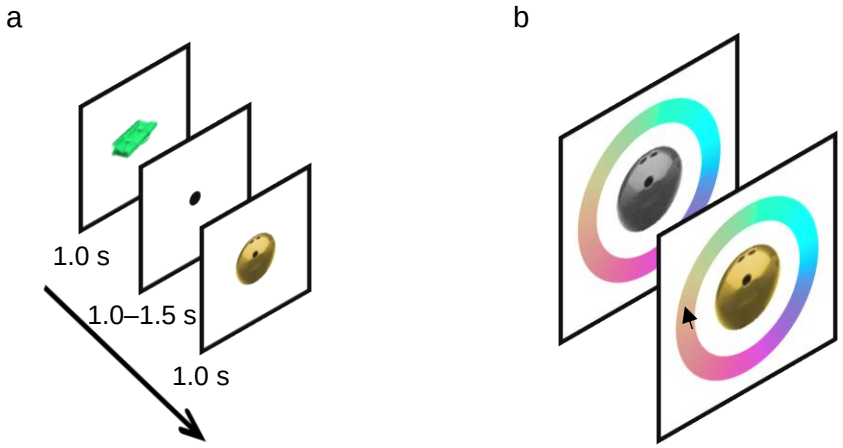

Figure 2.1. Item-colour task trials. **a.** Two consecutive learning trials each displaying an object tinted in one colour selected randomly from a wheel of 360 colours. **b.** Test trial probing an object from the learning period. Objects were shown desaturated of colour at the start of each trial. As the participant moved the computer mouse, the onscreen object was continuously updated to be tinted in the colour dictated by the current position of the pointer in respect to the colour wheel.

amount of stage N3 slow wave sleep, slow-wave EEG activity, and sleep spindles would positively associate with increases in memory accessibility attributable to the nap, and that the amount of stage R sleep would positively associate with increases in memory fidelity attributable to the nap.

Method

Participants

The study was completed by 39 young adults from the Brock University community who were recruited to participant in a study examining the effect of napping on cognitive and emotional performance. All participants reported being good sleepers with typical daily sleep of approximately seven or more hours roughly within 22:00 and 8:00 and did not report working late or overnight shifts frequently in the past six months or traveling over multiple time zones in the past three months. Participants reported being healthy with no history of psychiatric or neurological conditions and no current medications. Participants reported being non-smokers with only low-to-moderate daily caffeine consumption (approximately <300 mg per day). Participants either had English as their first language or had learned English before eight years of age, and all participants were right-handed and had normal or corrected-to-normal vision as well as normal colour vision. All participants were screened for evidence of depression, excessive sleepiness or fatigue, and sleep disorders through a package of questionnaires completed in the laboratory, including a sleep/wake and health questionnaire, the Beck Depression Inventory (Beck et al., 1961), the Epworth Sleepiness Scale (Johns, 1991), and a Fatigue questionnaire (Yoshitake, 1978). On the night prior to two laboratory visits, participants were asked to sleep from 23:00 to 7:00 and to abstain from caffeine, alcohol, vigorous exercise, and napping. All participants provided informed consent and received

a $60 honorarium for their participation.

The item-colour task was challenging, and some participants had to be completely removed from analyses due to poor performance on both sessions. See Data Analysis section for detailed explanation of the exclusions. In total, seven participants were completely removed from analyses of the item-colour task. Thus, the sample included 32 participants (17 female) with a mean age of 20.1 (*sd* = 2.9, minimum = 17, maximum = 26, Q_1 = 18, Q_2 = 19, Q_3 = 22). Seventeen participants (53.1%) were assigned to the nap condition before the control condition.

In the 90-min nap opportunity, some participants obtained very little or no REM sleep. A *post-hoc* grouping of participants based on the amount of stage R scored was used to examine the potential influence of REM sleep on memory performance. A REM+ group was defined as participants obtaining at least 10 stage R epochs (i.e., 5 min), and a REM- group was defined as participants obtaining less then 10 stage R epochs. Seventeen participants (53.1%) were grouped as REM+.

Item-Colour Memory Task

The item-colour task was programmed in PsychoPy, an open-source package for Python (Peirce, 2007). Stimuli were presented on a CRT monitor. The sizes of presented stimuli are provided in estimated degrees of visual angle (VA), referring to the size of the arc on the retina subtended by the image at the given distance. Participants completed the task using a standard keyboard and mouse while seated alone at a viewing distance of approximately 60 cm.

The item-colour task was modeled after the memory task used by Brady et al. (2013) to investigate the fidelity of long-term memories. Participants were asked to learn and later report on the colours of objects presented to them. Stimuli were 205 images of

distinct objects selected from Brady et al. (2008). The selected images were those depicting a single object that could be reasonably thought to exist in any colour and would be identified regardless of colour. The task consisted of an approximately ten-minute learning period and a self-paced, typically about fifteen-min test period that was administered approximately two hours after the learning period.

For the learning period, the 205 objects were each presented one at a time centered on a white background for 1.0 s (largest dimension approximately 6º VA) with a 1.0–1.5 s inter-trial interval displaying only a centered fixation dot (0.1º VA) separating each presentation (Figure 2.1a). Each object was completely tinted with one colour selected randomly from a wheel of 360 colours. Participants were asked to pay attention to the colours of the objects because they would be asked to recall the colours during the test. To ensure participants were paying attention to the objects, they were also asked to perform a repeat detection task, pressing the space bar whenever the same object was presented twice consecutively. Twenty-five of the 205 objects were presented twice consecutively with only the inter-trial interval separating the presentations. These repetitions were randomly distributed throughout the learning period, and repeated objects were not included in the later test. For feedback, participants were shown a centered circle (2.5º VA) for 1.0 s after object presentations for which a response was detected and before the regular inter-trial interval. The circle was green if a repetition was correctly identified or red in the case of a false alarm. A self-paced break was given after every 30 trials.

For the test period, 180 trials probed the memory for each of the 180 objects not repeated during the learning period. Each trial started with one object centered on the screen, desaturated of colour and surrounded by the colour wheel of possible tints

(diameter 15º VA) randomly rotated 0–359º (Figure 2.1b). Participants were asked to recall the object's colour from the learning period and use the computer mouse to report the appropriate tint from the colour wheel to match its previously presented colour. As participants moved the pointer around the screen, the colour of the object was continuously updated to be tinted in the colour dictated by the current position of the pointer in respect to the colour wheel. Participants progressed to the next object and trial via mouse click to report the currently applied colour as their response. The distance between the correct colour and the participant's response on the colour wheel was recorded as a continuous measure of error in degrees. Onscreen feedback of "OK," "Good," "Great," or "Perfect" was given for 1.0 s on a blank screen between trials when error was less than or equal to 15°, 10°, 5°, and 0° respectively. No feedback was given between trials when error was greater than 15º.

Electrophysiological Recording

Electrophysiology was recorded using Neuroscan SynAmps2 amplifiers with Scan 4.5.1 software (Compumedics Inc., Abbotsford, Australia). Recordings were conducted using gold-plated silver electrodes sampling at a rate of 1000 Hz filtered DC to 200 Hz with an additional notch filter at 60 Hz. EEG was recorded at 12 scalp sites (F3, Fz, F4, C3, Cz, C4, P3, Pz, P4, PO7, Oz, and PO8) placed according to the International 10-20 system (Pivik et al., 1993) with an online recording reference at Fpz, a ground reference placed at Afz, and additional offline reference electrodes at sites M1 and M2. To produce both a horizontal and vertical EOG signal, EOG was recorded on four electrodes: two 1 cm away from each outer canthus on the bicanthal plane, one 2 cm directly above the left orbit, and one 2 cm directly below the left orbit. A bipolar submental EMG channel was recorded with two electrodes placed under the chin. Electrical impedances were below 5

KΩ at all scalp sites and below 10 KΩ at all peripheral sites prior to recording.

Procedure

All procedures were cleared by the Brock University Bioscience Research Ethics Board

Recruitment, screening, and orientation. Participants were recruited primarily from the Brock University community through print and digital advertisements and classroom visits. Before enrollment in the study, interested individuals were asked, typically over the phone, questions regarding sleep and health in a 5–10-min interview for an initial check for eligibility. Those who were willing and eligible came to a 30-min laboratory session to give informed consent, perform the Ishihara test of colour deficiencies (Ishihara, 2014), and respond to detailed questionnaires regarding sleep behaviour and health. Willing participants who met eligibility criteria after the additional screening measures participated in a demonstration of the study tasks. The item-colour task demonstration consisted of an abbreviated learning period of only six randomly selected objects, two of which were repeated as an attention check, and four test trials about five minutes later probing memory for the non-repeated objects. Between learning and test, participants were given demonstrations of other tasks not reported here, including a motor sequence memory task, a visual short-term memory task, an emotional detection task with face stimuli, and a threat-detection task using threatening and non-threatening scenes as stimuli. At the end of the orientation and screening session, participants were scheduled for two experimental sessions.

Experimental sessions. Over the two experimental sessions, participants completed both a nap condition and a wake control condition in a random crossover design. The second experimental session took place 1–4 weeks after the first session with

the two sessions being scheduled exactly 1 week apart for most participants. Participants arrived at the laboratory at 13:00 and immediately underwent electrode application. At 13:30, participants were directed to their assigned bedroom for the learning portion of the two memory tasks. This learning period consisted of the learning period of a motor sequence task not reported here followed by the learning period of the item-colour task at approximately 13:40. The motor sequence task is not reported here because features of the task design and participant performance made it unsuitable for testing the SR2 hypothesis. More specifically, varied approaches to the inherent speed-accuracy trade-off made it difficult to obtain distinct measures of memory accessibility and memory fidelity. After the learning period, participants were told the condition to which they were assigned for the retention period, either the 90-min nap opportunity or wake. To minimize potential differences arising from expectations about condition assignment, participants were told that on both days they would be randomly assigned to one of three possible conditions: a short 20-min nap, a long 90-min nap, or no nap at all.

In the nap condition, participants were directed to bed immediately after the learning period of the item-colour task. Lights out occurred at 14:00. Participants were left undisturbed to sleep until 15:30, at which point they were woken by a researcher and escorted to a lounge with a selection of tabletop games to play with a researcher until 16:00. In the wake condition, participants were escorted to the lounge at 14:00 to play tabletop games with a researcher until 16:00. In both conditions, water and a light snack were offered to participants at 15:40. To reduce the amount of learning and interference that could take place during the retention period, participants were not allowed to play games they did not already know how to play.

At 16:00, participants resumed the study tasks, starting with the test trials of the

item-colour task. Following these test trials, participants were administered the remainder of the motor sequence memory task and the other tasks included in the study. Participants left the laboratory by 18:00.

Data Analysis

Data analysis scripts for this study are available from MacDonald (2020) via the Open Science Framework.

Memory performance measures. The outcome variables were defined to test the prediction that NREM and REM sleep processes would separately and respectively influence memory accessibility through reinforcement and memory fidelity through refinement. Memory accessibility and memory fidelity in the item-colour task were thought to be reflected, respectively, in the frequency in which reported colours approximated the target colours and the extent to which such reports were accurate to the exact target colours. Thus, for each participant and test, distributions of response errors (i.e., distance between report and target for each trial) were used to estimate recall percent, the percent of test trials for which at least the approximate target colour was recalled, and recall SD, the standard deviation of the distribution of response errors excluding those for which the approximate target colour was not recalled. These measures were estimated under the guiding premise that a distribution of response errors can be characterized by a mixture model including a uniform distribution of guessing error when there is a failure of recall and a normal (or circular normal) distribution of response errors centered around the target when recall is successful (Zhang & Luck, 2008). Tools developed for estimating the parameters of such mixed distributions in visual short-term and working memory data have been developed (Suchow et al., 2013). A two-step method was modeled after these examples, modified to better fit the data and task

demands of maintaining many representations over longer periods of time.

The first step to determining recall percent and recall SD was to calculate a preliminary recall proportion score using a criterion of defined recall failure set at 90º such that all responses greater than 90º away from the target would be classified as a failure of recall. For each distribution of response errors, the number of responses outside this criterion was counted and divided by the number of possible response options outside this criterion (i.e., 179) to obtain the number of responses per degree in the region defined as recall failure. Because the distribution of response errors during recall failure (i.e., guessing) is assumed to be uniform, this quotient was considered the height of a uniform distribution of guessing errors spanning the full 360º range of possible responses. The height of the uniform distribution was multiplied by 360 to get the approximated number of recall failures and then divided by the total number of responses to obtain an approximation of the proportion of trials for which recall failed and the complementary proportion of trials for which recall was successful. For example, if 30 of 180 responses were found to be outside of the 90º criterion, the approximated recall proportion would be 0.6648. This method of approximation assumes that the density of guesses outside the criterion is reflective of the overall number of guesses. However, there is uncertainty in this estimate. Simulations determined this method to be reliably within ±0.01 of known guess proportions for a wide range of possible guess proportions.

Consideration of multiple factors led to setting the defined recall failure criterion at 90º. First, in examination of the all-participant distribution of response errors, it was observed that the centered normal distribution appeared to meet the height uniform distribution at about ±45–60º, well short of ±90º. Second, a 90º shift from a given colour on the colour wheel returns a colour of a distinctly different category (e.g., blue to either

green or pink). Third, a larger defined recall failure criterion, such as 120º, would reduce the confidence in the uniform distribution height estimate given the smaller sample window. Thus, while it is possible that some responses with error greater than 90º reflect very low fidelity recall rather than total recall failure, the frequency of these responses appears to be very low, these responses would appear categorically incorrect, and their inclusion would come at a slight cost to reliability of the recall percent estimate.

The second step to obtaining recall percent and recall SD from the error distribution was to use the preliminary recall proportion and maximum likelihood estimation to fit a mixed normal and uniform probability density function to the distribution of response errors. This mixed distribution function could vary on the following parameters: the mean of the normal distribution (*mn*), the standard deviation of the normal distribution (*sdn*), and the proportion of samples taken from the normal distribution of recall success as opposed to the uniform distribution of recall failure (*rp*). Parameter *mn* was restricted to ±2.5º to allow for any slight shifts of the mean response error away from 0º. Parameter *sdn* was restricted to a wide range of 0.1–200º simply to allow the model to converge. Parameter *rp* was restricted to the previously approximated proportion of recall successes ±0.01 to account for imprecision in the approximation. Maximum likelihood estimation was conducted using *stats4* in R (R Core Team, 2017) and used to determine these parameters three separate times. The starting values for *mn* and *rp* were respectively 0º and the approximated proportion of recall success for each of the three runs; 1º, 20º, and 100º were used as starting values for *sdn* across the three runs of maximum likelihood estimation. The parameters from the best fitting of the three resulting functions, as determined by loglikelihood value, were selected. The performance measures of interest, recall percent and recall SD, were thus defined as the selected *rp*

(multiplied by 100) and *sdn*, respectively.

Sleep measures. PSG records from the 90-min nap opportunity were subjected to several procedures to obtain measures of sleep architecture, EEG power spectra, and sleep spindles.

Sleep scoring. Sleep stages and arousals were scored in 30-s epochs by K. M. according to American Academy of Sleep Medicine criteria (Berry et al., 2015). Lateral EEG channels were referenced to contralateral mastoids. A bipolar horizontal EOG channel was created from the two outer canthi EOG electrodes, and a bipolar vertical EOG channel was created from the EOG electrodes above and below the left eye. EEG and EOG channels received a 0.1–35 Hz bandpass filter with 12 dB per octave roll-off. The EMG channel received a 10–100 Hz bandpass filter with 12 dB per octave roll-off. Scoring was conducted while viewing F3, F4, C3, C4, PO7, PO8, horizontal EOG, vertical EOG, EMG, M1, and M2 in Neuroscan. For each participant, time spent in each sleep stage, total time spent asleep, the percent of sleep time spent in each stage, and the total number of arousals (arousals plus transitions to wake from sleep) were obtained from the scored files.

EEG power spectra. Pre-processing steps for obtaining EEG power data from the sleep records included re-referencing all EEG channels to the average of the two mastoid channels and applying a 0.5–100 Hz bandpass filter with 12 dB per octave roll-off using Neuroscan software. Using MNE, a Python toolbox used to analyze EEG data (Gramfort et al., 2013), each sleep file was resampled to 256 Hz. Epochs of sleep preceded or proceeded by an epoch of wake or stage N1 sleep were excluded from power spectral analyses along with stage R epochs adjacent to either stage N2 or stage N3 epochs as well as any stage N2 or stage N3 epochs adjacent to a stage R epoch. For epochs containing an

arousal, the largest continuous section of EEG unmarked by that arousal was subjected to power spectral analysis if it was at least 10 s in duration. These exclusions ensured epochs containing transitions between stages were not subjected to EEG power spectrum analysis. If the largest artifact-free section of data was less than 10 s in duration, the entire epoch was excluded. In addition, rare sections of bad data (e.g., from a lost electrode) were identified through visual inspection in Neuroscan, and data from the affected channels at the affected times were removed. Power spectral density was computed for each channel in each eligible epoch using Welch's method in MNE with 2-s, 75% overlapping windows, obtaining a frequency resolution of 0.5 Hz. The frequency spectrum was divided into the delta (1–3.5 Hz), theta (4–7.5 Hz), alpha (8–11.5 Hz), sigma (12–15.5 Hz), and beta (16–29.5 Hz) bands. Absolute power per unit of frequency (μV^2/Hz) was obtained for each frequency band, channel, and epoch. Averages combining all eligible stage R epochs gave measures of absolute power in R sleep and averages combining all eligible stage N2 and stage N3 epochs gave measures of absolute power in N2 and N3 sleep combined (N2/3 sleep). Full scalp measures were then obtained by averaging over each available channel. Each of these measures then received logarithmic (base 10) transformation. Tables and figures reporting descriptive statistics of average EEG power measures or analyses including average EEG power measures used full scalp measures unless otherwise specified.

Spindle detection. Spindles were identified through visual inspection of EEG in stages N2 and N3 sleep at site Cz. To aid identification, a wide band-pass (0.5–35 Hz) and narrow band-pass (12–16 Hz) filter were applied separately to two copies of the EEG at Cz, and both filtered channels were examined simultaneously. Spindles were defined as bursts of 12–16 Hz EEG activity lasting at least 0.5 s and consisting of a fusiform shape.

Spindle count (i.e., the number of spindles) and spindle density (i.e., the number of spindles per minute of sleep in stages N2 and N3) were recorded.

Statistical analyses. Statistical analyses were conducted in R (R Core Team, 2017) using base functions and a variety of specialized packages, including *reshape2* (Wickham, 2007) for handling data structures and *ggplot2* (Wickham, 2016), *grid* (R Core Team, 2017), and *gridExtra* (Auguie, 2016) for creating figures.

Data exclusion and adjustment. For statistical analysis and parameter estimation, effort was made to maximize the data included for each analysis while excluding data that were problematic by nature of poor performance or by extreme deviance from the rest of the sample. Rather than excluding individuals from all analyses for poor performance, participant scores were removed for the specific measures (recall percent or recall SD) and specific performances (first session and second session) for which their score was deemed problematic. Extreme values were not broadly excluded from analyses. Instead, checks for overly influential cases were conducted for each analysis and exclusions or adjustments occurred for influential cases on an analysis-by-analysis basis. The method for these exclusion and adjustment procedures are reported here.

When test performance in the item-colour task for a session was deemed to not be greater than chance, data from that session was excluded from all analyses of memory performance. Greater than chance performance was defined as having either significantly more responses within either 20º or 60º of the target than would be expected by random guessing at α of .05 (i.e., at least 28 responses within 20º or 71 responses within 60º). These two chance criterions were selected to ensure retention of two types of non-random response patterns: 1) very few items remembered with high accuracy and 2) some items remembered with poor accuracy. Seven participants were excluded from all analyses and

statistics because they did not perform significantly greater than chance on either session. Another two participants had only the data from their first session was excluded (one nap and one control), and another two participants had only the data from their second session was excluded (both control). Alternatively, of the 39 participants who completed the study, 32 were included in analyses because they performed above chance levels on at least one session. Of these 32 participants, 28 contributed recall percent scores for both sessions, and 4 participants contributed a recall percent score for only one session: a total of 60 recall percent scores.

Even when performance was greater than chance, recall SD scores were excluded from analyses if their corresponding recall percent was estimated to be less than 16.67% (i.e., 30 recall successes over 180 trials). Recall SD scores were excluded in these cases because an estimated standard deviation of a normal distribution of response errors was considered unreliable if based on less than 30 trials with recall success. Of the 28 participants above chance level performance for both sessions, 8 had both of their recall SD scores excluded, 6 had only their first session excluded (3 nap and 3 control), and 3 had only their second session excluded (1 nap and 2 control). None of the four participants contributing only one recall percent scores contributed a recall SD score. In summary, 11 participants contributed a recall SD score for both sessions, and 9 participants contributed a recall SD score for only one session: a total of 31 recall SD scores from 20 participants.

All additional adjustment or removal of data was done to fit regression models without influential cases. Cook's distance was used to identify influential cases (Christensen et al., 1992; Cook, 1977) and was calculated using *HLMdiag* (Loy & Hofmann, 2014). A participant was an influential case in a model if they had a Cook's

distance greater than both 0.25 and the mean Cook's distance multiplied by 3. When a participant was identified as such, their data was examined for extreme values on variables included in the model. An extreme value was one that deviated from the mean value by more than 1.5 times the interquartile range (1.5 IQR) of the sample full sample. When a participant was both influential in a model and had extreme values on predictors in the model, these extreme values were winsorized to 1.5 IQR, and the models were refit with the adjusted data. When a participant remained influential in a model after winsorization, when a participant was influential in a model despite not having an extreme value, or when their extreme value was on a memory performance measure, the model was refit with that participant removed. This process was repeated for a model until it was fit without influential cases, in which was the final adjusted model was considered, or until the number of excluded data points on the outcome measure was greater 10% of the number of included data points on the outcome measure, in which case only no final model was obtained. All regression models were subjected to this process, but not all these analyses are reported. For regression models predicting performance measures, the models fit without influential cases were the focus of follow-up analyses and evaluations of statistical significance. For tests of group differences on sleep measures, results from unadjusted data are reported. When effects or associations are reported as significant in models fit to adjusted data, the specific exclusions and adjustments are reported.

Statistical tests. Linear mixed-effect regression models were used to examine how sleep properties and experimental conditions (i.e., session or condition) associated with memory performance. Mixed-effect models were constructed using *lme4* (Bates et al., 2015). In all such models, a random intercept was allowed for each participant. Simple,

non-mixed-effect regression models were used when each participant offered only one data point for each measure. The *oneway.test* function was used to compare group means (e.g., REM+ vs. REM-); degrees of freedom were adjusted using the Welch method (Welch, 1951) when Levene's test indicated the groups had unequal variances. Tests of associations, or "effects", in simple linear regression were conducted using the *Anova* function in the *car* package (Fox & Weisberg, 2011) to perform analyses of variance *F*-tests using type II sums of squares. For linear mixed-effect models, the same function was used to perform type II Wald *F*-tests with denominator degrees of freedom approximated through the Kenward-Roger method (Kenward & Roger, 1997). All tests of interaction terms were conducted in models containing the corresponding lower-order terms. Simple effects and slopes were estimated at the 25^{th} and 75^{th} percentiles of the interacting predictor variables. Continuous predictors in regression models were standardized ($m = 0$, $sd = 1$). *B* values (i.e., regression coefficients) from standardized predictors on unstandardized outcomes are reported to show the magnitude and direction of effects from linear models. Categorical variables session and condition were effect coded such that coefficients indicated the effect of session 2 compared to session 1 and the effect of a nap compared to control.

Statistical significance was defined through *p* value less than α of .05 for tests concerning hypotheses and measures of considerable practical or theoretical interest. More specifically, this was the only significance criterion when testing the effect of session or condition and when testing the association between memory performance and the following sleep measures: total sleep duration, N3 duration, R duration, and REM group (+/-). Other theoretically interesting measures were considered as sets of related measures. One set included N2/3 delta power at each channel and the full scalp measure,

and another set included spindle count, spindle density, and N2/3 sigma power at each channel and the full scalp measure. Within these defined sets of measures, false discovery rate was controlled for using the Benjamini–Hochberg method (Benjamini & Hochberg, 1995) to limit false discovery rate to 0.10. Thus, for these measures, statistical significance was defined as having a *p* value less than .05 and the critical *Q* determined to limit false discovery to 0.10. Tests on recall percent were always considered in a different set than tests on recall SD. Alpha was lowered to .01 for tests of other sleep measures, including total sleep duration, N1 duration, N2 duration, the percent of total sleep duration spent in each sleep stage, arousal count, arousal density, average EEG power in other frequency bands of N2/3 sleep, and average EEG power in stage R sleep. For these EEG power measures, the Benjamini–Hochberg method was applied to limit false discovery rate to .05 within each similarly constructed set with a separate set for each remaining frequency band within stage N2 sleep and stage N3 sleep and each frequency band within stage R seep. Brief explorations of results of marginal significance, defined as those meeting the criterion of limit false discovery but exceeding α by no more than 0.04, are reported.

Due to the repeated-measure design, participant memory performance could differ significantly between experimental sessions based on their level of experience with the task. The effect of session was examined in regression models to quantify this effect. When testing the effect of the condition or testing sleep measures as predictors of performance or moderators of a condition effect on performance, models including a session term were examined to statistically control for possible practice effects. Models without the session term were also examined. Because the session in which participants napped was counterbalanced, the effect of including the term was generally minimal. For

simplicity, only the models with this statistical control are reported here.

Results

Sleep

Overall, the 90-min nap opportunities were well-used; only three participants obtained less than 20 min of sleep and no participant obtained less than 10 min of sleep. Descriptive statistics for sleep architecture, spindle count, and spindle density are reported in Table 2.1, and descriptive statistics for average EEG power measures are reported in Table 2.2. Approximately half of the participants obtained at least 10 stage R epochs and were hence grouped as REM+. There were 15 REM+ participants and 17 REM- participants. The REM+ group had a greater total sleep duration, $F(1, 17) = 9.32$, $p = .007$, a greater duration of stage R sleep, $F(1, 16) = 142.88$, $p < .001$, a greater percent of sleep time spent in stage R, $F(1, 18) = 141.34$, $p < .001$, and less arousal density, $F(1, 30) = 6.09$, $p = .020$, than the REM- group, but the groups did not significantly differ from REM- participants in duration of stage N1 sleep, $F(1, 30) = 0.03$, $p = .855$, stage N2 sleep, $F(1, 30) = 0.07$, $p = .795$, or stage N3 sleep, $F(1, 30) = 0.74$, $p = .398$; percent of sleep time in stage N1, $F(1, 22) = 3.36$, $p = .080$, stage N2, $F(1, 30) = 3.06$, $p = .090$, or stage N3, $F(1, 30) = 0.25$, $p = .619$; or arousal count, $F(1, 30) = 1.26$, $p = .272$. The REM+ group had more spindles, $F(1, 30) = 9.45$, $p = .005$, and greater spindle density, $F(1, 30) = 7.46$, p = .010, than the REM- group. These groups did not significantly differ on full scalp measures of N2/3 delta power, $F(1, 30) = 0.77$, $p = .388$, theta power, $F(1, 26) = 0.23$, $p = .633$, alpha power, $F(1, 30) = 0.41$, $p = .527$, sigma power, $F(1, 30) = 0.28$, $p = .600$, or beta power, $F(1, 30) < 0.01$, $p = .949$. Pearson correlations among these sleep measures are depicted in Figure 2.

Table 2.1
Descriptive statistics for sleep architecture measures in each participant grouping

Sample	Measure		n	*m*	*sd*	min.	max.	Q_1	Q_2	Q_3
All	Sleep dur.	*	32	69.70	19.98	10.50	87.50	65.75	75.75	82.00
	N1 dur.		32	11.06	6.74	2.50	37.00	7.00	9.00	13.75
	N2 dur.		32	36.72	17.57	6.00	74.00	23.50	33.75	46.75
	N3 dur.		32	15.83	15.04	0.00	45.00	2.25	10.25	30.00
	R dur.	*	32	6.09	6.42	0.00	18.00	0.00	3.00	11.75
	N1 %		32	18.33	12.90	3.45	52.94	9.92	12.92	22.53
	N2 %		32	52.57	17.22	23.84	89.70	37.21	54.08	61.08
	N3 %		32	21.37	20.30	0.00	59.60	3.02	12.33	39.33
	R %	*	32	7.73	8.00	0.00	23.68	0.00	4.79	14.68
	Spindles		32	192.62	107.48	9.00	453.00	95.00	192.50	272.50
	Spindle den.		32	3.66	1.62	1.06	7.02	2.20	4.06	4.86
	Arousals		32	7.25	5.00	1.00	26.00	4.00	6.00	9.00
	Arousal den.	*	32	0.11	0.08	0.01	0.39	0.06	0.08	0.13
REM+	Sleep dur.		15	79.43	4.48	72.50	86.00	74.50	80.00	83.50
	N1 dur.		15	11.30	5.24	2.50	23.50	7.50	10.00	14.50
	N2 dur.		15	37.60	11.93	20.00	55.00	24.50	43.00	45.50
	N3 dur.		15	18.27	13.83	0.00	40.00	8.50	13.00	34.50
	R dur.		15	12.27	3.60	6.50	18.00	8.50	13.00	15.00
	N1 %		15	14.21	6.74	3.45	30.92	10.14	12.66	19.19
	N2 %		15	47.08	14.00	27.59	67.90	31.54	52.33	58.90
	N3 %		15	23.31	18.29	0.00	55.17	10.18	15.95	40.83
	R %		15	15.40	4.40	8.72	23.68	11.64	15.57	20.12
	Spindles		15	247.73	67.06	76.00	326.00	232.00	264.00	287.00
	Spindle den.		15	4.41	1.11	2.20	6.33	4.12	4.44	5.10
	Arousals		15	6.20	2.68	1.00	12.00	4.00	6.00	8.00
	Arousal den.		15	0.08	0.03	0.01	0.14	0.05	0.07	0.10
REM-	Sleep dur.		17	61.12	24.27	10.50	87.50	62.00	66.50	76.00
	N1 dur.		17	10.85	7.99	3.50	37.00	6.50	8.50	12.50
	N2 dur.		17	35.94	21.73	6.00	74.00	23.50	32.50	48.00
	N3 dur.		17	13.68	16.13	0.00	45.00	1.00	5.50	23.00
	R dur.		17	0.65	1.16	0.00	4.00	0.00	0.00	1.50
	N1 %		17	21.97	15.90	5.56	52.94	9.71	16.56	28.12
	N2 %		17	57.41	18.70	23.84	89.70	47.06	57.14	70.00
	N3 %		17	19.66	22.34	0.00	59.60	1.32	10.67	37.10
	R %		17	0.96	1.78	0.00	6.35	0.00	0.00	1.82
	Spindles		17	144.00	114.39	9.00	453.00	75.00	111.00	179.00
	Spindle den.		17	2.99	1.73	1.06	7.02	1.64	2.42	4.00
	Arousals		17	8.18	6.35	1.00	26.00	4.00	6.00	11.00
	Arousal den.		17	0.14	0.09	0.06	0.39	0.08	0.10	0.20

* measures for which the REM+ and REM- groups differed significantly ($\alpha = .05$)
Note. Duration (dur.) reported in minutes. Density (den.) reported as count per minute of sleep (arousal density) or minute of stage N2 sleep and stage N3 sleep (spindle density).

Table 2.1
Descriptive statistics for EEG power measures in each participant grouping

Sample	Measure	n	*m*	*sd*	min.	max.	Q_1	Q_2	Q_3
All	N2/3 delta	32	-9.92	0.32	-10.42	-9.15	-10.13	-10.02	-9.65
	N2/3 theta	32	-10.90	0.21	-11.27	-10.50	-11.08	-10.94	-10.73
	N2/3 alpha	32	-11.35	0.23	-11.72	-10.62	-11.48	-11.35	-11.20
	N2/3 sigma	32	-11.50	0.22	-11.85	-11.01	-11.73	-11.47	-11.35
	N2/3 beta	32	-12.46	0.18	-12.74	-11.98	-12.57	-12.52	-12.35
	R delta	20	-10.40	0.44	-10.93	-8.72	-10.60	-10.45	-10.30
	R theta	20	-11.13	0.30	-11.46	-10.08	-11.34	-11.18	-11.03
	R alpha	20	-11.45	0.20	-11.73	-11.01	-11.63	-11.48	-11.40
	R sigma	20	-11.90	0.30	-12.24	-10.85	-12.03	-11.97	-11.88
	R beta	20	-12.30	0.46	-12.72	-10.48	-12.49	-12.40	-12.32
REM+	N2/3 delta	15	-9.87	0.24	-10.30	-9.47	-10.08	-9.92	-9.60
	N2/3 theta	15	-10.88	0.15	-11.10	-10.67	-11.00	-10.86	-10.73
	N2/3 alpha	15	-11.38	0.18	-11.70	-11.08	-11.47	-11.40	-11.21
	N2/3 sigma	15	-11.48	0.18	-11.82	-11.25	-11.57	-11.44	-11.31
	N2/3 beta	15	-12.46	0.18	-12.72	-12.03	-12.58	-12.47	-12.36
	R delta	15	-10.37	0.50	-10.93	-8.72	-10.65	-10.46	-10.27
	R theta	15	-11.15	0.32	-11.46	-10.08	-11.37	-11.20	-11.10
	R alpha	15	-11.50	0.19	-11.73	-11.09	-11.64	-11.49	-11.41
	R sigma	15	-11.91	0.32	-12.24	-10.85	-12.03	-11.97	-11.88
	R beta	15	-12.28	0.52	-12.72	-10.48	-12.50	-12.39	-12.31
REM-	N2/3 delta	17	-9.97	0.38	-10.42	-9.15	-10.24	-10.08	-9.70
	N2/3 theta	17	-10.92	0.26	-11.27	-10.50	-11.13	-10.97	-10.73
	N2/3 alpha	17	-11.32	0.27	-11.72	-10.62	-11.49	-11.29	-11.20
	N2/3 sigma	17	-11.52	0.25	-11.85	-11.01	-11.76	-11.49	-11.37
	N2/3 beta	17	-12.47	0.19	-12.74	-11.98	-12.57	-12.53	-12.33
	R delta	5	-10.47	0.21	-10.80	-10.25	-10.53	-10.42	-10.34
	R theta	5	-11.08	0.26	-11.46	-10.83	-11.22	-10.97	-10.93
	R alpha	5	-11.32	0.21	-11.51	-11.01	-11.45	-11.44	-11.20
	R sigma	5	-11.86	0.25	-12.05	-11.45	-12.02	-11.98	-11.81
	R beta	5	-12.36	0.23	-12.55	-11.97	-12.49	-12.40	-12.37

Note. REM+ and REM- groups did not differ significantly (α = .05) on combined stage N2 and N3 measures; groups were not compared on stage R measures.

Memory Performance

Of all included performances (60), participants approximated the target colour on 21.66% of test trials (*sd* = 16.24, minimum = 2.66, maximum = 78.89, Q_1 = 11.06, Q_2 = 17.75, Q_3 = 26.70). The mean recall SD of the 31 performances in which at least 30 target

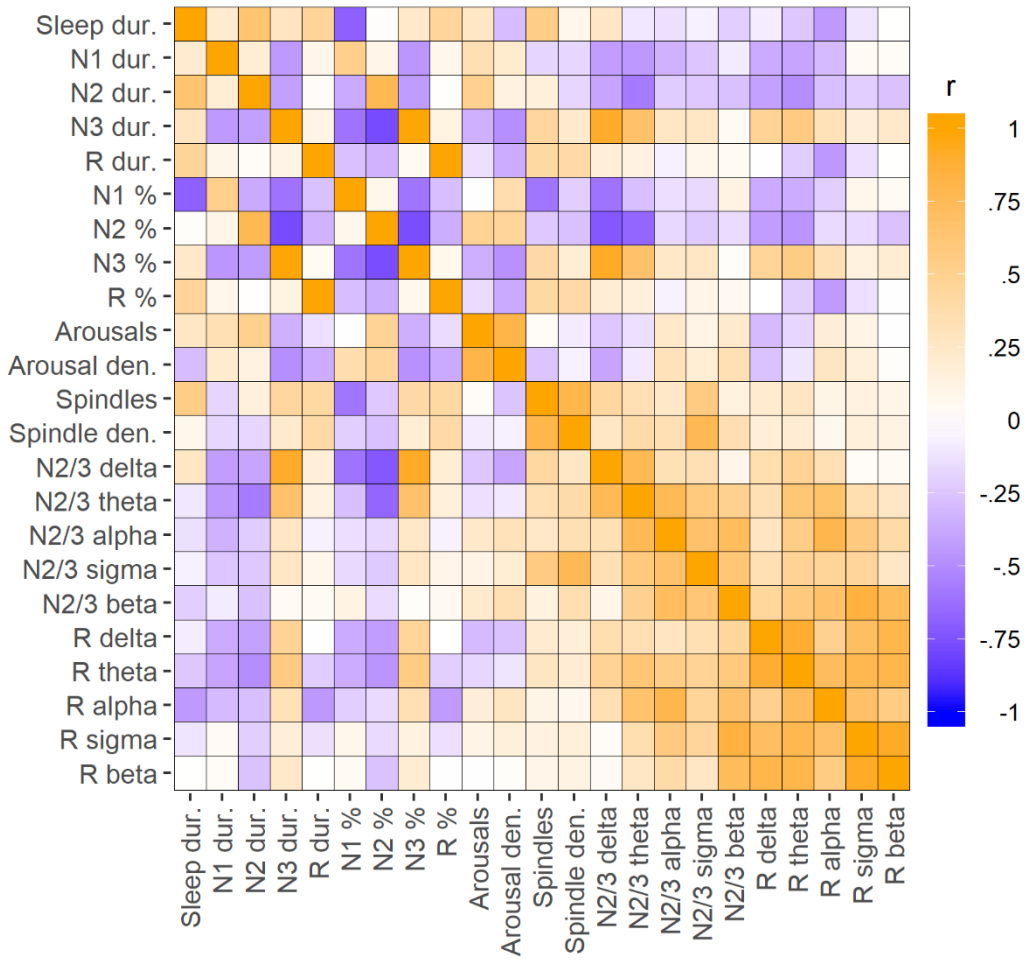

Figure 2.2. Pearson correlation matrix of sleep measures with *r* values depicted on a coloured scale. Dur. = duration. Density (den.) measured as count per minute of sleep (arousal density) or stage N2 sleep and stage N3 sleep (spindle density). Pairwise deletion was used.

colours were approximated (i.e., recall percent ≥ 16.67%) was 24.20º (*sd* = 8.60, minimum = 10.43, maximum = 47.77, Q_1 = 19.61, Q_2 = 24.24, Q_3 = 28.68). Descriptive statistics for recall percent and recall SD scores, by session and condition are reported in Table 2.3. The distributions of response errors for each session and experimental condition are displayed in Figure 2.3. A mixed-effect model predicting recall percent

Table 2.3
Descriptive statistics for item-colour task performance

Measure	n	*m*	*sd*	min.	max.	Q_1	Q_2	Q_3
S1 recall %	30	18.80	15.60	2.66	71.07	9.38	14.53	24.14
S1 recall SD	14	27.64	9.33	14.20	47.77	22.82	26.29	28.93
S2 recall %	30	24.52	16.62	7.15	78.89	13.85	20.67	33.08
S2 recall SD	17	21.37	6.99	10.43	33.61	15.05	20.94	25.78
Control recall %	29	23.06	17.42	3.00	78.89	12.73	18.07	28.61
Control recall SD	15	25.86	8.38	13.74	43.26	19.61	25.78	30.28
Nap recall %	31	20.35	15.22	2.66	71.07	9.38	17.44	24.14
Nap recall SD	16	22.65	8.77	10.43	47.77	18.58	21.88	26.29

Note. S1 = session 1. S2 = session 2. Recall percent and recall SD differed significantly ($\alpha = .05$) from S1 to S2. No significant effects of condition (control vs. nap) were detected. Recall SD measured in degrees of error.

from recall SD identified no significant association between the two measures in an unadjusted sample, $F(1, 16) = 0.64$, $p = .437$, $B = 0.18$, $se = 0.22$, although the sample of partcipants with valid measures of SD was small and a model without influential cases could not be determined without excluding too many cases.

Session 2 vs. session 1. Participants generally performed better on the item-colour task on their second session relative to their first. The effect of session on recall percent was significant, $F(1, 28) = 6.12$, $p = .020$, with an estimated 5.38-point ($se = 2.17$) increase in recall percent between sessions, from 18.39% on their first session to 23.77% on their second session. The effect of session for recall SD was not significant in an adjusted sample, $F(1, 16) = 2.99$, $p = .103$, with an influential case with extremely high recall SD on session 1 (+3.08 IQR) excluded. Recall SD was estimated to decrease 4.72º ($se = 2.68$) between days, from 28.46º on session 1 to 19.01º on session 2.

Nap vs. control. Recall percent was not significantly affected by condition, $F(1, 27) = 0.74$, $p = .397$. Against predictions, the nap, compared to control, was associated with a non-significant 1.88 point ($se = 2.19$) decrease in recall percent. A model without

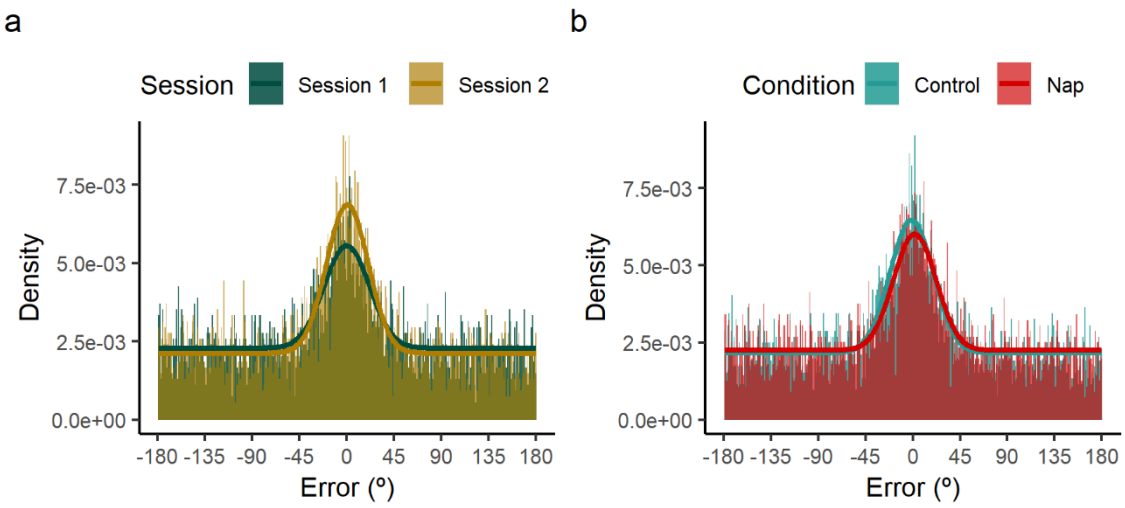

Figure 2.3. Distributions of responses measures as error in angular distance from target colour. Solid lines indicate density of a mixed distribution composed of a uniform distribution of guesses and a normal distribution of successful recall with parameters fit to the corresponding distribution of errors. **a.** Response errors by session. **b.** Response errors by condition.

influential cases could not be fit for recall SD without removing too many cases (two for this sample). After exclusion of two influential cases, one with extremely high recall SD on their nap day (+2.78 IQR) and another with extremely high recall SD on their control day (+1.59 IQR), there were no further influential cases; there was no sign of an effect of condition in this adjusted sample, $F(1, 14) = 1.57$, $p = .232$, $B = -3.05$, $se = 2.38$.

Sleep measures. The following tests examined whether measures of sleep were general predictors memory performance on both sessions or moderators of a condition effect. A session term (session 2 vs. session 1) was included in each regression model as a statistical control.

Predicting performance on both sessions. Regression models were constructed to

Table 2.4
Tests of sleep measures as predictors of recall percent on both sessions and as moderators of an effect of nap condition on recall percent

	As Predictor						As Moderator of Condition				
Measure	n	*F*	df_r	*p*	*B*	*se*	n	*F*	df_r	*p*	*B*	*se*
Sleep dur.	32	0.15	29	.701	-1.08	2.78	32^{a1}	0.01	25	.943	0.17	2.33
N1 dur.	31^{r1}	0.53	34	.473	-1.40	1.93	31^{r1}	2.08	31	.159	-3.91	2.67
N2 dur.							32	0.77	25	.389	-1.97	2.25
N3 dur.	31^{r1}	1.15	28	.292	-2.24	2.09	31^{r1}	6.73	25	.016	5.33	2.05
R dur.	32	0.41	29	.527	1.75	2.74	32	0.11	26	.739	-0.76	2.26
REM Grp.	30^{r2}	0.00	27	.997	-0.01	3.10	32	0.05	26	.823	1.05	4.64
Arousals	31^{a1r1}	0.00	29	.989	0.03	2.10	31^{a1r1}	0.29	26	.598	1.26	2.34
Arous. den.	31^{r1}	0.95	28	.337	2.06	2.10	30^{a1r2}	0.29	25	.592	1.27	2.34
Spindles	32	0.05	30	.832	0.59	2.74	32	0.75	26	.393	1.95	2.24
Spin. den.	32	0.72	30	.401	2.29	2.69	32	0.14	26	.716	-0.83	2.25
N2/3 delta	31^{r1}	1.03	30	.318	-2.08	2.04	31^{r1}	5.41	26	.028	5.02	2.15
N2/3 theta	32	0.07	31	.796	0.70	2.67	32	2.77	26	.108	3.74	2.24
N2/3 alpha	32	0.11	30	.739	0.92	2.72	32	0.27	26	.606	1.21	2.32
N2/3 sigma	31^{r1}	0.60	29	.446	1.63	2.11	32	0.12	26	.730	-0.80	2.29
N2/3 beta	32	0.79	30	.381	2.41	2.71	32	0.12	26	.733	-0.78	2.26
R delta	31^{r1}	1.03	30	.318	-2.08	2.04	31^{r1}	5.41	26	.028	5.02	2.15
R theta	32	0.07	31	.796	0.70	2.67	32	2.77	26	.108	3.74	2.24
R alpha	32	0.11	30	.739	0.92	2.72	32	0.27	26	.606	1.21	2.32
R sigma	31^{r1}	0.60	29	.446	1.63	2.11	32	0.12	26	.730	-0.80	2.29
R beta	32	0.79	30	.381	2.41	2.71	32	0.12	26	.733	-0.78	2.26

Note. Dur. = duration. Grp. = group. Density (den.) measured as count per minute of sleep (arousals) or minute of stage N2 sleep and stage N3 sleep (spindles). Numbers after superscript "a" indicate the number of influential cases that had predictor values winsorized for a given test. Numbers after superscript "r" indicate the number of influential cases removed from for a given test. Empty cells are shown for tests for which a model without influential cases could not be fit.

examine whether sleep measures or REM group predicted either recall percent or recall SD on both sessions, regardless of condition. A selection of *F*-tests from these models predicting recall percent and recall SD are reported in Table 2.4 and Table 2.5, respectively.

None of sleep duration, N1 duration, N2 duration, N3 duration, R duration, REM group, spindle density, spindle count, arousal density, and arousal count and no measures of the percent of sleep time spent in a sleep stage or EEG power from either stages N2

Table 2.5
Tests of sleep measures as predictors of recall SD on both sessions and as moderators of an effect of nap condition on recall SD

	As Predictor						As Moderator of Condition				
Measure	n	*F*	df_r	*p*	*B*	*se*	n	*F*	df_r	*p*	*B*	*se*
Sleep dur.	20	0.86	14	.368	1.14	1.21						
N1 dur.	20	0.08	20	.775	0.54	1.83	20	0.13	19	.724	-1.65	4.41
N2 dur.	20	0.25	13	.628	-0.66	1.33	20	0.16	12	.694	1.18	2.86
N3 dur.												
R dur.							19[r1]	0.02	18	.896	0.43	3.16
REM Grp.	19[r1]	2.60	15	.128	4.60	2.79	19[r1]	0.06	16	.817	1.53	6.34
Arousals	20	0.38	12	.551	-0.74	1.19	20[a1]	0.00	11	.966	0.12	2.65
Arous. den.	20	1.28	12	.279	-1.34	1.17	20[a1]	0.72	15	.408	-2.39	2.74
Spindles	19[r1]	3.65	14	.077	2.88	1.47						
Spin. den.	19[r1]	1.25	14	.282	1.79	1.57	20	0.12	14	.736	1.18	3.34
N2/3 delta	19[r1]	0.20	19	.663	0.70	1.54						
N2/3 theta	20	0.11	15	.743	0.58	1.70						
N2/3 alpha	20[a1]	0.04	16	.844	-0.33	1.61						
N2/3 sigma	20	0.07	14	.798	-0.37	1.39	20	0.08	13	.785	-0.82	2.89
N2/3 beta												
R delta	19[r1]	0.20	19	.663	0.70	1.54						
R theta	20	0.11	15	.743	0.58	1.70						
R alpha	20[a1]	0.04	16	.844	-0.33	1.61						
R sigma	20	0.07	14	.798	-0.37	1.39	20	0.08	13	.785	-0.82	2.89
R beta												

Note. Dur. = duration. Grp. = group. Density (den.) measured as count per minute of sleep (arousals) or minute of stage N2 sleep and stage N3 sleep (spindles). Numbers after superscript "a" indicate the number of influential cases that had predictor values winsorized for a given test. Numbers after superscript "r" indicate the number of influential cases removed from for a given test. Empty cells are shown for tests for which a model without influential cases could not be fit.

and N3 sleep or stage R sleep were significant predictors of recall percent on both sessions. Likewise, none of these measures were significant predictors of recall SD on both sessions.

Moderating an effect of condition. Regression models were constructed to examine whether sleep measures or REM group moderated an effect of nap condition relative to control for either recall percent or recall SD. A selection of *F*-tests from these models predicting recall percent and recall SD are reported in Table 2.4 and Table 2.5,

respectively.

For recall percent, N3 duration was found to be a significant moderator of the condition effect as the N3 duration × condition interaction was significant in an adjusted sample excluding one influential case with an extreme recall percent score for the nap condition (+3.63 IQR) and the control condition (+3.17 IQR). At 2.0 min of N3 sleep, the nap was estimated to decrease of recall percent by 6.77 points (*se* = 2.77) relative to control, *t* = -2.51. At 28.0 min of N3 sleep, the nap was estimated increase recall percent by 2.70 points (*se* = 2.70), *t* = 1.00. This interaction was further illustrated by a positive correlation, *r*(25) = .47, *p* = .014, between N3 duration and the nap minus control change in recall percent (Figure 2.4a), though this correlation excluded an additional four

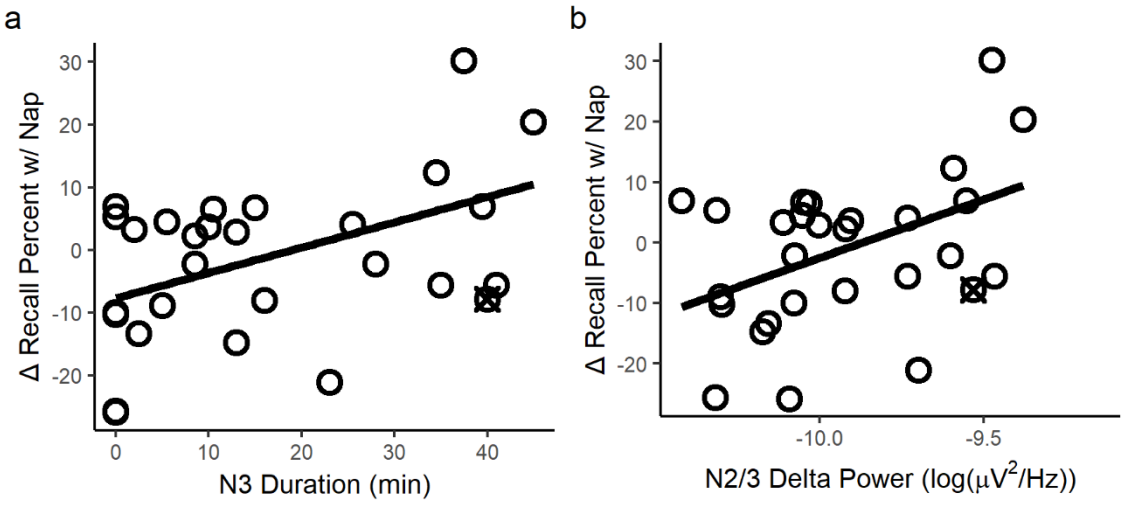

Figure 2.4. Change in recall percent (nap - control) predicted by two measures of NREM sleep slow wave activity. The influential case excluded from linear models in figure and report is marked by the × symbol. **a.** Change in recall percent as a function of N3 duration. **b.** Change in recall percent as a function of average delta band (1–3.5 Hz) electroencephalographic power in stage N2 and N3 epochs.

participants for whom a difference score could not be computed due to them only having one valid recall percent score.

Similarly, the N2/3 delta power × condition interaction effect was significant in predicting recall percent at channels F3, F4, and C4 in the unadjusted sample without influential cases, at channels Cz, P4, Pz, and PO8 and the full scalp measure in an adjusted sample excluding the participant with extremely high recall percent for nap and control, at C3 in an adjusted sample excluding a separate non-extreme influential case, and at P3 in an adjusted sample excluding both of these cases. The effect was largest, by *B* value at channel P3, $F(1, 25) = 11.87$, $p = .002$, $B = 6.81$, $se = 1.97$, at channel C3, $F(1, 25) = 11.32$, $p = .002$, $B = 6.46$, $se = 1.91$, at channel Oz, $F(1, 26) = 6.14$, $p = .020$, $B = 5.34$, $se = 2.14$, and at channel PO8, $F(1, 26) = 6.06$, $p = .021$, $B = 5.31$, $se = 2.14$. With the full scalp measure, at relatively low N2/3 delta power, a nap was estimated to decrease recall percent by 5.37 points ($se = 2.55$), $t = -2.10$. There was less clear of an effect at relatively high N2/3 delta power, $t = 0.80$, where a nap was estimated to increase recall percent by 2.15 points ($se = 2.68$). This interaction was further illustrated by a positive correlation, $r(26) = .40$, $p = .033$, between N2/3 delta power and the nap minus control change in recall percent (Figure 2.4b), though this correlation again excluded four participants who only had one valid recall percent score.

The percent of sleep time spent in stage N3 sleep also, understandably, moderated the effect of condition on recall percent in an adjusted sample excluding the same influential case with extremely high recall percent on both sessions. This interaction was the same nature as that of N3 duration and N2/3 delta power, and technically did not meet the more restrictive significance criterion (α of .01) for variables of only secondary focus.

Neither REM group nor any other sleep measure, including spindle measures and

EEG power measures from stages N2 and N3 sleep and stage R sleep, were significant moderators of condition for recall percent.

None of the examined sleep measures or REM group were significant moderators of a condition effect on recall SD. However, it should be noted that due to the small sample of valid scores for recall SD (31 scores, 20 participants), many regression models could not be fit without influential cases before too much data had been excluded.

Discussion

This study was designed to test the SR2 hypothesis proposal that sleep supports both the accessibility of newly acquired memories for retrieval and the fidelity at which these memories may be retrieved. It was predicted that, compared to a period of wakefulness, a 90-min nap opportunity would result in better memory accessibility, measured as the proportion of items for which the approximate colour was recalled, and better memory fidelity, measured as the accuracy to the specific target colours among those approximated. A non-significant decrease in error of recall with a nap relative to control was directionally consistent with the hypothesis; however, for the proportion of item colours recalled, there was worse performance after the nap relative to control. Sleep in the form of a nap was not effective in improving accessibility and fidelity of memory for the item-colour task.

The effect of the nap opportunity was moderated by the extent of NREM sleep slow wave activity in the nap. Participants with relatively less slow wave activity, reflected in less delta power in stages N2 and N3 or less stage N3 sleep, generally recalled fewer colours after a nap than in the control condition. There was no clear difference between nap and control performance for participants with relatively more slow wave activity in their nap, but participants with the greatest nap-related gains in the number of colours

recalled (i.e., nap - control) tended to have the most NREM slow wave activity (Figure 2.4). This interaction is accordant with the prediction that memory accessibility benefits attributable to the nap would associate with properties of NREM sleep linked to memory consolidation. An association between memory benefits and greater NREM sleep slow wave activity is consistent with past research (Diekelmann et al., 2012; Holz et al., 2012; Huber et al., 2004; Landsness et al., 2009; Ruch et al., 2012; Schabus et al., 2005; Schönauer et al., 2017; Tamaki et al., 2013), but it is unclear why wake condition performance was greater than nap condition performance for those with less NREM sleep slow wave activity in their nap. It may be that playing competitive games in the control condition generally increased performance motivations, resulting in a benefit to performance equivalent to or greater than benefits that could be accrued from all but the most beneficial naps. This control condition was conceptualized as an ecologically valid alternative activity one might pursue on an afternoon in lieu of taking a nap. Whether alternative control conditions such as quiet wakefulness might produce other results is an open question.

Even for the participants with the greatest nap-related gains in the number of item colors recalled and the most NREM sleep slow wave activity, it cannot be said with certainty that NREM sleep slow wave activity affected performance. Slow wave activity of NREM sleep is considered a physiological marker of sleep need (Borbély & Achermann, 1999). Thus, this pattern may be the result of participants who regularly have a subtle sleep debt both performing poorly at recall on the control day due to sleepiness and having relatively more NREM sleep slow wave activity during their nap. Experimental manipulations of sleep may be used to better understand the relationship between properties of NREM sleep and memory accessibility. Study 2 utilized TMR

during NREM slow wave sleep to more directly examine whether benefits to memory accessibility are driven by memory processing mechanisms of NREM sleep.

Despite predictions, there was no indication that REM sleep was associated with nap-attributable benefits to memory fidelity. This question was examined both using the amount of scored stage R sleep as a continuous measure and using REM group as a categorical grouping of those with at least 5 min of stage R sleep (REM+) and those without 5 min of stage R sleep (REM-). The occurrence of roughly equal numbers of participants in the REM+ and REM- groups was favourable for addressing this part of the SR2 hypothesis. A factor that was not favourable for addressing this part of the SR2 hypothesis was the small sample of recall error scores after exclusions. Recall SD scores were excluded if they were obtained from performances for which too few colours were recalled to properly measure error of recall, leaving only 31 recall SD scores in total and only 11 participants with a recall SD score for both the nap and control conditions. It is also worth noting that the amount of stage R sleep obtained was quite low even among the REM+ group (m = 12.27, sd = 3.60), and it may be that greater amounts are required to produce benefits to memory performance. Study 3 was better suited to identify a possible association between REM sleep and memory fidelity as it included an experimental manipulation of REM sleep in the form of a REMD protocol.

The difficulty of the item-colour task was a major limitation of this study. Nearly one quarter of performances (23%), including all data from seven (18%) participants, were excluded from all analyses due to performance that was not significantly greater than chance. Even among the greater-than-chance performances, the percent of item colours approximated was under 12% for one quarter of performances and under 27% for three quarters of performances. Due to the generally low number of item colours recalled,

estimates of recall error were excluded from 48% of the performances deemed above chance and 60% of all performances recorded. This high level of data exclusion not only reduced the likelihood of observing significant effects for both memory accessibility and memory fidelity measures, but it may have also impacted the generalizability of the results. For instance, the interaction showing the extent of nap-related effects to vary by NREM sleep slow wave activity in the nap may only hold for a population of relatively high-motivation individuals willing to persevere in the task despite the demotivating effect of not recalling approximately 80% of the colours. The generally poor performance in approximating item colours at recall was likely the result of shallow encoding of many items and perhaps a generally poor ability to retain colours information over the roughly two-hour retention period. Future investigations would benefit from ensuring items are sufficiently encoded and the features tested may be well-remembered over the retention period examined. Study 2 and Study 3 used item location as the feature of interest and ensured adequate encoding by including a performance criterion to be met during the learning period.

Chapter 3

Study 2: The Effect of Targeted Memory Reactivation

The primary focus of Study 2 was to investigate the NREM sleep component of the SR2 hypothesis: the notion that NREM sleep processes reinforce newly acquired memories, increasing their accessibility for later retrieval. The technique of TMR during NREM slow wave sleep was selected for this purpose. This study was built on the rationale that increasing the extent to which some memories are processed during NREM slow wave sleep (through TMR), one could examine whether processes of NREM slow wave sleep have specific effects on memory accessibility.

TMR during NREM slow wave sleep has been frequently shown to improve memory performance or at least protect memories from degradation or interference for both declarative and procedural memory tasks despite considerable variation in the TMR procedure (Oudiette & Paller, 2013; Schouten et al., 2017). Both odour cues (Diekelmann et al., 2011; Diekelmann et al., 2012; Rasch et al., 2007; Rihm et al., 2014) and sound cues (Creery et al., 2015; Antony et al., 2012; Oudiette et al., 2013; Rudoy et al., 2009; Schreiner et al., 2015) have been successful in producing this effect. Benefits to visuospatial memory performance (i.e., memorizing the location of individual items on a screen) have been reported most often (Bendor & Wilson, 2012; Creery, et al., 2015; Diekelmann et al., 2011; Diekelmann et al., 2012; Oudiette et al., 2013; Rasch et al., 2007; Rihm et al., 2014; Rudoy et al., 2009). In some cases, TMR has been effective in producing memory benefits for specific items or learners. Although, first reports of sound based TMR for visuospatial material revealed general benefits for cued items relative to non-cued items (Rudoy et al., 2009), subsequent investigation found this cueing effect particularly for high learners or, at the individual item level, for items learned with an

intermediate-to-high degree of error (Creery et al., 2015). While benefits for cued items over non-cued items are typical, there has been at least one case of sound based TMR benefiting both cued and non-cued items in a visuospatial memory task. Specifically, Oudiette et al. (2013) found that low-value items that would otherwise tend to be forgotten over sleep were rescued from forgetting by TMR during slow wave sleep regardless of whether the specific items were cued or not. Thus, while memory benefits of TMR during sleep, particularly slow wave sleep, appear quite robust to variations in method, characteristics of the learning experience may moderate its effectiveness.

Some studies have found TMR cueing to produce changes in EEG power spectra during sleep, and, in some cases, these changes have been associated with cueing benefits to memory performance. Rihm et al. (2014) found that slow wave sleep presentation of an odour paired with learned material before sleep, but not a different odour, increased delta power over frontal electrodes and increased fast sigma power over parietal electrodes relative to moments before odour presentation. They also found that only presentation of the paired odour produced relatively faster negative-to-positive slopes of slow oscillations at frontal electrodes. Furthermore, TMR-produced increases in negative-to-positive slow oscillation slopes were associated improved memory performance over sleep, suggesting that changes in NREM sleep EEG power spectra are associated with a beneficial reactivation of memory representations (Rihm et al., 2014). Similarly, Creery et al. (2015) found the benefit of sound based TMR cueing during slow wave sleep in a nap to be associated with greater delta power over frontal electrodes during slow wave sleep and greater density of fast spindles over parietal electrodes during the cueing period. Lastly, Schreiner et al. (2015) found benefits of cueing vocabulary learning during NREM sleep were associated with greater sound-cue-induced theta and sigma power at frontal

electrodes. Together, these studies suggest that effectiveness of TMR cueing may depend on properties of sleep or be predicted by EEG responses to TMR cues.

These studies examined the effects of TMR during slow wave sleep on memory performance and many used a continuous measure of error from target during recall, but none, to the best of the author's knowledge, have dissected the memory benefits from cueing along characteristics akin to memory accessibility and memory fidelity. To do so using the analysis procedure introduced and used in Chapter 2 would require more items (approximately 100 per experimental condition) than has been used in previous research on the effect of TMR during sleep.

The current study used a visuospatial memory task and TMR procedure like those shown to produce beneficial effects of TMR during sleep (e.g., Rudoy et al., 2009), but adapted to measure both the accessibility and fidelity of memory. To improve performance in this item-location task over that of the item-colour task of Study 1, the item-location task included an active learning procedure with a high learning criterion. Participants were shown 20 blocks of 10 items each with each item being a circular-frame image in a specific location on a circular grid. Participants repeatedly put these images in their locations until they passed the learning criterion for each block. Throughout the task, each image was repeatedly paired with a unique and semantically related sound during both the learning period and a test before an 8-hr nocturnal sleep period. Half of the items were selected for sound based TMR during NREM slow wave sleep. In the morning, participants were again tested on their ability to place all images in their proper locations. While a general decline in memory performance from the pre-sleep test to the post-sleep test was expected, it was predicted, given the hypothesis that NREM sleep processes increase the accessibility of memories, that the ability to correctly place images in their

approximate location after sleep would be greater for cued items than control items. This study also provided an opportunity to examine relationships between properties of a full night of sleep and overnight changes in memory accessibility and memory fidelity. Based on the SR2 hypothesis, it was predicted that NREM slow wave activity obtained over the night, reflected in amounts of N3 slow wave sleep and delta-band EEG power during NREM sleep, would be associated with less overnight decline or overnight gains in memory accessibility. It was predicted that measures of REM sleep, particularly the amount of stage R sleep obtained would be associated with less overnight decline or overnight gains in memory fidelity.

Method

Participants

The study was completed by 38 young adults from the Brock University community who were recruited to participant in a study examining the role of sleep in memory for locations. All participants reported being good sleepers with typical daily sleep of approximately seven or more hours roughly within 22:00 and 9:00 and did not report working late or overnight shifts frequently in the past six months or traveling over multiple time zones in the past three months. All participants reported no history of psychiatric condition or head injury and reported not having any current medical condition or regular substance use other than caffeine (e.g., medications, alcohol, cannabis) affecting sleep or cognitive function. All participants reported only low-to-moderate daily caffeine consumption (approximately <300 mg per day). On the days on the experiment, participants were asked to consume no substance affecting cognitive function or sleep (e.g., alcohol or caffeine), to take no naps, to obtain no vigorous exercise, and to eat adequate meals so as to not be hungry before sleep. For the sleep

period preceding those in the lab, participants were asked to go to sleep no later than 23:30 and to wake up at 7:30. All participants reported normal hearing and normal or corrected-to-normal vision. Normal colour vision was confirmed through completion of the Ishihara colour test (Ishihara, 2014), and normal hearing between 500 and 2000 Hz was confirmed through a hearing test with an audiometer conducted by a researcher. All participants provided informed consent and received an honorarium of up to $60 or course credit for participation in the study.

One participant was excluded because they reported consuming caffeine before the experimental night and this caffeine appeared to have detrimental effects as they had a near-minimum amount of N3 sleep and they frequently woke up to sound cues during the night. With this participant excluded, the sample was 37 participants (31 female) with a mean age of 20.38 (sd = 2.60, minimum = 18, maximum = 28, Q_1 = 18, Q_2 = 19, Q_3 = 22). Of the 31 female participants, 17 were taking hormonal contraceptives, 6 were estimated to be in the follicular phase of the menstrual cycle, 7 were estimated to be in the luteal phase of the menstrual cycle, and 1 did not provide useful data on the menstrual cycle questionnaire. In contrast to performance on the item-colour task of Study 1, all participants were found to perform at levels significantly greater than chance at recall on the item-location task, and thus no participants were excluded from all memory task analyses due to poor performance.

Item-Location Memory Task

The memory task was programmed in Psychopy (Peirce, 2007). Visual elements of the task were presented on an LED computer monitor. Task sounds were played through desktop computer speakers placed to the sides of the monitor. Participants completed the task using a standard keyboard and mouse while seated alone in their

assigned bedroom at a viewing distance of approximately 60 cm. The item-location memory task can be divided into multiple parts: the learning period, the recall tests, the recognition test, and the sound discrimination test. Scripts for each part are available from MacDonald (2020) via the Open Science Framework.

Stimuli. The stimuli used for the memory task items were 210 (200 for the experiment, 10 for practice and demonstration) easily identifiable and distinct images of animals (e.g., bat or cat), environments (e.g., crashing waves or baseball game), human actions (laughing or marching), musical instruments (e.g., bagpipes or guitar), vehicles (e.g., train or helicopter), and other objects (e.g., shopping cart or microwave) obtained from various online repositories of free-to-use and attribution-free images (e.g., pixabay.com and pexels.com). Each image was paired with a 1-s, representative sound clip (e.g., cat–meow). Some sound files were obtained from the freely available 110-item set of sound clips provided by Hocking et al. (2013) for neuroimaging research. Additional sound files were in the public domain and obtained from online repositories (e.g., freesound.org). Sounds were selected and edited to be acoustically distinct from each other and to recognizably represent their paired image within their short duration. Normalization of loudness between the sound files was preformed using the normalization tool in Audacity® with subsequent amplitude adjustments made to files that were subjectively quieter or louder than the others. Sound files were played at approximately 50 dB sound pressure level during computer tasks and at approximately 37 dB sound pressure level during the sleep period, both measured from the likely head position of participants at those times (i.e., respectively, 60 cm from the computer monitor and centered on a pillow at the top of the bed). Sound intensity of 37 dB was selected to match the intensity used in previous studies using TMR during slow wave

sleep (Creery et al., 2015; Oudiette et al., 2013; Rudoy et al., 2009).

During the task, the images appeared on a circular grid composed of a grey circle with a 9.8º VA radius, 35 thin black radial lines each cutting a 10º arc of the circle, and 4 thin black concentric rings with radii 20%, 40%, 60%, and 80% the radius of the grey circle. For each participant, each image was assigned a target location on the circular grid (Figure 3.1a). Target location assignment was random with the stipulation that for no item could more than half of the other items in the same block be within 90º angular distance from that item and for no item could more than half of all other items in the task be within 90º angular distance from that item. Locations varied in both circular angle and distance from the centre of the circle with a minimum distance from centre of 5.3º VA and a maximum distance from centre of 8.6º VA. Images appeared on the circular grid, filling a thin, black, and circular frame that was centered on their target location and occluded exactly 15º of the circular grid at the assigned distance from centre. The limits placed on the possible distances from centre allowed some variation in item locations while restraining images to, at maximum distance, be at the edge of the circular grid without leaving it and to not be too small at the minimum distance.

Learning. The learning period was divided into 20 blocks with 10 items each. Each block began with the onscreen prompt “Memorize card locations”, before the 10 images were sequentially shown alone at their target location on the circular grid for 3.0 s while the paired sound was played (Figure 3.1a). There was a randomly determined 0.5–1.0 s interval between items. Participants then began the first of up to 10 rounds of placing the items on the grid. At the start of the round, the screen displayed the round number and the prompt, “Move cards to proper locations.” In a random order, the first of the 10 images appeared in the center of the circular grid. At the same time, its paired

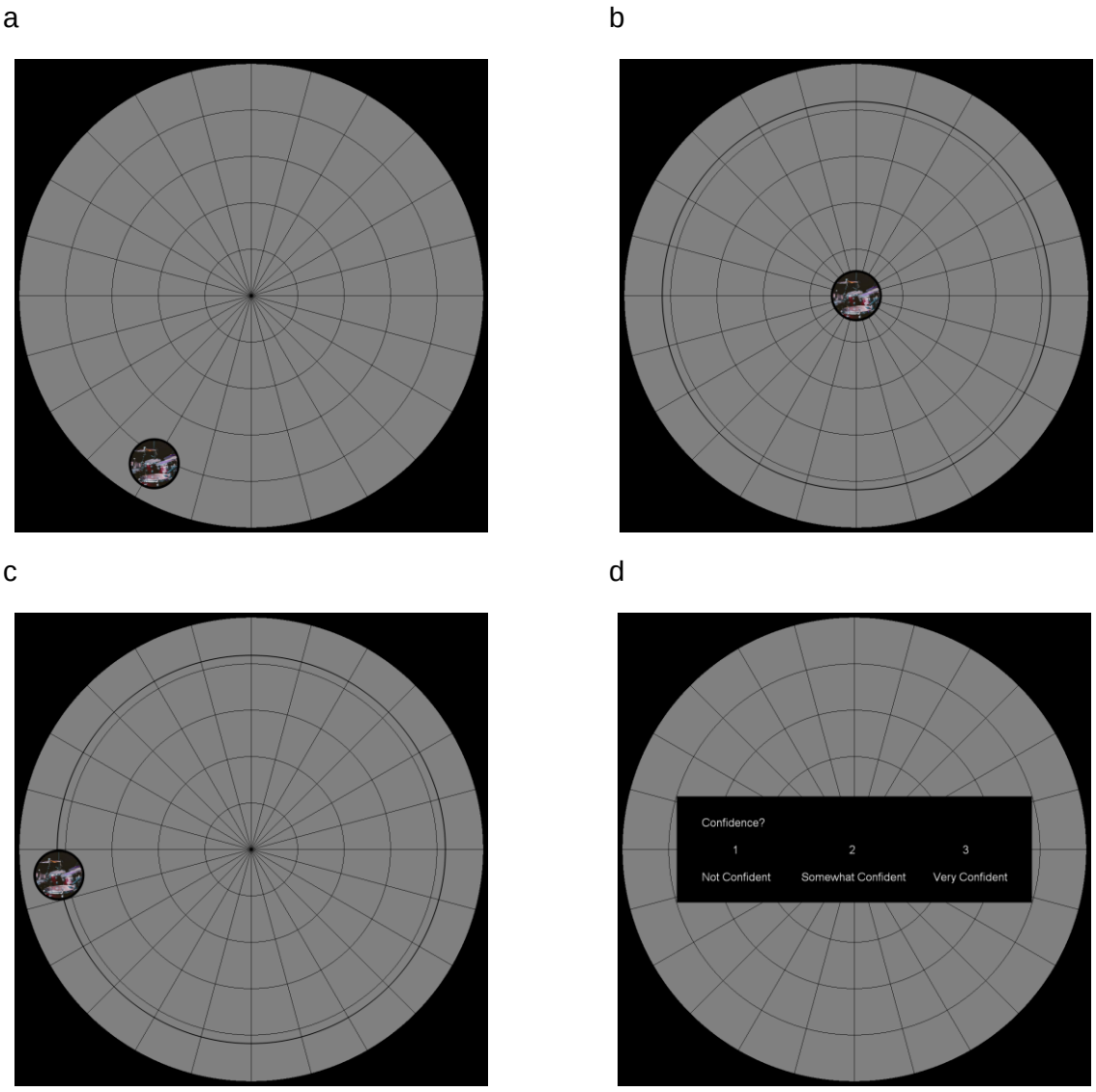

Figure 3.1. Item-location memory task displays. **a.** Image in target location for 3.0 s for initial or feedback presentations during learning. Paired sound played at initiation. **b.** Recall probe during learning or recall test. Image was centered for 0.2 s before it could be moved. Paired sound played at initiation. Thicker black ring indicated possible response locations. **c.** Response display during learning or recall test. Image could be moved along the black ring until the image location was submitted as the response location. **d.** Prompt probing confidence in response location following its submission during recall test.

sound was played and a black ring twice the thickness as those forming the grid appeared centered on the grid and intersecting the target location (Figure 3.1b). After 0.2 s, the mouse pointer appeared over the image, signalling that participants could then use the mouse to move the image on the grid. As the participant moved the mouse, the image moved along the black ring so that all possible responses were the correct distance from the centre of the circle but varied on the angular distance from the target location (Figure 3.1c). Participants selected their response location at their own pace with a mouse click. After the response, there was a 0.1 s delay followed by the simultaneous disappearance of the image at the response location, reappearance of the image at the target location, and presentation of the paired sound. This feedback display (Figure 3.1a) remained for 3.0 s before being cleared from the screen at the start of a randomly determined 0.5–1.0 s pause before this process was repeated for the other items in the block. The angular distance between the response location and target location was recorded for each response. Participants were then shown a prompt informing them of the number of items they placed correctly. A response was considered correct when it was within 10º angular distance of the target location. This entire process was then repeated for round two with the items being given in a new random order. For round three and all subsequent rounds, participants were only given an item if they had not previously placed it correctly in two consecutive rounds. The number of items remaining for the upcoming round out of the 10 total items in each block was included in the end-of-round prompts of round two through to round nine. Once all 10 items of a block were learned (i.e., placed correctly in two consecutive rounds) or all 10 rounds were completed, participants moved on to the next block until all 20 blocks were completed. Items that were not learned in 10 rounds were recorded as such.

Recall tests. Participants were given one recall test between the learning period and sleep period and given another recall test in the morning after the sleep period. Apart from a separate randomization of item order, the two recall tests were identical. The recall test was structured similarly to the learning period with a few exceptions: All 200 items were now presented in a random order and were no longer organized by block, participants were asked to place each item only once, no feedback display was given after responses, and participants were asked to report their confidence in their response immediately after they placed each item. For confidence reports, the image disappeared from the screen and the prompt, "Confidence?" appeared centered over the circular grid with an indication of the following response options: 1 for "Not Confident", 2 for "Somewhat Confident", and 3 for "Very Confident" (Figure 3.1d). Participants responded to the confidence probe using the number row on the keyboard. Input of a confidence response removed the probe from the screen and initiated a randomly determined 0.5–1.0 s pause before the next item appeared until the task was complete. For each recall test item, the angular distance between response location and target location was recorded along with the corresponding response time and confidence rating.

Recognition test. The recognition test was given only once, immediately following the post-sleep recall test. In the recognition test, each of the 200 images appeared on screen in their target location, one at a time, in a random order with their paired sounds playing once as they appeared. On each trial, a duplicate image appeared in a decoy location at the same time as the image in the target location. After 0.2 s, the mouse pointer appeared at the center of the screen and participants could use the mouse to respond. Participants were asked to click on the image in the correct location. Participant response removed both images from the screen and initiated randomly determined 0.5–

1.0 s pause before the next item was presented until the task was complete. Separately for cued and control items, the distance between the decoy location and the target location was randomly selected, without replacement, from a 100-item list with the following composition: eight of each multiple-of-five distance from 30–50º (i.e., eight decoy distances of 30º distance, eight decoy distances of 35º, etc.), four of each multiple-of-five distance from 55–70º, and two of each multiple-of-five distance from 75–180º. Response time and accuracy were recorded for each trial.

Sound discrimination test. The sound discrimination test was given after the recognition test and after participants were informed that half of the tasks' sounds were played to them during sleep. In the sound discrimination test, participants were asked to indicate which sounds were played during the night. Participants heard each sound sequentially in a random order, without its paired image. For each sound, participants responded by key press whether they believed the sound was played during the night (right arrow key) or not (left arrow key). Participant response initiated a randomly determined 1.0–1.5 pause before the next sound was played until the task was complete. Before responding, participants could also make a key press to hear the sound again (down arrow key). Accuracy was recorded and a d' score computed to measure the ability to discriminate between cued and control items.

Electrophysiological Recording

Electrophysiology was recorded using Neuroscan SynAmps2 amplifiers with Scan 4.5.1 software (Compumedics Inc., Abbotsford, Australia) and gold-plated silver electrodes sampling at a rate of 1000 Hz filtered DC to 200 Hz with an additional notch filter at 60 Hz. Two electrode montages were used: one montage for screening and adaptation sessions used to screen for sleep disorders and adapt participants to sleeping in

the laboratory and another montage for the experimental sessions. For screening and adaptation nights, EEG was recorded at four scalp sites (C3, C4, O1, and O2) placed according to the International 10-20 system (Pivik et al.,1993) with an online recording reference at Fpz, a ground reference placed at Afz, and additional offline reference electrodes at sites M1 and M2. EOG was recorded from two electrodes: both 1 cm away from each outer canthus on the bicanthal plane with the right EOG electrode placed 1 cm above the plane and the left EOG electrode placed 1 cm below the plane. A bipolar submental EMG was recorded with two electrodes placed under the chin. A bipolar EMG channels was created for each leg with two electrodes placed over both the left and right lower anterior tibialis. Dual-band respiratory inductance plethysmography was also recorded. The montage differences for the experimental sessions included EEG being recorded at 12 scalp sites (F3, Fz, F4, C3, Cz, C4, P3, Pz, P4, O1, Oz, and O2), and EOG being recorded from four electrodes: two 1 cm away from each outer canthus on the bicanthal plane, one 2 cm directly above the left orbit, and one 2 cm directly below the left orbit. Furthermore, leg EMG and respiration were not monitored for the experimental sessions. Electrical impedances were below 5 KΩ at all scalp sites and below 10 KΩ at all peripheral sites prior to recording.

Procedure

All procedures were cleared by the Brock University Bioscience Research Ethics Board.

Recruitment, screening, and orientation. Participants were recruited primarily from the Brock University community through print and digital advertisements and classroom visits. Advertisements indicated that the study was examining the role of sleep in memory for locations. Before enrollment in the study, interested participants were

asked, typically over the phone, questions regarding sleep and health to ensure they met the eligibility criteria. Those who were willing and eligible attended a 30-min laboratory screening session to give informed consent, perform the Ishihara test of colour deficiencies (Ishihara, 2014), perform a hearing test using an audiometer, and respond to a more detailed screening questionnaire regarding sleep behaviour and health. Consenting participants who met the eligibility criteria after these screening measures were then scheduled for two consecutive overnight sessions in the laboratory, the first of which was the screening and adaptation night and the second of which was the experimental session. On the first night, participants came to the lab at 22:30 and immediately completed a short task of visual short-term memory before the electrophysiological recording equipment was applied. Afterwards, participants were taken to their assigned bedroom where they were left to sleep, lights off, from 23:30 to 7:30. For the entire sleep period, low-level white noise was played from desktop speakers on the side ends of the headboard. White noise was played at 37 dB sound pressure level measured from a normal head position centered and near the top of the bed. If necessary due to electrodes becoming unfixed, a researcher would enter the room, awaken the participant, and quickly re-apply the electrodes.

After the sleep period of the screening and adaptation night, recording equipment was removed and participants were able to leave the laboratory at 8:00. Sleep records from the screening and adaptation nights were examined during the day for evidence of sleep disorder or difficulty. Participants were excluded from further participation when evidence of sleep disorder (e.g. periodic limb movement disorder) or sleep difficulty (less than 6 hr of sleep or less than 80% sleep efficiency) was discovered. Otherwise, participants returned to the laboratory later that evening for the experimental session.

Experimental session. The experimental session began at 22:30 the evening following the screening and adaptation night and continued until approximately 10:00 the next morning. For organization, the experimental session may be divided into the before-sleep period, the sleep period, and the after-sleep period.

Before sleep. The experimental protocol began at 20:30 with the application of the electrophysiological recording equipment. At approximately 21:00, participants started a practice block of the item-location task under observation of a researcher who verbally reinforced the onscreen instructions. The practice block was a full 10-item block of item locations delivered through the same procedure used for the learning period of the main task. For some participants who were behind schedule (e.g., due to arriving late to the laboratory), the practice block was terminated without completion, but not before it was clear that they fully understood the task. Before the start of the learning period, participants were informed that they would never need to remember the location of the practice block items, but that they would be asked to recall the locations of the real task items both on a test before sleep and in the morning. They were also instructed to take a quick guess if they did not know the location of an item, pay attention to the feedback of correct locations, and be as precise as they could with their responses when they did know the correct location of an item. Participants then started the main learning period no later than 21:30. Given that participants learned to a set criterion, the duration of the learning period was variable, but it was designed to be typically completed within 75 and 05 min. After the learning period, participants took a break of at least 5 min before starting the test. Break duration was longer for participants who completed the learning period ahead of schedule because the start time of the pre-sleep recall test was set to be no earlier than 23:00.

After completion of the pre-sleep recall test designed to be typically completed within 10 and 25 min, participants got in bed and completed a short questionnaire asking them to report any unusual events during the day, whether they had caffeine, whether they napped, and their current mood state and level of sleepiness. The questionnaire was given within 5 min of the start of the seep period, which began at 23:30 for most participants (i.e., those who finished the test early or on schedule), or as soon as possible for the few participants who were behind this schedule.

Sleep period. The 8-hr sleep period started at approximately 23:30 and ended at approximately 7:30 for most participants with these times shifted later for those who competed the pre-sleep recall task behind schedule. To start the sleep period, a researcher turned on the white noise set to play throughout the sleep period, turned off the light and computer monitor, and closed the door, leaving them to sleep alone in the bedroom. White noise was played from two desktop speakers on the side ends of the headboard at 37 dB sound pressure level measured from a normal head position centered and near the top of the bed.

As the white noise was initiated, the computer determined which 100 of the 200 items were assigned to be cued during the night. Items were randomly assigned to cued and control items repeatedly until the two lists of items were matched within each block and in total on the average of the item-level average error of responses when learned (i.e., the mean of the absolute values of the distances between target and response on the final two correct responses required for an item to be considered learned), the average of number of presentations of each item during learning, the number of items learned, the average distance of target locations from centre, the average error of responses during the pre-sleep recall test, and the average confidence rating received during the post-sleep

recall test. The matching protocol progressed block-by-block, assigning half of the learned items and half of the non-learned items to be cued then comparing the to-be-cued and control items on the previously listed measures. The separate averages for the to-be-cued and control lists were considered matched if they were within 0.5 times their average standard deviation of each other. If they were not matched, random assignment was repeated to a maximum of 2500 attempts. After every 250 attempts, the factor multiplied by the average standard deviation of the two lists was increased by a factor of 2 (e.g., 0.5 to 1.0). Once each block was resolved, the complete lists of to-be-cued and control items were compared. The to-be-cued and control lists were considered matched if their separate averages were within 0.5 times their average standard deviation of each other and the counts learned items in each list were within 1 of each other. If they were not matched, the entire process of randomly assigning conditions within each block was repeated to a maximum of 200 attempts. After every 25 attempts, the factor multiplied by their average standard deviation of the two complete lists was increased by a factor of 2 and the allowable difference in the number of learned items in each complete list was increased by 0.2. For most participants, the complete lists of to-be-cued and control items were matched by attempt 30.

During the night, the TMR procedure was enacted. The TMR procedure was modeled after the procedure used by Rudoy et al. (2009), altered to fit the full night of sleep and a greater number of task items. Sleep was monitored online by K. M. for the presence of stage N3 sleep. When a participant maintained unambiguous N3 sleep for 2 min, the researcher initiated sound cueing. First, three 600 Hz, 1.0-s tones were played at 35 dB sound pressure level with a 6.0-s inter-stimulus interval to test whether sounds could be played without arousing the participant. If the participant did not show signs of

arousal to these tones, the sound cues from the task followed. The sound cues were played in random order with a 6.0-s inter-stimulus interval. From -0.5 to -0.1 s before the sound cue was played, the volume of the white noise was gradually reduced, reaching a minimum of 31 dB sound pressure level and remaining at that level for 0.1 s before the cue was played, its full 1.0 s duration, and 0.1 s after the cue before returning to its original level over the next 0.4 s. Effectively, these changes meant the white noise played at 37 dB sound pressure level for the majority of the 5.0-s interval between sounds and was lowered to 31 dB sound pressure level while the sound cues were played at approximately 37 dB sound pressure level. After all 100 cues were played, there was a 2-min pause of only white noise before beginning again with the three 600 Hz tones and again playing each sound cue in a new random order. The cueing procedure was temporarily halted when the continuously monitored PSG of the participants showed evidence of a transition to wake or R sleep. The cueing procedure was not halted when participants transitioned out of N3 sleep into N2 sleep without signs of arousal, though the occurrence of this event was relatively rare. Cues were played a maximum of five times each. The TMR procedure was considered complete after five presentations of the 100-item list. Participants did not receive the maximum of 500 cues if they did not achieve enough uninterrupted N3 sleep or if they repeatedly transitioned to wake in response to the cues. Although there was some subjectivity in determining when to continue attempts to play the cues and when to terminate the TMR procedure, the goal was to play as many cues as possible without resulting in significant disruption of sleep. Generally, the cueing procedure was terminated before the maximum was reached if there were three-to-five attempts at cueing that resulted in transitions to wake or after the fifth hour of the sleep period because cues beyond that point were likely to wake participants.

The script managing both the matching of to-be-cued and control items and the TMR procedure is available from MacDonald (2020) via the Open Science Framework.

After sleep. Upon awakening at approximately 7:30, participants were given a short questionnaire to report on their sleep experience, current mood, and level of sleepiness. They were then taken to a separate room where they ate instant oatmeal and water for breakfast while watching their choice of popular movie from a selection of generally low-arousing options. They watched the movie until approximately 8:55 when they were taken back to the bedroom to begin the remaining parts of the item-location task at 9:00, including the post-sleep recall test identical to the pre-sleep recall test and the recognition test designed to be typically completed within 8 and 12 min. After both tests, participants were probed with the question, "did you notice anything weird or unusual last night?" to give them an opportunity to freely report that they had heard the sound cues being played or noticed anything else about the TMR procedure. They were then informed that half of the sounds from the task were played to them during the night, embedded inside the white noise. They were then asked to complete to sound discrimination test to determine if they could reliably indicate which sounds were played during the night. The sound discrimination test was designed to be typically completed within 8 and 12 min. After this last task, female participants were given a short menstrual cycle questionnaire and electrophysiological recording equipment was removed while participants were further debriefed about the nature of the study. Participants left the laboratory at approximately 10:00.

Data Analysis

Data analysis scripts for this study are available from MacDonald (2020) via the Open Science Framework.

Memory performance measures. Performance on the recall tests was calculated using largely the same method described in Chapter 2 for the item-colour task. Memory accessibility and memory fidelity were thought to be reflected, respectively, in the frequency at which reported locations at least approximated the target locations and the extent to which such reports were accurate to the exact target locations. For each participant, recall test (pre-sleep and post-sleep), and group of stimuli (cued, control, and both combined), a distribution of response errors (i.e., distance between report and target for each trial) was collected. Items that were not learned to criterion during the learning period were excluded from these distributions. Recall percent—the percent of test trials for which at least the approximate target location was recalled—and recall SD—the standard deviation of the distribution of response errors excluding those for which the approximate target location was not recalled—were estimated for each distribution of response errors. As for the item-colour task, estimation of recall percent and recall SD was completed in a two-step process including the determination of a preliminary recall proportion score from a criterion of defined recall failure and the subsequent fitting of a mixed normal and uniform probability density function to a distributions of response errors. Readers may refer to Chapter 2 for a description of this process. The only difference between the process used for the item-colour task and the process used for the item-location task was a change in the criterion of defined recall failure from 90º in the item-colour task to 80º in the item-location task. This change was made to address a tendency for some participants to show a non-normal clustering of response errors around 90º away from the target location. Although these orthogonally misplaced items reflect a level of low-fidelity recall, the analysis procedure assumes a normal distribution of errors for recalled items, and thus an 80º criterion was selected to allow for low-fidelity

responses up to 80º error while avoiding this non-normal clustering which may negatively impact estimates of error of recall.

The recognition test given after the post-sleep recall test was administered as an alternative measure of memory performance in the case that performance on the post-sleep recall test was problematically low as was the case for the item-colour task reported in Chapter 2. Recall performance was sufficiently good for nearly all participants, and thus recall performance was analyzed. Recognition performance was not thoroughly analyzed and is not reported.

Sleep measures. PSG records from the sleep period of the experimental night were subjected to several procedures to obtain measures of sleep architecture, EEG power spectra, and sleep spindles.

Sleep scoring. Lateral EEG channels were referenced to contralateral mastoids. A bipolar horizontal EOG channel was created from the two electrodes on the bicanthal plane, and a bipolar vertical EOG channel was created from the EOG electrodes above and below the left orbit. The electrode placements used for the experimental session did not include the American Academy of Sleep Medicine recommended EOG electrodes placed approximately 1 cm above and to the right of the right eye and approximately 1 cm below and to the left of the left eye. To correct for this oversight and still allow automated sleep scoring, these EOG channels were estimated using a combination of the vertical and horizonal EOG channels available. Sleep records were scored through a combination of automated and manual scoring. Neuroscan software was used to apply a 0.3 to 100 Hz bandpass filter to all channels in each file before decimating the files to 250 Hz and subjecting them to the *Michele* sleep scoring system, an automated sleep scoring system developed by Younes Sleep Technologies, which has been shown to be valid for scoring

sleep and associated events, particularly when combined with manual editing (Malhotra et al., 2013). *Michele* scored each 30-s epoch as wake, stage R, stage N1, stage N2, or stage N3 and marked arousals with start and end markers. After automated scoring, K. M. examined each sleep file in its entirety and the scoring of sleep stages and arousals was manually edited according to American Academy of Sleep Medicine criteria (Berry et al., 2015). Manual editing was conducted within the *Michele* sleep viewer with only C3, C4, O1, and O2 EEG channels, submental EMG, the two estimated EOG channels, and the two mastoid channels visible. Within the viewer, EEG channels were referenced to the contralateral mastoid, EOG channels were referenced to the right mastoid, a 0.3–100 Hz bandpass filter was applied to EEG and EOG channels, and a 1–100 Hz bandpass filter was applied to the submental EMG channel. The primary focus of manual editing was to confirm sleep onset and review the staging of epochs for which the algorithm was uncertain. Beyond this focus, common edits included adding or removing arousal markers, moving arousal markers to optimally encompass EEG artifact introduced by the arousals, and small adjustments to sleep staging around transitions between wake, stage R, stage N1, and stage N2 epochs. Sleep architecture variables of total sleep duration, the time spent in each sleep stage, the percent of total sleep duration spent in each stage, and the total number of arousals (arousals plus transitions to wake from sleep) were obtained from the scored files.

EEG power spectra. Measures of average, log-transformed absolute power in delta, theta, alpha, sigma, and beta frequency bands per epoch for stage N2, stage N3, and stage R sleep for each EEG channel and a full scalp measure were acquired through methods nearly identical to those described in Chapter 2 with the only difference being that separate measures were acquired for N2 and N3 sleep rather than combining them

into N2/3 measures. Readers may refer to Chapter 2 for a description of these methods. As for Chapter 2, tables and figures reporting descriptive statistics of average EEG power measures or analyses including average EEG power measures used full scalp measures unless otherwise specified.

Spindle detection. Spindles on channels C3 and C4 were identified by the *Michele* sleep scoring system. The spindle count used for analyses was the average combining the number of spindles identified in N2 and N3 sleep at channel C3 and the number of spindles identified in N2 and N3 sleep at channel C4. A spindle density measure was calculated by dividing the spindle count by the total minutes of stage N2 sleep and stage N3 sleep combined.

Induced power changes to cues. The neurological response to the sounds played during the night was examined by looking at changes in power spectral density occurring over time in response to the cue. Cueing events followed by arousal or other artifact were identified through visual inspection in Neuroscan using sleep records referenced and filtered for sleep scoring and subsequently excluded from induced power analysis. With neuroscan software, EEG channels were re-referenced to the average of the two mastoids and then subjected to a 0.5–100 Hz bandpass filter. Using the MNE Python toolbox (Gramfort et al., 2013), EEG data was then cut into 6.0-s epochs, -2.0–4.0 s around each cue onset. These epochs were resampled to 256 Hz. Time-frequency representations with a time resolution of 0.1 s were then computed from these epochs for each channel using Morlet waves for each whole number frequency using f / 2 cycles where f is equal to the frequency. Due to analysis artifact that occurs at the edges of the time-frequency representations, the time range examined was reduced to -1.5–3.5 s around cue onset. Each time-frequency representation then underwent baseline correction using -1.5–0.0 s

as the baseline period and the *zlogratio* method which, for each frequency, divided induced power values by the mean of the baseline values, took the log, and divided the values by the standard deviation of the log-transformed baseline values.

For each channel, time-frequency representations from all participants were averaged to form a grand average of the induced power response. This grand average was visually examined in combination with individual averages collected from each participant to identify time-frequency bands of interest indicating the major components of the typical induced power response to the sound cues during sleep. The induced power response in these time-frequency bands of interest were quantified within each participant at the channels for which the given induced power response was greatest in the grand average. For each participant and time-frequency band response of interest, an additional "maximum channel" measure was created, capturing the maximum induced power value over the quantified channels regardless of the specific site the value came from. Tables and figures reporting descriptive statistics of induced EEG power measures or analyses including induced EEG power measures used maximum channel measures unless otherwise specified.

Statistical analyses. Statistical analyses were conducted in R (R Core Team, 2017) using base functions and a variety of specialized packages, including *reshape2* (Wickham, 2007) for handling data structures and *ggplot2* (Wickham, 2016), *grid* (R Core Team, 2017), and *gridExtra* (Auguie, 2016) for creating figures.

Data exclusion and adjustment. The methods used to exclude or adjust data due to poor recall test performance and efforts to fit regression models without influential cases were nearly identical those described in Chapter 2 with the only exception being an increase in the estimated instances of recall success required to allow the inclusion of the

recall SD measure of a given performance. For the current study, recall SD scores were excluded from analyses if they were based on fewer than 50 instances of recall success. This criterion is greater than the criterion of 30 recall successes used in analysis of Study 1 because the higher level of performance on item-location task afforded the opportunity of increasing the criterion to a more desirable level. Readers may refer to Chapter 2 for a description of data exclusion and adjustment methods. As for Study 1, regression models fit without influential cases were the focus of follow-up analyses and evaluations of statistical significance when predicting performance measures, and, when effects or associations are reported as significant in models fit to adjusted data, the specific exclusions and adjustments are reported.

All recall test performances (i.e., each participant for each recall test for cued items, control items, and both item types combined) were found to be greater than chance, and thus no performances were excluded for this reason. No recall SD scores for performance measures combining both cued and control items were excluded based on too few instances of recall success. A total of eight participants had at least one of their recall SD scores excluded for performance measures involving either cued or control items alone. One of them had both their cued item and control item pre-sleep recall SD scores and their post-sleep control item recall SD score excluded. Another one of the eight had their pre-sleep cued item recall SD score and both their cued item and control item post-sleep recall SD scores excluded. Three of them had only their post-sleep control item recall SD score excluded, and three of them had only their post-sleep cued item recall SD score excluded. Thus, no data were excluded due to poor performance for analyses combining both cued and control items but recall SD scores were excluded for 12 of 148 (8.1%) performances concerning only cued or only control items; these

exclusions came from 8 of 37 (21.6%) participants.

Statistical tests. The primary focus of statistical testing was to identify how various factors such as cueing during sleep or properties of sleep may have affected or related to changes in memory performance between the two tests. Overnight change scores for recall percent, recall SD, and recall confidence were calculated by subtracting pre-sleep scores from post-sleep scores. Overnight change scores were not computed for cases in which either the pre-sleep or post-sleep measure was excluded; such cases were considered to have missing data for analyses of overnight change. Analyses not concerned with possible effects of cueing were conducted on performance measures computed from the full distribution of response errors collapsing across cued and control items due to the greater confidence placed in performance estimates computed from larger distributions of response errors.

The strategies guiding the formation of regression models and statistical tests were like those for Study 1. A combination of simple (non-mixed-effect) linear regression models and mixed-effect linear regression models were used to examine how pre-sleep performance, sleep properties, and cueing associated with memory performance. Simple regression models were used to examine associations between predictors and performance when each participant offered only one data point for each measure (e.g., overnight change in recall percent for all items), and mixed-effect regression models were used when participants had more than one data point each for a measure (e.g., recall percent at pre-sleep and post-sleep tests or overnight change in recall SD for cued items and for control items). Mixed-effect models were constructed using *lme4* (Bates et al., 2015). In all such models, a random intercept was allowed for each participant. Tests of associations, or “effects”, in simple regression models were conducted using the *Anova*

function in the *car* package (Fox & Weisberg, 2011) to perform analyses of variance *F*-tests using type II sums of squares. For mixed-effect regression models, the same function was used to perform type II Wald *F*-tests with denominator degrees of freedom approximated through the Kenward-Roger method (Kenward & Roger, 1997). All tests of interaction terms were conducted in models containing the corresponding lower-order terms. Simple effects and slopes were estimated at the 25th and 75th percentiles of the interacting predictor variables. Continuous predictors in regression models were standardized ($m = 0$, $sd = 1$). *B* values from standardized predictors on unstandardized outcomes are reported to show the magnitude and direction of effects from linear models. Categorical variables of test time and condition were effect coded such that coefficients indicated the effect of post-sleep test relative to pre-sleep test and the effect of cueing compared to control.

The following additional tests were also conducted. A one-sample *t*-test against 0 was used to test whether participants could discriminate between cued and non-cued sounds in the morning. Pearson's correlation coefficient was calculated to quantify the association between memory performance measures.

Statistical significance was defined through *p* value less than alpha of .05 for tests concerning hypotheses and measures of considerable practical or theoretical interest. More specifically, this was the only significance criterion when testing the effect of test time or condition and when testing the association between memory performance and either N3 duration or R duration. Other theoretically interesting measures were considered as sets of related measures. One set included N3, the N3 duration × N3 delta power interaction for each channel and the full scalp measure; another set included spindle count, spindle density, the N2 duration × N2 sigma power interaction for each

channel and the full scalp measure, and the N3 duration × N3 sigma power interaction for each channel and the full scalp measure. For measures of the induced power response to cues during the night, each quantified time-frequency band response of interest had its own defined set consisting of the different EEG channels (including the maximum channel measures) for which the time-frequency band response was quantified. Within these defined sets of measures, false discovery rate was controlled for using the Benjamini–Hochberg method (Benjamini & Hochberg, 1995) to limit false discovery rate to 0.10. Thus, for these measures, statistical significance was defined as having a *p* value less than .05 and the critical *Q* determined to limit false discovery to 0.10. Tests on recall percent were always considered in a different set than tests on recall SD. Alpha was lowered to .01 for tests of other sleep measures, including total sleep duration, N1 duration, N2 duration, the percent of total sleep duration spent in each sleep stage, arousal count, arousal density, average EEG power in other frequency bands of stages N2 and N3 sleep, and average EEG power in stage R sleep. For these average EEG power measures, the Benjamini–Hochberg method was applied to limit false discovery rate to .05 within each similarly constructed set with a separate set for each remaining frequency band within stage N2 and stage N3 sleep and a separate set for each frequency band within stage R sleep. Brief explorations of results of marginal significance, defined as those meeting the criterion of limit false discovery but exceeding α by no more than 0.04, are reported.

Results

Sleep

Descriptive statistics for sleep architecture, average EEG power, and induced EEG power variables are reported in Table 3.1. Correlations among select sleep measures and

Table 3.1
Descriptive statistics for sleep architecture and select EEG power measures

Measure	n	*m*	*sd*	min	max	Q_1	Q_2	Q_3
Sleep dur.	37	450.72	15.56	392.00	473.00	443.00	454.50	461.50
N1 dur.	37	34.74	14.63	13.50	78.50	25.50	32.50	40.00
N2 dur.	37	233.99	28.95	138.00	292.50	218.00	238.00	251.00
N3 dur.	37	90.18	25.00	27.00	144.00	74.00	88.50	107.50
R dur.	37	91.81	25.14	33.50	155.00	78.50	93.50	104.00
N1 %	37	7.71	3.21	3.09	16.61	5.54	7.31	9.68
N2 %	37	52.02	7.03	30.60	63.52	47.30	52.76	56.24
N3 %	37	19.93	5.33	6.89	31.65	15.64	19.10	23.92
R %	37	20.34	5.43	7.56	34.37	17.91	20.81	22.49
Arousals	37	98.32	26.68	60.00	175.00	74.00	96.00	119.00
Arousal den.	37	0.22	0.06	0.13	0.39	0.16	0.21	0.26
Spindles	37	974.74	548.60	0.00	2383.50	505.00	914.00	1380.00
Spindle den.	37	3.00	1.64	0.00	7.27	1.60	2.93	4.20
N2 delta	37	-10.22	0.14	-10.45	-9.77	-10.30	-10.26	-10.14
N2 theta	37	-11.11	0.13	-11.37	-10.90	-11.18	-11.12	-11.02
N2 alpha	37	-11.51	0.23	-11.96	-10.99	-11.66	-11.54	-11.40
N2 sigma	37	-11.63	0.23	-12.12	-11.14	-11.83	-11.59	-11.45
N2 beta	37	-12.66	0.20	-13.07	-12.23	-12.77	-12.64	-12.54
N3 delta	37	-9.61	0.13	-9.93	-9.38	-9.72	-9.59	-9.51
N3 theta	37	-10.88	0.13	-11.16	-10.69	-10.97	-10.84	-10.78
N3 alpha	37	-11.48	0.22	-11.97	-11.03	-11.62	-11.51	-11.34
N3 sigma	37	-11.77	0.22	-12.22	-11.39	-11.98	-11.75	-11.61
N3 beta	37	-12.75	0.21	-13.20	-12.23	-12.88	-12.77	-12.62
R delta	37	-10.79	0.14	-11.07	-10.47	-10.88	-10.78	-10.72
R theta	37	-11.33	0.16	-11.73	-10.99	-11.41	-11.37	-11.21
R alpha	37	-11.66	0.25	-12.05	-11.20	-11.86	-11.71	-11.48
R sigma	37	-12.20	0.20	-12.61	-11.79	-12.35	-12.21	-12.08
R beta	37	-12.61	0.21	-13.04	-12.12	-12.77	-12.62	-12.44
Ind. delta	37	4.45	2.35	0.90	11.96	2.95	4.55	5.16
Ind. theta	37	4.94	2.36	0.72	9.54	3.33	4.70	6.00
Ind. sigma	37	5.41	2.80	1.05	11.35	3.17	5.11	6.80

Note. Duration (dur.) reported in minutes. Density (den.) reported as count per minute of sleep (arousal density) or minute of stage N2 sleep and stage N3 sleep (spindle density). Induced (ind.) power measured at central channels as *z*-scores relative to pre-stimulus interval.

the number of cues played during the night are depicted in Figure 3.2.

Response to Cueing

On average 441.24 (*sd* = 87.73, Q_1 = 402, Q_2 = 500, Q_3 = 500) sound cues were

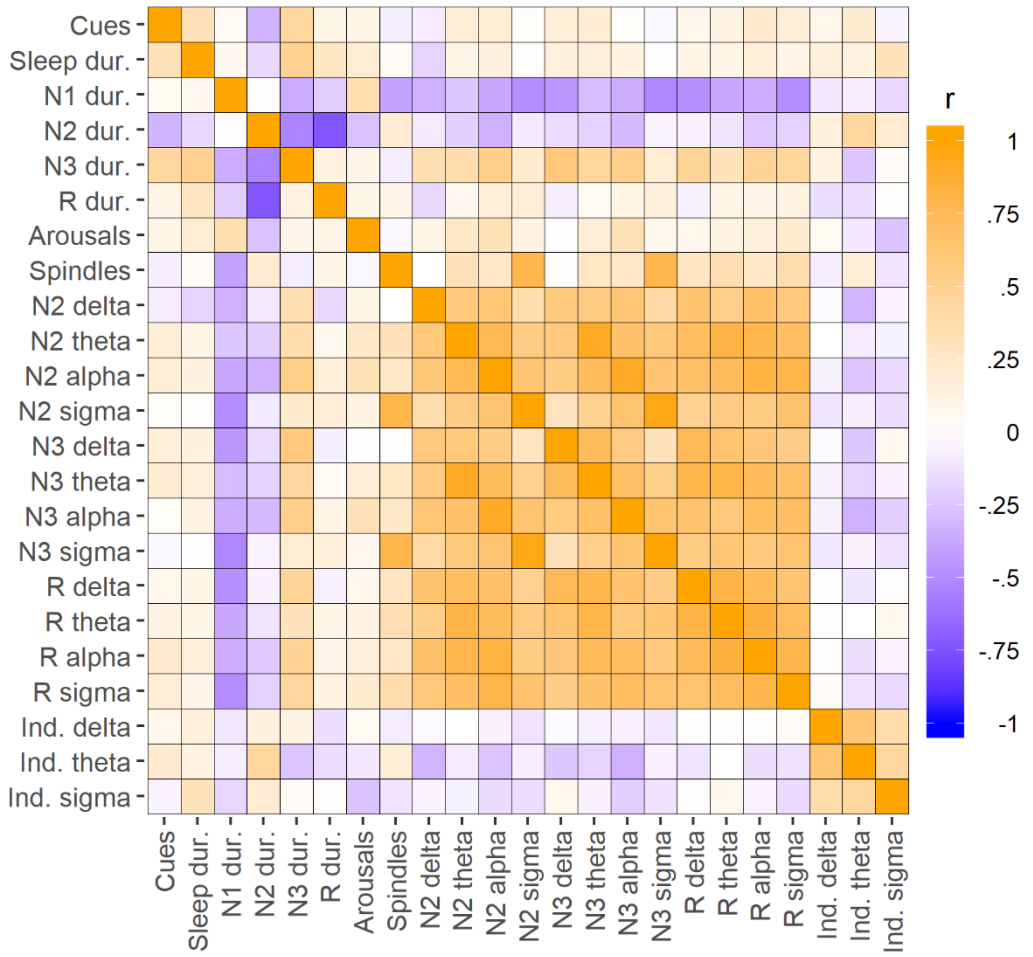

Figure 3.2. Pearson correlation matrix of cue count and sleep measures with *r* values depicted on a coloured scale. Dur. = duration. Ind. = induced. Cues indicates the number of sound cues played during the night.

played to participants during the night. Of the 37 participants included in the sample, 21 participants received all 500 sound cues, and 3 participants received less than 300 sound cues (191, 204, and 257), meaning not all the 100 cues were played at least three times. Of all sound cues played, 92.1% were played in epochs later marked as stage N3, 7.6% were played in epochs later scored as stage N2, 0.1% were played in epochs later marked as stage N1, and 0.2% were played in epochs later scored as wake with the latter two

categories the result of arousals immediately following the cue. Before being briefed about the playing of sounds during the night, only one participant reported hearing sounds when probed with the question of, “did you noticed anything strange or unusual last night?” This participant did not specifically report hearing sounds from the memory task, but only “noises”, and they were not able to reliably indicate which sounds were played

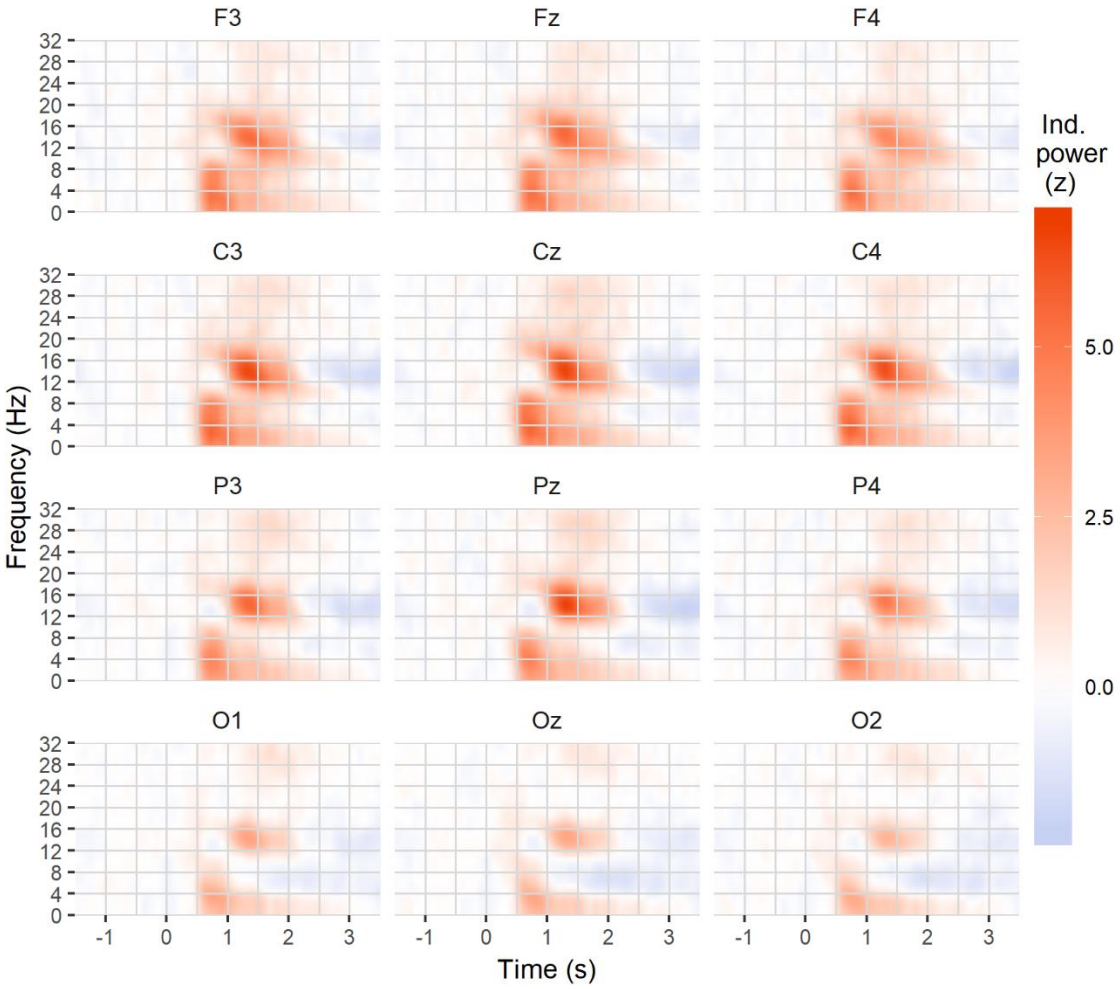

Figure 3.3. Average induced electroencephalographic power response to cues across all 37 included participants at 12 scalp channels. Induced power values with a resolution of 0.1 s and 1 Hz were calculated as z-scores relative to the -1.5–0.0 s pre-stimulus interval and are represented on a colour scale.

during the night in the sound discrimination test (d' = -0.05). Participants, on average, were also unable to reliably indicate which sounds were played during the night with the mean d' score in the sound discrimination test being -0.01 (*sd* = 0.18, minimum = -0.35, maximum = 0.33), $t(35) = -0.07$, $p = .941$.

Grand averages for the induced EEG power response to sound cues during sleep are depicted in Figure 3.3. Visual inspection of the grand and individual averages identified an increase in delta power 0.5–1.5 s post cue, an increase in theta power 0.5–1.0 s post cue, and an increase in sigma power 1.0–1.7 s after the cue that was most prominent across central channels (C3, Cz, and C4). Thus, the induced power response within these bands at these sites was quantified for each participant along with a "Cmax" measure capturing each participant's maximum induced power value over these channels. Descriptive statistics for these induced power variables are reported in Table 3.1.

Learning Performance

The average duration of the learning period was 91.46 min (*sd* = 16.25, minimum = 71.32, maximum = 145.65, $Q_1 = 80.65$, $Q_2 = 89.03$, $Q_3 = 97.00$). On average, participants learned 198.97 of the 200 items (*sd* = 3.66). One participant learned only 178 (89%) of the items; all other participants learned at least 196 (98%) of the items, and 27 (73%) participants learned all 200 items to criterion.

Pre-Sleep and Post-Sleep Recall

Descriptive statistics for pre-sleep and post-sleep recall measures are reported in Table 3.2. The distributions of response errors for all learned items, separated by recall test and item condition are shown in Figure 3.4. The average duration of the pre-sleep recall test was 18.07 min (*sd* = 4.26, minimum = 11.55, maximum = 31.55, $Q_1 = 15.02$, $Q_2 = 17.62$, $Q_3 = 19.85$). The average duration of the post-sleep recall test was 18.12 min

Table 3.2
Descriptive statistics for memory performance in recall tests

Items	Measure	n	*m*	*sd*	min	max	Q_1	Q_2	Q_3
All	Pre-sleep recall %	37	75.42	11.08	46.54	92.67	70.06	77.18	83.62
	Pre-sleep recall SD	37	11.84	3.33	6.21	20.50	9.45	11.60	13.45
	Pre-sleep confidence	37	2.12	0.22	1.59	2.51	2.00	2.10	2.30
	Post-sleep recall %	37	69.54	12.83	34.97	88.15	63.72	72.77	78.20
	Post-sleep recall SD	37	12.90	4.86	6.37	33.33	10.24	12.14	13.83
	Post-sleep confidence	37	2.05	0.26	1.37	2.60	1.91	2.04	2.23
	Overnight Δ recall %	37	-5.88	5.10	-16.99	5.43	-9.95	-5.43	-1.81
	Overnight Δ recall SD	37	1.06	2.79	-3.98	12.83	-0.40	0.69	2.25
	Overnight Δ confidence	37	-0.07	0.15	-0.37	0.25	-0.14	-0.06	-0.01
Control	Pre-sleep recall %	37	75.89	11.44	47.31	91.76	70.86	77.07	84.53
	Pre-sleep recall SD	36	12.20	4.48	6.35	27.28	9.39	11.70	12.57
	Pre-sleep confidence	37	2.12	0.22	1.60	2.50	2.01	2.12	2.32
	Post-sleep recall %	37	70.16	12.97	40.08	88.15	65.63	71.86	79.10
	Post-sleep recall SD	32	12.81	4.55	7.09	26.53	9.89	11.74	14.35
	Post-sleep confidence	37	2.04	0.27	1.24	2.58	1.90	2.07	2.25
	Overnight Δ recall %	37	-5.72	7.11	-25.27	12.66	-8.52	-5.48	-3.43
	Overnight Δ recall SD	32	1.21	3.59	-7.77	14.40	-0.93	0.99	2.52
	Overnight Δ confidence	37	-0.08	0.16	-0.38	0.25	-0.16	-0.08	-0.03
Cued	Pre-sleep recall %	37	75.13	11.17	42.92	93.57	70.22	77.29	82.72
	Pre-sleep recall SD	35	11.30	3.04	5.72	18.98	9.34	11.25	12.94
	Pre-sleep confidence	37	2.11	0.22	1.58	2.52	1.98	2.11	2.28
	Post-sleep recall %	37	69.05	14.12	22.32	89.95	60.63	73.67	77.29
	Post-sleep recall SD	33	12.35	4.33	5.37	24.63	9.81	11.69	13.73
	Post-sleep confidence	37	2.05	0.26	1.50	2.61	1.90	2.07	2.23
	Overnight Δ recall %	37	-6.07	6.37	-20.60	3.62	-9.05	-5.43	-1.81
	Overnight Δ recall SD	32	0.83	2.43	-5.31	7.15	-0.60	0.74	1.96
	Overnight Δ confidence	37	-0.06	0.15	-0.39	0.26	-0.14	-0.05	0.03

Note. All-item recall percent, recall SD, and confidence differed significantly ($\alpha = .05$) from pre- to post-sleep. No significant effects of condition (control vs. cued) were detected. Recall SD measured in degrees of error. Confidence rated on a 1–3 scale.

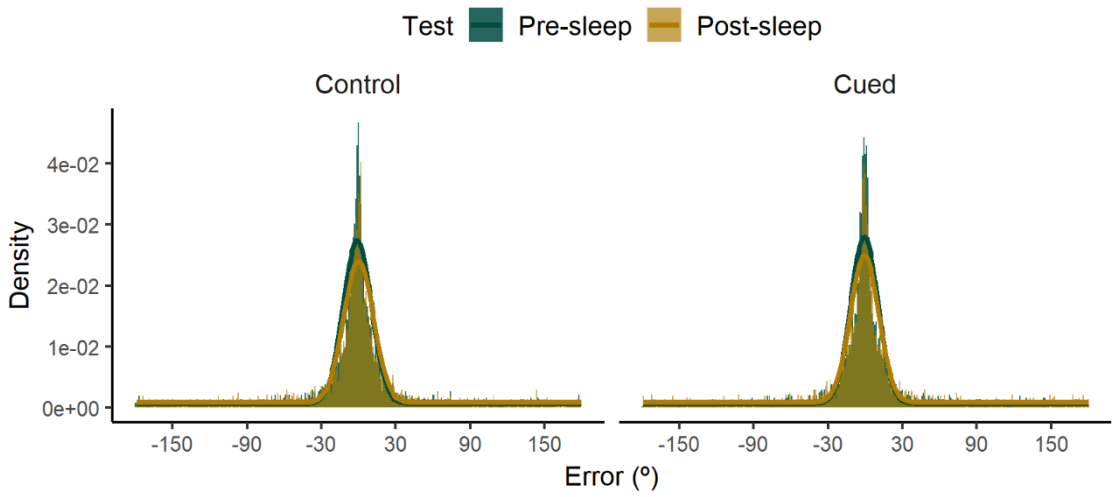

Figure 3.4. Distributions of responses measured as error in angular distance from target location. Solid lines indicate density of a mixed distribution composed of a uniform distribution of guesses and a normal distribution of successful recall with parameters fit to the corresponding distribution of errors. Data separated by condition and coloured by time of recall test.

(*sd* = 5.19, minimum = 9.07, maximum = 31.88, Q_1 = 14.10, Q_2 = 17.28, Q_3 = 20.67).

At the pre-sleep test, participants, on average, recalled 75.42% percent of learned item locations with a recall SD of 11.84º and a confidence rating of 2.12. Pre-sleep recall percent was not significantly correlated with pre-sleep recall SD, $r(34) = -.02$, $p = .929$, or pre-sleep confidence, $r(34) = .16$, $p = .366$, both in an adjusted sample with one influential case with extremely low pre-sleep recall percent (-1.73 IQR) and extremely high pre-sleep recall SD (+1.77 IQR) excluded. After exclusion of this case and two additional influential cases without extreme scores, pre-sleep recall SD and pre-sleep confidence were significantly negatively correlated, $r(32) = -.44$, $p = .009$; this correlation was similar the unadjusted sample, $r = -.30$. At pre-sleep test, control and cued items were

not significantly different on recall percent, $F(1, 36) = 1.35$, $p = .254$, recall SD, $F(1, 34) = 1.43$, $p = .240$, or confidence, $F(1, 34) = 0.04$, $p = .833$, with two influential cases excluded for the test for confidence.

At the post-sleep test, on average, participants recalled 69.54% of learned item locations with a recall SD of 12.90º and a confidence rating of 2.05. After exclusion of an influential and extreme case with low post-sleep recall percent (-1.99 IQR), high post-sleep recall SD (+5.43 IQR), and low post-sleep confidence (-1.71 IQR), post-sleep recall percent and post-sleep recall SD were not significantly correlated, $r(34) = -.28$, $p = .098$. In the same adjusted sample, the correlation between post-sleep recall percent and post-sleep confidence was not significant, $r(34) = .23$, $p = .182$. Post-sleep recall SD was not significantly correlated with post-sleep confidence, $r(32) = -.15$, $p = .387$, in an adjusted sample excluding the previously described influential and extreme case, another influential case with extremely high post-sleep recall SD (+2.01 IQR), and another influential case without extreme scores. Cued and control items did not differ significantly on post-sleep recall percent, $F(1, 36) = 0.63$, $p = .433$, post-sleep recall SD, $F(1, 31) = 1.29$, $p = .265$, or post-sleep confidence, $F(1, 36) = 0.52$, $p = .478$.

Overnight Change in Recall Performance

The focus of analyses and predictions was on changes in memory performance measures over the night from the pre-sleep recall test to the post-sleep recall test as these measures were thought to be more sensitive to effects of sleep processes. Descriptive statistics for overnight change scores are reported in Table 3.2.

The percent of learned item locations recalled was significantly lower post-sleep compared to pre-sleep, $F(1, 36) = 49.27$, $p < .001$, with a 5.88-point ($se = 0.84$) decrease in recall percent over the night. Participants also had significantly worse fidelity of recall

post-sleep, $F(1, 35) = 4.88$, $p = .064$, with a 0.73º ($se = 0.33$) increase in recall SD over the night estimated even in an adjusted sample excluding the same, previously noted influential case with an extreme overnight change in recall SD (+3.99 IQR). Participants were also less confident at the post-sleep test, $F(1, 36) = 8.14$, $p = .007$, with a rating decrease of 0.07 ($se = 0.02$). With and only with the same influential case excluded, overnight change in recall SD significantly correlated with overnight change in recall percent, $r(34) = .36$, $p = .030$, indicating an association between overnight declines in recall percent and overnight gains in fidelity of recall. Overnight change in recall confidence was not significantly corelated with overnight changes in either recall percent, $r(35) = .02$, $p = .921$, or recall SD, $r(34) = -.01$, $p = .954$, with the latter correlation also calculated without the same influential case.

The following sections will first examine the effect of cueing on overnight change in recall performance then examine various measures as potential predictors of overnight changes on their own or as moderators of a cueing effect.

Cued vs. control. It was predicted that the cueing during N3 sleep would improve memory accessibility, resulting in better retention of approximate item locations for cued items relative to control items. However, overnight change in recall percent did not significantly differ between cued and control items $F(1, 36) = 0.06$, $p = .807$, $B = -0.35$, $se = 1.44$. There was also no significant difference between cued and control items for overnight change in confidence, $F(1, 36) = 2.14$, $p = .153$, $B = 0.02$, $se = 0.02$, or recall SD, $F(1, 30) = 0.94$, $p = .340$, $B = -0.49$, $se = 0.50$, with the latter test conducted in an adjusted sample excluding two influential cases. Mixed-effect regression models including both recall test (pre-sleep and post-sleep) and condition (control and cued) confirmed this lack of effect with non-significant test × condition interaction effects for

recall percent, $F(1, 108) = 0.05$, $p = .825$, recall SD, $F(1, 97) = 0.03$, $p = .871$, and confidence, $F(1, 108) = 0.54$, $p = .463$

Pre-sleep performance. The following tests examined whether pre-sleep recall percent, pre-sleep recall SD, or pre-sleep confidence predicted overnight change in memory performance or confidence.

Recall percent. Overnight change in the percent of learned item locations recalled was not predicted by either pre-sleep recall percent, $F(1, 35) = 0.71$, $p = .406$, $B = 0.72$, $se = 0.85$, or pre-sleep confidence, $F(1, 35) = 0.08$, $p = .779$, $B = 0.24$, $se = 0.86$, but pre-sleep recall SD was a significant predictor, $F(1, 35) = 10.32$, $p = .003$, with as estimated 2.43-point ($se = 0.76$) decrement to recall percent for every 3.33º (1 SD) increase in pre-sleep recall SD (Figure 3.5a). Thus, overnight decrements to recall percent were greatest

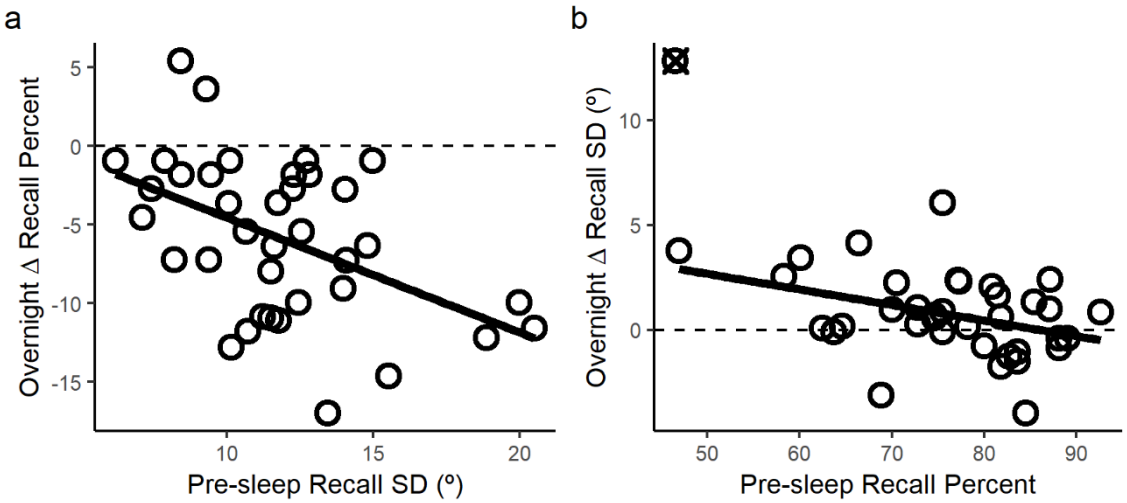

Figure 3.5. Overnight change in performance predicted by pre-sleep performance. **a.** Overnight change in recall percent predicted by pre-sleep recall SD. **b.** Overnight change in recall SD predicted by pre-sleep recall percent. The influential case excluded from linear models in figure and report is marked by the × symbol.

for those with lower fidelity recall before sleep.

Pre-sleep recall SD was a significant moderator of the cueing effect on recall percent, as indicated by a significant pre-sleep recall SD × condition interaction, $F(1, 33) = 5.70$, $p = .023$, $B = 3.13$, $se = 1.31$, in an adjusted sample excluding one influential case with extremely high pre-sleep recall SD (+1.77 IQR) and an extreme overnight change score for cued-item recall percent (-1.60 IQR) and another influential case without extreme scores on these measures. However, cueing did not have a clear effect at either low pre-sleep recall SD (9.38º), where it was estimated to decrease recall percent by 1.78 points ($se = 1.61$), $t = -1.11$, or at high pre-sleep recall SD (12.82º) where it was estimated to increase recall percent by 2.08 points ($se = 1.47$), $t = 1.42$. This interaction suggests cueing resulted in greater retention of approximate item locations for participants who had lower fidelity of recall before sleep, but it was not a strong interaction, and the two excluded cases show a clear contradicting pattern (Figure 3.6). Another piece of evidence that the effect of cueing on recall percent was moderated by pre-sleep performance was the pre-sleep recall percent × condition interaction being marginally significant after exclusion of the influential and extreme case with the lowest pre-sleep recall percent (-1.71 IQR), $F(1, 34) = 3.86$, $p = .058$, $B = 2.75$, $se = 1.40$. There was no indication of a cueing effect on recall percent at high pre-sleep recall percent (83.62%), $B = 1.36$, $se = 1.73$, $t = 0.79$. At low pre-sleep recall percent (70.30%), although the cueing effect was not particularly clear, $t = -1.41$, cueing was estimated to decrease recall percent by 2.28 points ($se = 1.61$). Pre-sleep confidence did not significantly moderate a cueing effect on recall percent, $F(1, 35) = 1.84$, $p = .184$, $B = 1.94$, $se = 1.43$.

Recall SD. Overnight change in recall SD was significantly predicted by pre-sleep recall percent, $F(1, 34) = 5.61$, $p = .024$, in an adjusted sample excluding an influential

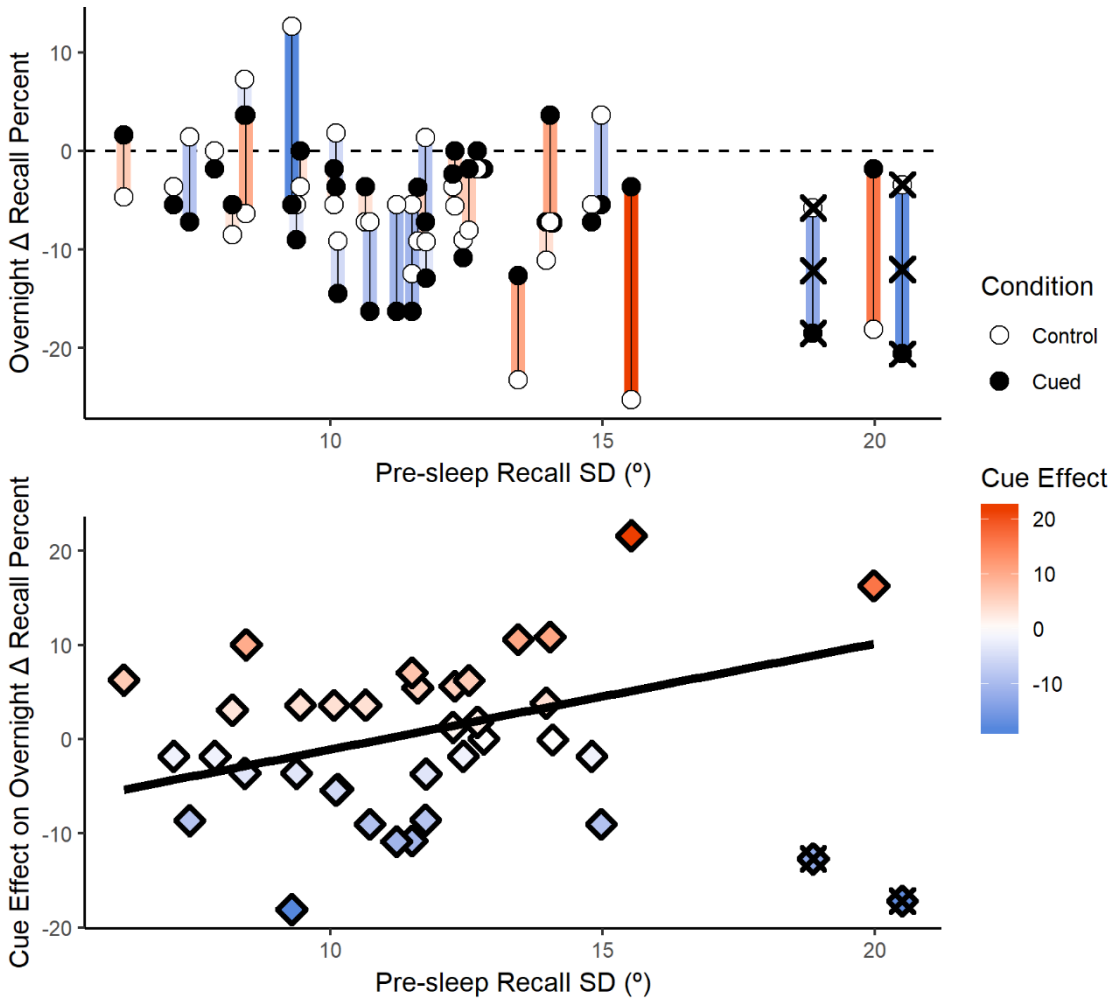

Figure 3.6. Pre-sleep recall SD moderating the effect of cueing (cued - control) on overnight change (post-sleep - pre-sleep) in recall percent. Vertical lines in top panel connect overnight retention scores for cued and control items within each participant. Cueing effect depicted in colour and on y-axis in bottom panel. Influential cases excluded from the linear models in figure and report are marked by the × symbol

and extreme case with low pre-sleep recall percent (-1.73 IQR), high pre-sleep recall SD (+1.76 IQR), and a high overnight change score for recall SD (+3.99 IQR). A 10.09-point (1 SD) increase in pre-sleep recall percent was associated with a 0.75° (*se* = 0.32) decrease in the recall SD change score, indicating less overnight decline or overnight

gains in recall fidelity (Figure 3.5b). Pre-sleep recall SD was marginally significant as a predictor of overnight change in recall SD, only after excluding the same influential and extreme case and two other influential cases without extreme scores on these measures, $F(1, 32) = 3.37$, $p = .076$. With these cases excluded, a 2.78º (1 SD) increase in pre-sleep recall SD was associated with a 0.56º ($se = 0.30$) decrease in the recall SD change score, suggesting those with less recall fidelity before sleep were more likely to maintain or increase their fidelity over the night compared to those with greater pre-sleep recall fidelity. Pre-sleep confidence was not a significant predictor of overnight change in recall SD, $F(1, 34) = 0.23$, $p = .631$, $B = -0.16$, $se = 0.34$, in the adjusted sample excluding the same influential and extreme case.

There was little evidence that pre-sleep performance moderated a cueing effect on recall SD. Neither the pre-sleep recall percent × condition interaction, $F(1, 38) = 0.39$, $p = .538$, $B = 0.61$, $se = 0.98$, nor the pre-sleep recall SD × condition interaction, $F(1, 28) = 0.02$, $p = .892$, $B = -0.08$, $se = 0.57$, were significant in predicting overnight change in recall SD, but the pre-sleep confidence × condition interaction was marginally significant, $F(1, 27) = 3.78$, $p = .062$, $B = 0.95$, $se = 0.49$. The latter two tests were conducted in adjusted samples excluding an influential case with extremely high pre-sleep recall SD (+1.63 IQR) and extreme overnight change scores for control item recall SD (-1.98 IQR), cue item recall SD (+2.02 IQR), and control item confidence (-1.69 IQR) and another influential case with an extreme overnight change score for control item recall SD (+3.44 IQR). Relative to control, cueing was associated with an estimated 1.12º ($se = 0.58$) decrease in recall SD at relatively low pre-sleep confidence (2.00 rating), $t = -1.93$, and had no clear effect on recall SD at relatively high pre-sleep confidence (2.30 rating), $B = 0.23$, $se = 0.61$, $t = .379$.

Sleep measures. The following tests examined whether measures of sleep were predictors of overnight change in memory performance.

Recall percent. A selection of the *F*-tests used to examine whether sleep measures were predictors of overnight change in recall percent own their own or as moderators of

Table 3.3
Tests of sleep measures as predictors of overnight change in recall percent on their own or as moderators of an effect of cueing

	As Predictor						As Moderator of Cueing Effect					
Measure	n	*F*	df_r	*p*	*B*	*se*	n	*F*	df_r	*p*	*B*	*se*
Sleep dur.	37[a1]	0.55	35	.462	0.64	0.85	37[a1]	0.91	35	.348	-1.38	1.45
N1 dur.	36[r1]	7.98	34	.008	-2.22	0.79	36r1	6.51	34	.015	3.48	1.36
N2 dur.	37	0.41	35	.526	-0.55	0.86	36[r1]	0.70	34	.409	-1.13	1.35
N3 dur.	37	0.60	35	.445	0.66	0.85	37	4.57	35	.040	-2.95	1.38
R dur.	37	1.59	35	.215	1.06	0.84	36[r1]	1.70	34	.201	1.74	1.34
Arousals	37	1.32	35	.258	-0.97	0.85	37	0.84	35	.365	-1.33	1.45
Spindles	37	1.42	35	.242	1.00	0.84	37	0.03	35	.857	-0.27	1.47
Ind. delta	37[a1]	2.29	35	.139	-1.26	0.83	37[a1]	0.44	35	.512	-0.97	1.46
Ind. theta	37	0.10	35	.751	-0.28	0.86	37	0.08	35	.780	0.41	1.47
Ind. sigma	37	3.02	35	.091	-1.44	0.83	37	0.07	35	.787	-0.40	1.47
N2 delta	37	0.72	33	.403	-0.82	0.97	36[r1]	0.69	32	.412	1.23	1.48
N2 theta	37	0.79	33	.380	-0.97	1.09	37	0.59	33	.450	-1.43	1.86
N2 alpha	34[r3]	6.85	30	.014	-2.98	1.14	34[r3]	0.02	30	.887	0.30	2.13
N2 sigma	36[r1]	0.05	32	.820	0.25	1.10	36[r1]	0.07	32	.796	0.46	1.75
N2 beta	35[r2]	4.11	31	.051	2.17	1.07	35[r2]	0.02	31	.900	-0.25	1.98
N3 delta	37	4.47	33	.042	1.63	0.77	37	0.01	33	.939	0.10	1.35
N3 theta	37	8.98	33	.005	2.87	0.96	37	1.00	33	.326	1.72	1.73
N3 alpha							35[r2]	2.15	31	.153	2.12	1.45
N3 sigma	37	0.21	33	.647	-0.52	1.13	37	0.81	33	.373	1.62	1.79
N3 beta	34[r3]	1.96	30	.171	-1.97	1.41	34[r3]	0.26	30	.612	1.26	2.46
R delta	35[r2]	1.32	31	.259	1.70	1.48						
R theta	37	0.00	33	.986	-0.02	1.16	35[r2]	1.13	31	.296	-1.95	1.83
R alpha	36[r1]	1.14	32	.294	-1.14	1.07	36[r1]	0.02	32	.887	-0.27	1.90
R sigma	37	0.03	33	.875	-0.15	0.97	36[r1]	0.02	32	.888	-0.22	1.59
R beta	37	0.55	33	.464	-0.87	1.17	36[r1]	1.03	32	.319	-2.00	1.98

Note. Dur. = duration. Ind. = induced. Average electroencephalographic power measures tested as interactions with corresponding sleep stage durations. Numbers after superscript "a" indicate the number of influential cases that had predictor values winsorized for a given test. Numbers after superscript "r" indicate the number of influential cases removed from for a given test. Empty cells are shown for tests for which a model without influential cases could not be fit.

an effect of cueing are reported in Table 3.3.

N3 duration was not a significant predictor of overnight change in recall percent on its own; however, it did significantly moderate a cueing effect, as indicated by the significant N3 duration × condition interaction, $F(1, 35) = 4.57$, $p = .040$, $B = -2.95$, $se =$

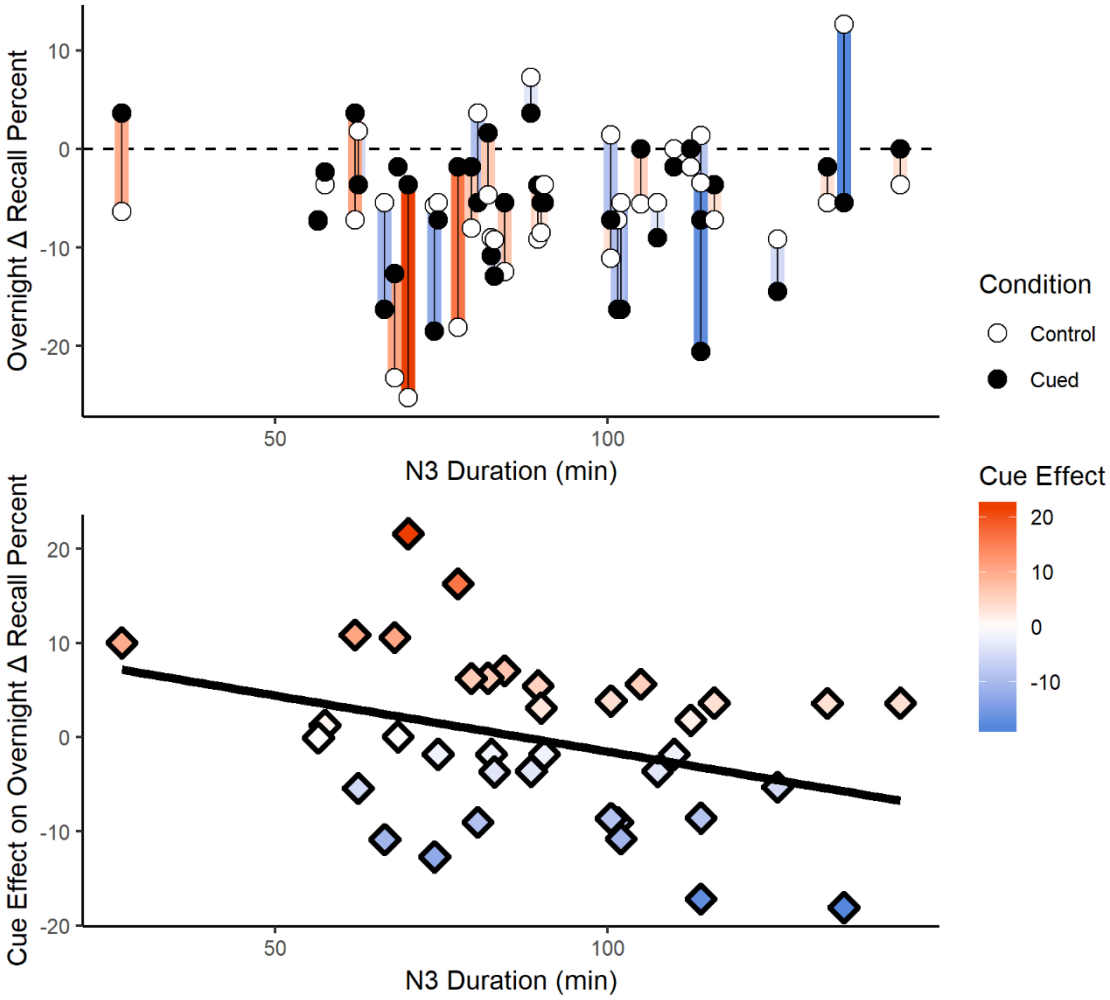

Figure 3.7. N3 duration moderating the effect of cueing (cued - control) on overnight change (post-sleep - pre-sleep) in recall percent. Vertical lines in top panel connect overnight retention scores for cued and control items within each participant. Cueing effect depicted in colour and on y-axis in bottom panel.

1.38 (Figure 3.7). There was no clear effect of cueing at either low N3 duration of 74.00 min, t = 0.96, or high N3 duration of 107.50 min, t = -1.44, but relative to control items, retention of recall percent for cued items was estimated to be 1.57 points (*se* = 1.64) greater for those with less N3 sleep and 2.41 points (se = 1.68) lesser for those with more N3 sleep. R duration was not a significant predictor of overnight change in recall percent, either on its own or in interaction with condition.

It was predicted that greater retention of the ability to recall item locations would be associated with having both high N3 duration and high average delta power in stage N3 epochs. This prediction was indeed reflected in the results as the N3 duration × N3 delta power interaction was significant at the full scalp measure and channels F4, Fz, P3, P4, O1, and O2. This interaction was marginally significant F3, C4, Cz, Pz, and Oz. An influential, non-extreme case was excluded for tests at P3, Pz, and P4, and a separate influential, non-extreme case was excluded for the test at F4. The effect was largest, by *B* value, at channel O2, *B* = 1.98, *se* = 0.77, channel P3, *B* = 2.46, *se* = 0.99, channel O1, *B* = 1.87, *se* = 0.81, and channel P4, *B* = 1.85, *se* = 0.85. A composite measure averaging N3 delta power over these four channels was created, and the interaction was also significant with this composite measure, $F(1, 33) = 5.41$, $p = .026$, *B* = 1.92, *se* = 0.82 (Figure 3.8). For those with low N3 delta power, N3 sleep had little association with overnight change in recall percent, *B* = -0.14, *se* = 1.17, t = -0.12; however, for those with high N3 delta power, the a 25.00-min (1 SD) increase in N3 sleep duration was associated with an estimated 2.68-point (*se* = 1.02) increase in the overnight retention of recall percent, t = 2.62. There was no evidence that the N3 duration × N3 delta power interaction moderated an effect of cueing on recall percent.

There were no significant predictors of overnight change in recall percent among

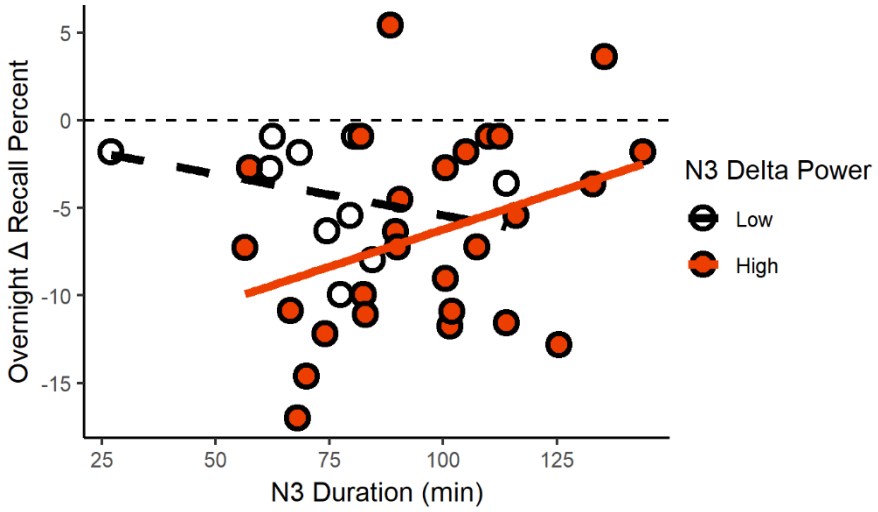

Figure 3.8. Overnight change in recall percent (post-sleep - pre-sleep) predicted by N3 duration on a median split of average delta (1–3.5 Hz) electroencephalographic power in stage N3 epochs over channels P3, P4, O1, and O2.

the set of spindle measures consisting of spindle count, spindle density, the N2 duration × N2 sigma power interactions, and the N3 duration × N3 sigma power interactions.

The remaining sleep architecture variables, including total sleep duration, N1 duration, N2 duration, the percent of total sleep duration spent in these sleep stages, arousal count, and arousal density, were also examined as predictors of overnight retention of recall percent on their own and as moderators of a cueing effect. Of these measures, only N1 duration was a significant predictor of overnight change in recall percent. The results observed for the percent of sleep time spent in each stage essentially duplicated those for the measures of sleep stage durations.

N1 duration was a significant predictor of overnight change in recall percent on its own, $F(1, 34) = 7.98$, $p = .008$, after exclusion of an influential case with extremely high N1 sleep duration (+2.66 IQR, 78.5 min). An additional 13.50 min (1 SD) of N1 sleep

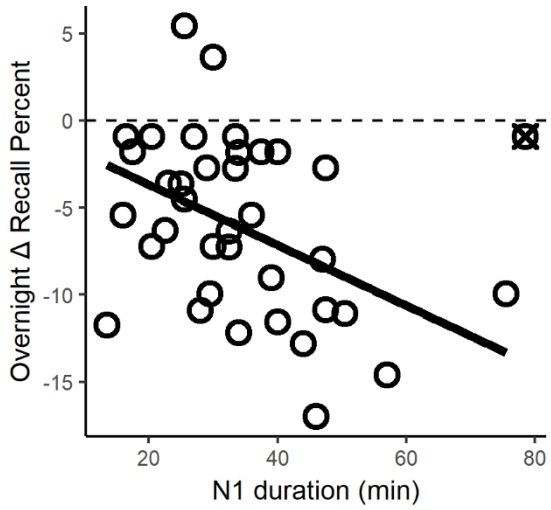

Figure 3.9. Overnight change in recall percent predicted by N1 sleep duration. The influential case excluded from linear models in figure and report is marked by the × symbol.

was estimated to decrease overnight retention of recall percent by 2.22 points (se = 0.79; Figure 3.9). N1 duration as a moderator of a cueing effect did not quite meet the stricter significance criterion with the same influential case excluded, $F(1, 34) = 6.51$, $p = .015$, $B = 3.48$, $se = 1.36$. This interaction was consistent with the interactions described for N3 sleep duration as it was opposite in direction. That is, an effect of cueing was not very clear at either low N1 duration of 25.25 min, $t = -1.47$, or high N1 duration of 40.00 min, $t = 1.09$, but relative to control items, retention of recall percent for cued items was estimated to be 1.66 points ($se = 1.52$) greater for those with more N1 sleep and 2.38 points ($se = 1.62$) lesser for those with less N1 sleep.

Other EEG power spectrum measures in the delta, theta, alpha, sigma, and beta bands from stages N2, N3, and R sleep beyond those already reported were also examined as predictors of overnight change in recall percent. In predicting this outcome measure,

only for N2 alpha power was there a significant sleep stage duration × average EEG power interaction, and only for R theta power was there a significant sleep stage duration × average EEG power interaction × condition interaction. There was also notable evidence of an N3 duration × N3 theta power that did not quite meet the criterion to limit false discovery rate. Briefly, the N3 duration × N3 theta power interaction was consistent with the previously discussed N3 duration × N3 delta power interaction though even more robust across multiple channels. It was largest by *B* value with the full scalp measure and characterized by an additional 25.00 min (1 SD) of stage N3 sleep being associated with an estimated 2.22-point (*se* = 0.95) increase in overnight retention of recall percent at high N3 delta power, $t = 2.33$, while having a less clear, negative association with this outcome at low N3 delta power, $B = -2.03$, $se = 1.33$, $t = -1.52$.

The N2 duration × N2 alpha power interaction was significant in adjusted samples at channel Pz, $F(1, 31) = 9.73$, $p = .004$, $B = -4.25$, $se = 1.36$, channel O1, $F(1, 31) = 10.47$, $p = .003$ $B = -3.19$, $se = 0.99$, channel Oz, $F(1, 31) = 9.12$, $p = .005$, $B = -3.43$, $se = 1.14$, and channel O2, $F(1, 30) = 15.77$, $p < .001$, $B = -3.62$, $se = 0.91$. All these tests were fit to adjusted samples excluding an influential case with extremely low N2 duration (-2.42 IQR) and another, non-extreme influential cases. An additional non-extreme influential case was excluded for the test at O2. A composite measure averaging N2 alpha power over these channels was created, and the interaction was significant, $F(1, 30) = 18.13$, $p < .001$, $B = -4.40$, $se = 1.03$, in an adjusted sample excluding the same three influential cases excluded for the test at O2 (Figure 3.10). The interaction was characterized by an additional 28.95 min (1 SD) of N2 sleep being associated with an estimated 4.16-point (*se* = 1.22) increase in retention of recall percent for those with low N2 alpha power, $t = 3.40$, and no clear association but an estimated 1.90-point (*se* = 0.83)

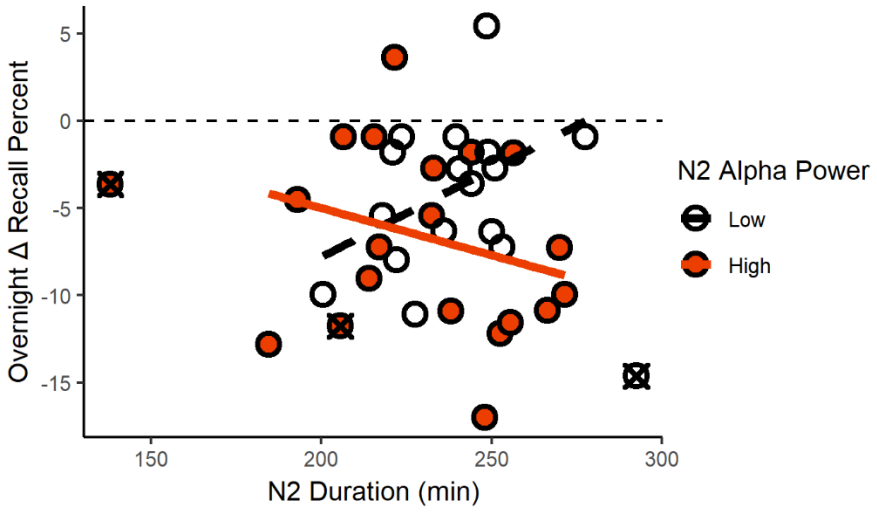

Figure 3.10. Overnight change in recall percent (post-sleep - pre-sleep) predicted by N2 duration on a median split of average alpha (8–11.5 Hz) electroencephalographic power in stage N2 epochs over channels Pz, O1, Oz, and O2. Influential cases excluded from linear models in figure and report are marked by the × symbol.

decrease in retention of recall percent for those with high N2 alpha power, t = .029.

The R duration × R theta power × condition interaction was significant in predicting overnight change in recall percent, indicating moderation of a cueing effect, specifically at channel P4, $F(1, 32) = 10.37$, p = .003, B = -5.85, se = 1.82, in an adjusted sample excluding one influential case with an extreme overnight change score for control item recall percent (+3.16 IQR). At low R theta power at P4, there, was a clear R duration × condition interaction that indicated a greater benefit of cueing (cue - control) on recall percent for those with greater R duration, B = 5.40, se = 2.10, t = 2.58, and was further characterized by an estimated cueing effect of -2.49 points (se = 1.78) at low R duration of 78.75 min, t = -1.40, and an estimated cueing effect of 3.08 points (se = 1.98) at high R duration of 104.75 min, t = 1.55. At high R theta power at P4, a R duration × condition

interaction effect indicated a reduced benefit of cueing on recall percent for those with greater R duration, B = -2.48, se = 1.51, t = -1.64, and was characterized by an estimated

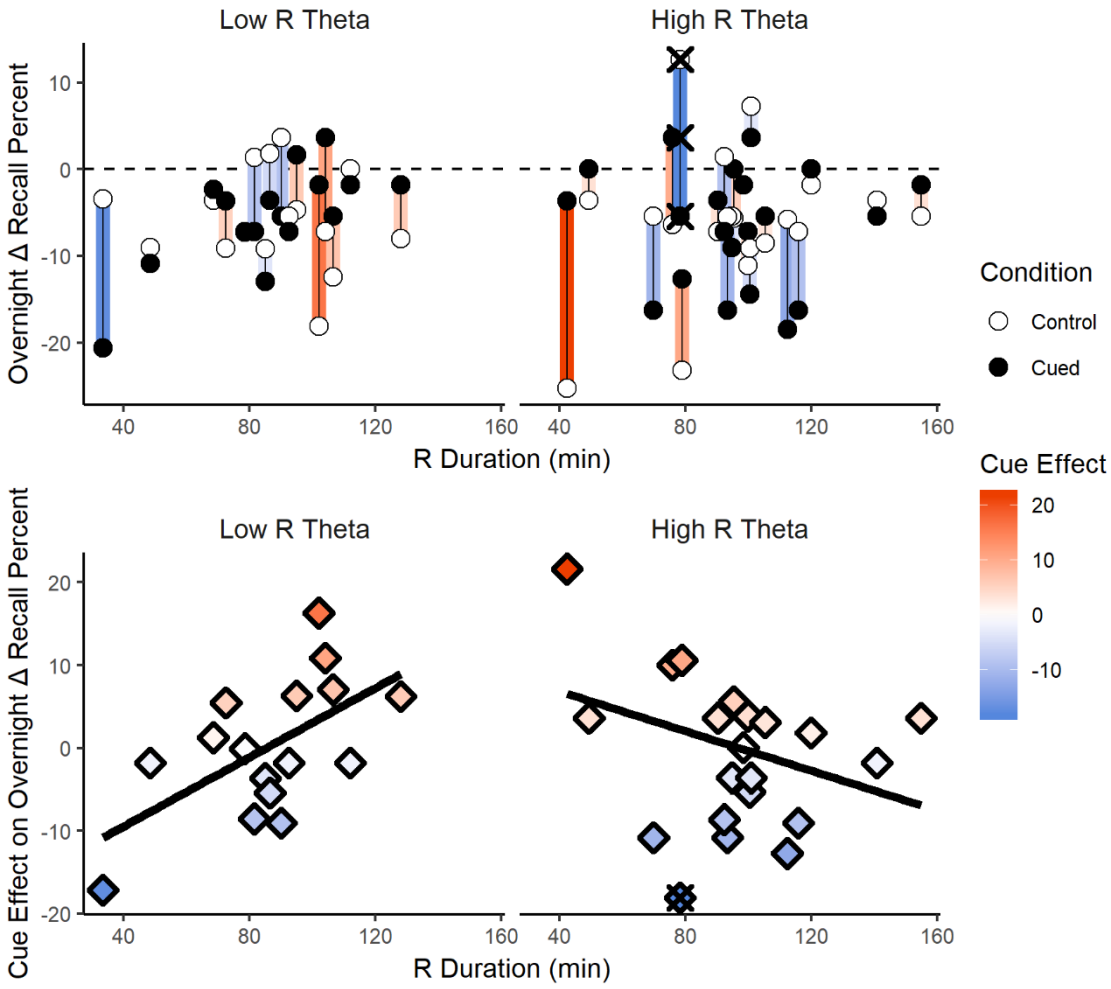

Figure 3.11. R duration moderating the effect of cueing (cued - control) on overnight change (post-sleep - pre-sleep) in recall percent by a median split of average theta (4–7.5 Hz) electroencephalographic power in stage R epochs at channel P4. Vertical lines in top panels connect overnight retention scores for cued and control items within each participant. Cueing effect depicted in colour and on y-axis in bottom panels. The influential case excluded from linear models in figure and report is marked by the × symbol.

cueing effect of 2.94 points (*se* = 1.89) at low R duration, *t* = 1.55, and a negligible estimated cueing effect of 0.38 points (*se* = 1.60) at high R duration, *t* = 0.24. Thus, there appeared to be an effect of the cueing to increase retention of approximate item locations for those with high amounts of stage R sleep with low theta power at P4 and for those with low amounts of stage R sleep with high theta power at P4 (Figure 3.11). Notably, R duration moderation was most clear for relatively low R theta power at P4, and the inclusion of the excluded case would further disrupt the weaker R duration moderation for relatievely high R theta power at P4.

Recall SD. A selection of the *F*-tests used to examine whether sleep measures were predictors of overnight change in recall SD own their own or as moderators of an effect of cueing are reported in Table 3.4.

Neither N3 duration nor R duration were significant predictors of overnight change in recall SD on their own or as moderators of a cueing effect on recall SD. A positive association between R duration and overnight change in recall SD was, however, marginally significant, $F(1, 34) = 3.54$, $p = .068$, with removal of one influential case with extremely low R duration (-1.76 IQR, 33.5 min), extremely high recall SD at pre-sleep (+1.76 IQR) and post-sleep (+5.44 IQR) tests, and an extreme overnight change score for recall SD (+3.99 IQR). This marginally significant association was characterized by an additional 23.45 min of R sleep (1 SD) predicting an estimated 0.61º (*se* = 0.32) increase in recall SD. Notably, this association is opposite in direction to the predictions made from the SR2 hypothesis in that greater duration of stage R sleep was associated with a relatively greater increase in error of recall over the night.

None of spindle count, spindle density, arousal count, arousal density, total sleep duration, N1 duration, N2 duration, the percent of sleep time spent in each stage, or any

Table 3.4
Tests of sleep measures as predictors of overnight change in recall SD on their own or as moderators of an effect of cueing

	As Predictor						As Moderator of Cueing Effect					
Measure	n	*F*	df_r	*p*	*B*	*se*	n	*F*	df_r	*p*	*B*	*se*
Sleep dur.	34[a1r3]	2.32	32	.137	0.43	0.28	33[r2]	1.74	29	.197	0.82	0.62
N1 dur.	35[r2]	2.36	33	.134	-0.46	0.30	33[r2]	2.67	34	.112	0.88	0.53
N2 dur.	36[r1]	1.10	34	.302	-0.35	0.34	34[a1r1]	0.90	33	.349	0.66	0.69
N3 dur.	36[r1]	0.72	34	.401	0.29	0.34	35	0.57	31	.455	-0.59	0.78
R dur.	36[r1]	3.54	34	.068	0.61	0.32	35	0.09	32	.763	-0.26	0.84
Arousals	35[r2]	0.56	33	.461	-0.24	0.32	34[r1]	1.01	34	.321	0.70	0.69
Spindles	35[r2]	0.42	33	.521	0.20	0.31	34[r1]	0.01	29	.941	0.05	0.65
Ind. delta	36[r1]	0.12	34	.734	0.12	0.34	35	0.23	33	.637	-0.38	0.79
Ind. theta	36[r1]	0.28	34	.601	-0.18	0.34	35	0.12	32	.728	0.27	0.78
Ind. sigma	36[r1]	0.00	34	.964	-0.02	0.34	34[r1]	0.02	31	.877	-0.11	0.68
N2 delta	36[r1]	0.14	32	.712	0.14	0.38						
N2 theta	37	0.00	33	.956	0.03	0.60						
N2 alpha												
N2 sigma	34[r3]	0.99	30	.328	-0.37	0.37						
N2 beta	35[r2]	0.32	31	.576	-0.25	0.44						
N3 delta												
N3 theta	34[r3]	2.56	30	.120	0.49	0.31	34[r1]	1.58	29	.218	-1.03	0.82
N3 alpha	35[r2]	0.75	31	.393	0.28	0.32	35	0.51	30	.480	-0.58	0.80
N3 sigma	35[r2]	0.07	31	.797	-0.10	0.39	33[r2]	1.30	28	.264	0.93	0.82
N3 beta	35[r2]	0.28	31	.599	0.27	0.50						
R delta	35[r2]	0.13	31	.719	-0.15	0.42						
R theta	35[r2]	0.14	31	.708	0.15	0.39	34[r1]	3.46	32	.072	-1.80	0.96
R alpha	36[r1]	0.40	32	.531	-0.27	0.42	35	1.60	29	.216	-1.33	1.05
R sigma	35[r2]	1.17	31	.287	-0.32	0.29	33[r2]	2.02	27	.167	-1.03	0.73
R beta	36[r1]	0.47	32	.500	-0.29	0.43	35	0.43	31	.517	-0.73	1.11

Note. Dur. = duration. Ind. = induced. Average electroencephalographic power measures tested as interactions with corresponding sleep stage durations. Numbers after superscript "a" indicate the number of influential cases that had predictor values winsorized for a given test. Numbers after superscript "r" indicate the number of influential cases removed from for a given test. Empty cells are shown for tests for which a model without influential cases could not be fit.

of the sleep stage duration × average EEG power measures were significant predictors of overnight change in recall SD.

Cue-induced power changes. The following tests examined whether the identified changes in delta, theta, and sigma EEG power induced by sound cues during

the night predicted overnight change in memory performance. *F*-tests used to examine whether the maximum channel measure for each induced power response predicted either recall percent or recall SD are reported in Table 3.3 and Table 3.4, respectively.

Recall percent. Induced delta power was a significant predictor of overnight change in recall percent on its own at channel C3, $F(1, 35) = 6.88$, $p = .013$, and channel Cz, $F(1, 34) = 5.30$, $p = .028$, $B = -1.89$, $se = 0.82$. An influential case with an extreme induced power response at C3 (+2.97 IQR) and Cz (+1.90 IQR) was winsorized to reduce influence in for the former case and excluded for the latter test. A 2.10 *z*-score (1 SD) increase in induced delta power from baseline at C3 was associated with an estimated 2.07-point ($se = 0.79$) decrease in overnight retention of recall percent (Figure 3.12a).

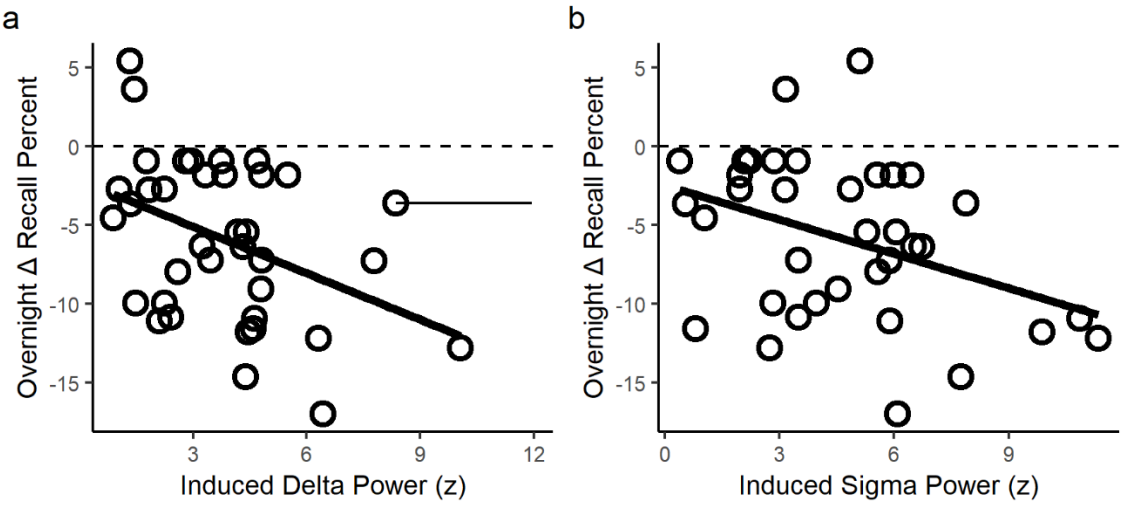

Figure 3.12. Overnight change in recall percent predicted by cue-induced increases in electroencephalographic power at C3. **a.** Overnight change in recall percent predicted by induced delta (1–3.5 Hz) power. Solid horizonal line connects an influential case that has been winsorized to its original value. **b.** Overnight change in recall percent predicted by induced sigma (12–15.5) power.

Induced sigma power at C3 was also significant predictor of overnight change in recall percent on its own, $F(1, 35) = 6.08$, $p = .019$. A 2.71 *z*-score (1 SD) increase in induced sigma power at C3 was associated with a similar estimated 1.96-point (*se* = 0.80) decrease in overnight retention of recall percent (Figure 3.12b). Thus, those who had a greater cue-induced increase in delta and sigma power retained fewer item locations over the night. Induced theta power was not a significant predictor of overnight change in recall percent on its own, and no induced power measure was as a significant moderator of cueing effect on recall percent.

Recall SD. No measures of induced delta power, induced theta power, or induced sigma power were significant predictors of overnight change in recall SD.

Discussion

Study 2 was designed to specifically test the SR2 hypothesis proposal that NREM sleep processes reinforce newly acquired memories, increasing their accessibility for later retrieval. TMR during slow wave sleep is thought to reactivate cued memory representations or at least bias reactivation events toward cued representations (Belal et al., 2018; Bendor & Wilson, 2012; Cairney et al., 2018) in a manner that often improves memory performance for cued items over control items (Oudiette & Paller, 2013; Schouten et al., 2017). It has been proposed that memory reactivation during NREM sleep is a mediator for the beneficial effect of sleep on memory (Diekelmann & Born, 2010; Rasch & Born, 2013), and spontaneous memory reactivation during NREM sleep has been associated with greater memory performance (Schönauer et al., 2017). Reactivation of memory representations during NREM sleep is proposed to be a mechanism of the reinforcing effects of NREM sleep outlined within the SR2 hypothesis. Thus, it was predicted that TMR during NREM slow wave sleep in the current experiment would

result in greater memory accessibility for cued items relative to control in the morning, and thus better overnight retention of cued item locations than those that were not cued during the night.

At first glance, the TMR procedure did not appear to affect memory performance, a finding inconsistent with earlier research showing positive effects of TMR during slow wave sleep on memory performance on visuospatial tasks (Rasch et al., 2007; Rudoy et al., 2009). However, identification of significant moderators of cueing effect suggests TMR had different effects within different people. Overnight retention of the approximate locations of cued items over control items tended to be greater among those with lower fidelity of recall before sleep (Figure 3.6), more stage R sleep so long as average theta power of stage R epochs was low, or less stage R sleep if their R epochs were high in average theta power (Figure 3.11). There was also slight evidence for positive cueing effect, or at least a less negative cueing effect, on recall percent among those with less stage N3 sleep (Figure 3.7) or more stage N1 sleep. Thus, there was some experimental evidence that memory reactivation in NREM sleep reinforced memory representations, maintaining their accessibility for retrieval.

That the TMR procedure may be more effective at increasing the retention of item locations among those who had lower recall fidelity before bed is consistent with past evidence of TMR-based improvement particularly for items learned with less accuracy (Creery et al., 2015). This pattern is also consistent with research showing low-value items that may otherwise be forgotten over sleep are better retained when they are cued during slow wave sleep (Oudiette et al., 2013). One account for why, in the current study, TMR during NREM slow wave sleep may have improved retention of item locations among those with lower pre-sleep recall fidelity is that low-fidelity memory

representations are likely at a greater risk of becoming inaccessible than high-fidelity representations and are thus more likely to benefit from reinforcement processes that would maintain their accessibility. This explanation is supported by another outcome of the current study: the finding that low pre-sleep recall fidelity was generally associated with overnight declines in ability to recall approximate item locations regardless of condition (Figure 3.5a).

The slight evidence for TMR during slow wave sleep providing a benefit to cued items for participants with relatively less slow wave sleep is consistent with past research showing that a short nap with odour based TMR during NREM sleep had an equivalent beneficial effect on memory performance as a longer 90-min nap opportunity (Diekelmann et al, 2012). These results, along with evidence that TMR cueing biases the content of memory reactivation (Bendor & Wilson, 2012) suggest TMR in the context of little NREM slow wave sleep supplements the effects of NREM slow wave sleep for cued items. In cases of greater amounts of NREM slow wave sleep, there may be a saturation of memory reinforcement that does not allow for further benefit from TMR cueing. The TMR method used in the current study may have been more effective in producing benefits if it had been delivered in a nap or other context in which relatively less slow wave sleep would be achieved. Although benefits of TMR during sleep are often observed in full-night sleep studies (e.g., Rasch et al., 2007; Rihm et al., 2014), this specific method of using sounds to cue memories in a visuospatial task has predominately been examined in nap studies (e.g., Creery et al., 2015; Oudiette et al., 2013; Rudoy et al., 2006).

It must be stressed that these and other instances of moderation were not well characterized by a TMR benefit in some participants (i.e., low pre-sleep recall fidelity or

low N3 duration) and an absence of an effect in their counterparts. Rather, there was evidence of a TMR decrement in some participants. Decreased ability to retrieve locations for cued items relative to control items was observed for participants with high pre-sleep recall fidelity (Figure 3.6), high pre-sleep recall percent, more stage N3 sleep (Figure 3.7), or less stage N1 sleep. Although selective benefits of TMR during sleep for some people or some items are more consistent with past research, the possibility of TMR-driven decrements to performance should evidently be considered. It may be that TMR instigated two processes: one which supplements endogenous memory reactivation but may saturate in effect and one which, in some conditions, decreases memory accessibility for cued items. One might envision a subsequent memory weakening process, perhaps occurring during REM sleep, that dampens memory representations that have become overly strong through, in the current study, a combination of deep learning, high stage N3 duration, and TMR during NREM sleep. Subsequent weakening of these memory representations relative to those for non-cued items might result in the observed TMR-associated decrements to memory accessibility.

The unexpected R duration × R theta power × condition interaction may further inform this speculation. This specificity for channel P4 is not easy to explain, but it is worth noting that this result withstood reasonable attempts to limit type I error, including an α of .01 and a .05 limit on the false discovery rate within the set testing this interaction at each channel and the full scalp measure. Within this interaction, greater stage R duration predicted TMR benefits for participants with low theta EEG power in stage R sleep but predicted relative TMR decrements for participants with high theta EEG power in stage R sleep (Figure 3.11). If REM sleep contains a refinement process of selective strengthening of dominant and weakening of non-dominant memory representations, as

proposed by the SR2 hypothesis, the enhanced cueing effect with greater R duration would be expected. Why then, for participants with high stage R theta power was greater R duration predictive of relative TMR decrements to memory accessibility? Poe et al. (2000) found that rat hippocampal place cells active in particular places fired during peaks of the hippocampal theta rhythm in REM sleep while those places were novel but fired during troughs of this rhythm in REM sleep after those places were familiar and proposed, based on evidence of phase-dependent bidirectional synaptic plasticity associated with the hippocampal theta rhythm (Hölscher et al., 1997; Huerta & Lisman, 1995; Pavlides et al., 1988), that the hippocampal theta rhythm in rat REM sleep may mediate a selective strengthening of novel memories and weakening of more familiar memories. In the current study, memories for cued items may have been treated as familiar due to repeated cued reactivations and thus weakened by theta-rich REM sleep. Further, non-cued items may have been treated as novel memories and strengthened during REM sleep theta peaks due to relative absence of reactivation in NREM sleep caused by a TMR-induced bias in memory reactivation (Bendor & Wilson, 2012). Either effect or both combined could have resulted in an absence of TMR benefit or a TMR decrement to memory accessibility. These TMR-decrements may have been present in the current study but not in the previous studies with a similar memory task and TMR procedure (e.g., Creery et al., 2015; Oudiette et al., 2013; Rudoy et al., 2006) because these studies were conducted using nap designs in which the opportunity for REM sleep would be greatly reduced.

TMR during NREM slow wave sleep separated cued and non-cued items on overnight retention of approximate item locations for some participants, but it had no significant impact on overnight change in recall SD either at large or in subsamples

evidenced by significant moderation. This pattern is consistent with the notion that the memory processing during NREM sleep is particularly suited to reinforcing memory traces and supporting accessibility rather than refining memory traces and supporting fidelity of recall.

A surprising outcome of the current study was an association between exhibiting greater cue-induced increases in delta and sigma EEG power and showing worse overnight retention of the ability to place items in their approximate locations (Figure 3.12). This finding is inconsistent with previous reports of cue-induced increases in fast spindles (Creery et al., 2015) and sigma power (Schreiner et al., 2015) being associated with positive effects of cueing on memory performance. Benefits of TMR during NREM sleep have likewise been associated with cue-induced increases in theta power (Schreiner et al., 2015). Notably, in the current study, the negative association between the response to cues and performance was not for cued items specifically but with memory for all items, and the cueing effect calculated relative control items was not associated with EEG response to cues. Memory representations may be reactivated in response to the sound cues, as demonstrated by EEG pattern-classification (Cairney et al., 2018), which may benefit the cued memories, but the current results suggest neural responsiveness to sound cues during NREM sleep may indicate or cause generally poor consolidation of newly acquired memories over the night. It may be that the cue-induced delta and sigma activity are indicative of efforts to inhibit arousals triggered by the sound cues given that both K-complexes and spindles (sleep events existing in delta and sigma ranges, respectively) are thought by many to protect sleep (Bastien et al., 2000; Dang-Vu et al., 2011; Dang-Vu et al., 2010; Jahnke et al. 2012). Arousal mechanisms may ultimately overcome these inhibitory efforts and disrupt important processes of sleep-dependent memory

consolidation. In contrast, participants who showed less induced delta and sigma activity may have been those who were less susceptible to the arousing potential of the sound cues.

Beyond the effects of TMR, it was predicted that NREM slow wave activity obtained over the night would be associated less overnight decline or overnight gains in memory accessibility, and this predicted result indeed came to pass. Significant N3 duration × N3 delta power interactions and similar N3 duration × N3 theta power interactions suggest that N3 sleep benefits overnight retention of approximate item locations so long as the scored N3 epochs are rich in low-frequency EEG activity (Figure 3.8). Notably, these interactions were most prominent at lateral, posterior channels, particularly P3, P4, O1, and O2. It was also found that participants with high N1 sleep duration or high N2 sleep duration with high alpha power tended to show a greater overnight decline in ability to place items in in their approximate locations (Figure 3.11). These findings are considered part of a general pattern in which greater retention of approximate item locations over sleep was associated with obtaining more relatively deep NREM sleep and less relatively shallow NREM sleep. Comparatively, there were no direct associations between properties of REM sleep and retention of approximate item locations observed, and properties of NREM sleep were not reliably associated with fidelity of recall. This pattern is consistent with the proposal that slow wave activity of NREM particularly supports memory accessibility.

The SR2 hypothesis proposes a memory refinement role for REM sleep. Thus, it was predicted that measures of REM sleep, particularly the amount of stage R sleep obtained, would be predictive of changes in fidelity of recall as captured by the measure of recall SD. This outcome was not observed. If anything, there was some indication that

better retention of recall fidelity was associated with obtaining less REM sleep. It may be that only a minimal amount of REM sleep is needed to refine newly acquired memory representations. Declines in memory performance being associated with greater REM sleep duration may be linked to the proposed selective strengthening and weakening processes proposed for REM sleep, and such declines are not unprecedented as Oudiette et al. (2013) found greater REM sleep duration to be associated with forgetting of low-value items.

One limitation of the current study lies in the large number of test items relative to previous work on TMR during sleep. The current study used a total of 200 items and cued 100 of them during sleep. To the author's knowledge, 200 is the largest number of items used to date in a TMR design and far exceeds the 50 items (25 cued) in the studies by Creery et al. (2015) and Rudoy et al. (2009) on which the current methodology was modeled. This large difference in the number of items limits the comparisons that can be made to past research of TMR during sleep. It is not clear how performance changes over sleep or the effect of cueing may differ as a result of the larger number of items and cues. However, as previously noted, the finding that cue-associated benefits to the recall of item locations were observed for those with low pre-sleep fidelity of recall is consistent with the results of Creery et al. (2015) despite the difference in number of items.

Another limitation present in the current, and most studies of TMR during sleep, is an inability to verify that memories are indeed reactivated by the cues. A memory benefit for items cued during sleep has been shown in various designs (Oudiette & Paller, 2013; Schouten et al., 2017). Evidence of reactivated memory representations following sound cues has been observed in rats through recordings of hippocampal neuronal firing patterns (Bendor and Wilson, 2012) and in humans through EEG pattern classification

(Belal et al., 2008). Thus, TMR during sleep is certainly possible. However, it is unknown whether the sound cues of the current study reliably reactivated memory representations for the paired item locations, and it is possible the effectiveness of sound cues for this purpose varied due to several factors. Some participants showed signs of arousal to sound cues while others appeared completely undisturbed by all 500 cues, inviting suspicion than sound cues were not loud enough to be processed at any level. Presumably, there is an ideal intensity at which sound cues may reliably produce TMR without causing an awakening or disrupting sleep, but this ideal intensity likely varies by accrued sleep time (Williams et al., 1964) and individual differences in sensitivity (Bastuji & García-Larrea, 1999; Zimmerman, 1970). Furthermore, the received intensity of sound cues of the current study varied by position of head relative to the headboard speakers. Varying the intensity of sound cues due to individual or situational factors in effort to achieve the ideal intensity may have resulted in clearer TMR effects. Alternatively, for some designs, odour based TMR may be preferred over sound based TMR due to a relatively low arousal threshold for odour stimuli (Carskadon & Herz, 2004).

Finally, as the current study had no experimental manipulation of REM sleep, the conclusions that can be made with respect to REM sleep are limited. Study 3 was specifically designed to test the components of the SR2 hypotheses related to REM sleep.

Chapter 4

Study 3: The Effect of Selective REM Sleep Deprivation

The primary focus of Study 3 was to investigate the REM sleep component of the SR2 hypothesis: the notion that REM sleep processes refine newly acquired memories, supporting fidelity of retrieval. For this purpose, participants received one of two manipulations: a REMD manipulation or a NREM sleep interruptions (NREMI) manipulation in which participants were given a pattern and number of awakenings during NREM sleep equal to the pattern and number of awakenings required for REMD. By comparing REMD and NREMI groups, the current study sought to identify the contributions of REM sleep to the processing of newly acquired memories with respect to memory accessibility and memory fidelity.

The SR2 hypothesis was formulated with attention to studies of sleep and memory using REMD. As argued in Chapter 1, REMD impairment of memory performance appears to be more likely when memory is tested in a manner requiring greater fidelity of recall. The sensitivity of more complex memory tasks to impairments from post-learning REMD has been noted for both avoidance task learning in rodents (Smith, 1985) and various memory tasks in humans (Smith et al., 2004; Tilley et al., 1992). For example, some studies have shown REMD after learning to impair memory for prose (Empson & Clarke, 1970; Tilley & Empson, 1978), but not significantly impair memory word pairs (Chernik, 1972; Ekstrand et al., 1971) or word lists (Empson & Clarke, 1970). Memory tasks with more complex material such as prose may draw more on memory fidelity than those with simple material such as word lists, but a more direct test of the SR2 hypothesis predictions regarding REM sleep refinement was required.

Potentially critical variations in method exist within past REMD studies in

humans, particularly in relation to the control group or condition used. Empson and Clarke (1970) compared REMD participants to paired control participants who would also be woke whenever their paired REMD participant showed signs of REM sleep. Tilley and Empson (1978) compared a group of REMD participants with a group of participants similarly deprived of Stage 4 NREM sleep. Chernik (1972) compared REMD participants to a group of performance-matched control participants who were given a similar number of awakenings and had a similar amount of sleep, and Lewin and Glaubman (1975) had a similar method but used a within-subjects design. Ekstrand et al. (1971) compared their REMD group to both a group deprived of Stage 4 NREM sleep and a group with a number of awakenings during lighter NREM sleep stages equal to the average of those needed to deprive yoked participants of REM sleep and Stage 4 NREM sleep. Furthermore, while the rest of these studies involved waking REMD participants at the sign of approaching REM sleep, Ekstrand et al. (1971) had participants wakened only at unambiguous REM sleep.

The current study expanded upon this past research by using the same item-location memory task used in Study 2 to assess changes in both memory accessibility and memory fidelity in response to REMD. For a REMD method, the current study implemented a procedure modeled after the one reported by Corsi-Cabrera et al. (2015). Of the memory studies noted, this REMD method was most like the method of Chernik (1972). This procedure involved repeatedly waking participants at the onset of REM sleep or suspected REM sleep to drastically limit the amount of REM sleep obtained during the night. This REMD group was compared to a control NREMI group who had a similar pattern of sleep interruptions throughout the night but had these interruptions occur during NREM sleep, specifically N1 and N2 sleep to avoid reductions in N3 sleep, of

which both groups would ideally obtain similar amounts. The current study also included the TMR procedure of Study 2 to be examined alongside the REMD and NREMI manipulations to determine whether an effect of cueing may vary depending on the presence of REM sleep. It was again expected that memory performance would decline over the night, but it was predicted that REMD participants would have greater overnight declines in fidelity of recall than NREMI participants given that REM sleep processes are hypothesized to support memory fidelity. As with Study 2, it was predicted, given the hypothesis that NREM sleep processes increase the accessibility of memories, overnight retention of approximate item locations would be associated with slow wave activity of NREM sleep and be greater for cued items than control items. Furthermore, slow wave activity of NREM sleep was predicted to negatively correlate with retention of recall fidelity in the REMD group. The rationale for this last predication was that, according to the SR2 hypothesis, a general amplification effect of NREM sleep processes may also amplify noise in the memory system and, without REM sleep refinement, this noise would persist and decrease fidelity of recall.

Method

Participants

The study was completed by 40 young adults from the Brock University community who were recruited to participant in a study examining the effect of fragmented sleep on memory for locations. All participants reported being good sleepers with typical daily sleep of approximately seven or more hours roughly within 22:00 and 9:00 and did not report working late or overnight shifts frequently in the past six months or traveling over multiple time zones in the past three months. All participants reported no history of psychiatric condition or head injury and reported not having any current

medical condition or regular substance use other than caffeine (e.g., medications, alcohol, cannabis) affecting sleep or cognitive function. All participants reported only low-to-moderate daily caffeine consumption (approximately <300 mg per day). On the days of the experiment, participants were asked to consume no substance affecting cognitive function or sleep (e.g., alcohol or caffeine), take no naps, obtain no vigorous exercise, and eat adequate meals so as to not be hungry before sleep. For the sleep period preceding those in the lab, participants were asked to go to sleep no later than 23:30 and to wake up at 7:30. All participants reported normal hearing and normal or corrected-to-normal vision. Normal colour vision was confirmed through completion of the Ishihara colour test (Ishihara, 2014), and normal hearing between 500 and 2000 Hz was confirmed through a hearing test with an audiometer conducted by a researcher. All participants provided informed consent, and they received an honorarium of up to $60 or course credit for participation in the study.

Two participants were excluded due to poor adherence to study directions and poor sleep on the experimental night. Both participants reported napping on the day of the experimental night and subsequently had poor sleep during the night. In both cases, less than the full set of 100 sound cues could be played during the night because participants frequently woke up to the sound cues. One of these participants was assigned to the REMD group and the other received a matching number of experimental awakenings in the NREMI group, but ultimately the two participants were excluded. An additional participant was excluded due to difficulty learning the item locations, having spent an exceptionally long time on the task and learning only 84% of the item locations. With these participants excluded, the total sample was 37 participants (25 female) with a mean age of 19.35 (*sd* = 1.83, minimum = 17, maximum = 25, Q_1 = 18, Q_2 = 19, Q_3 = 20).The

NREMI group had 18 participants (12 female) with a mean age of 19.33 (*sd* = 1.46, minimum = 17, maximum = 22, Q_1 = 18, Q_2 = 19, Q_2 = 19, Q_3 = 21, and the REMD group had 19 participants (13 female) with a mean age of 19.37, (*sd* = 2.17, minimum = 17, maximum = 25, Q_1 = 18, Q_2 = 18, Q_3 = 21. Of the 25 female participants, 7 were taking hormonal contraceptives (3 REMD), 11 were estimated to be in the follicular phase of the menstrual cycle (6 REMD), and 7 were estimated to be in the luteal phase of the menstrual cycle (4 REMD). All participants were found to perform at levels significantly greater than chance during recall tests.

Item-Location Memory Task

The memory tasks were nearly identical to those used in Study 2, including the learning period, recall tests, and recognition test of the item-location task and the sound discrimination test. As such, scripts for each part are available from MacDonald (2020) via the Open Science Framework. Readers may refer to Chapter 3 for task descriptions. One difference was that the minimum possible distances from the center of the circular gird for each item location was increased from 5.3º VA to 7.2º VA due to an observed tendency for greater error in responses for items that were closer to the centre of the grid and thus smaller on the screen. The only other difference in the tasks was the replacement of some picture and sound stimuli so that all stimuli were obtained from online repositories of free-to-use and attribution-free online repositories rather than having some stimuli come from sets made available for research purposes. These stimuli are made available from MacDonald and Schirmeister (2020) via the Open Science Framework. Again, normalization of loudness between the sound files was preformed using the normalization tool in Audacity® with subsequent amplitude adjustments made to files that were subjectively quieter or louder than the others.

Electrophysiological Recording

Electrophysiology was recorded using Neuroscan SynAmps2 amplifiers with Scan 4.5.1 software (Compumedics Inc., Abbotsford, Australia) and gold-plated silver electrodes sampling at a rate of 1000 Hz filtered DC to 200 Hz with an additional notch filter at 60 Hz. Recording parameters for the screening and adaptation night were identical to those used in Study 2. Recording parameters for the experimental sessions were nearly identical to those used in Study 2 with only two exceptions. The first of these exceptions was a replacement of the site O1 and site O2 scalp electrodes with electrodes at sites PO7 and PO8 to obtain wider scalp coverage. The second exception was the two EOG electrodes placed on the bicanthal plane were adjusted such that the left electrode was shifted 1 cm below the plane and the right electrode was shifted 1 cm above the plane. This change was made so that the recording montage was more compatible with the automated scoring software and more inline with American Academy of Sleep Medicine recommendations for sleep scoring (Berry et al., 2015). Readers may refer to Chapter 3 for more detailed descriptions of electrophysiological recording parameters. Electrical impedances were below 5 KΩ at all scalp sites and below 10 KΩ at all peripheral sites prior to recording.

Procedure

All procedures were cleared by the Brock University Bioscience Research Ethics Board. Recruitment, screening, and orientation procedures were nearly identical to those used in Study 2 with the only exception being that advertisements for the current study indicated that the study was examining the effects of fragmented sleep on memory for locations, whereas the advertisements for the previous study indicated the study was examining the role of sleep in memory for locations. Readers may refer to Chapter 3 for

other procedural details regarding participant recruitment, screening, and orientation.

Experimental session. The experimental session began at 22:30 the evening following the sleep screening and adaptation night and continued until approximately 9:45 the next morning. The experimental session procedures for before and after the sleep period were nearly identical to those used in Study 2 with the only exception being additional instruction immediately before the pre-sleep and post-sleep recall tests. Participants were encouraged to, if they felt they knew the location of a test item, to "take an extra moment" and be sure to place it in the exact location. This additional instruction was added due to concern that measurement of recall fidelity was limited by non-careful responding. Readers may refer to Chapter 3 for the other details of the before-sleep and after-sleep procedures.

Sleep period. The 8-hr sleep period started at approximately 23:30 and ended at approximately 7:30 for most participants with these times shifted later for those who competed the pre-sleep recall task behind schedule. Three participants received slightly longer sleep periods (0.5–2 min longer) to accommodate experimental awakenings near the end of their sleep period. One participant in the NREMI group had a 20-min longer sleep period because they were exhibiting REM sleep at the scheduled end of their sleep period and, to that point, they appeared to have obtained very little REM sleep despite being in the NREMI group. Immediately before the start of the sleep period, a researcher turned on the low-level white noise set to play throughout the sleep period, turned off the light and computer monitor, and guided participants through a short demonstration of the experimental awakening procedure that they would experience throughout the sleep period. The researcher then left the room and closed the bedroom door, leaving them to sleep until interrupted by experimental awakenings.

The TMR procedure, including the intensity of the white noise and sound cues, the matching of to-be-cued and non-cued items on learning and recall test variables, and the playing and termination of sound cues during sleep, was identical to that used in Study 2. Readers may refer to Chapter 3 for details of the TMR procedure.

Participants were assigned to be in either the REMD group or the NREMI group after the first bout of substantial N3 sleep. The first five participants in the study were assigned to the REMD group so that later participants could be assigned to the NREMI group, matched to a REMD participant, and given a number and pattern of experimental awakenings like that of their matched REMD participant. Participant matching was not random, but rather directed by the estimated number of sound cues that would be played during the night. For example, a participant who received 400–500 sound cues during their first bout of N3 sleep would be matched with a previous REMD participation who, over the course of their night, received all or nearly all 500 of their sound cues. Alternatively, if a participant was more easily aroused by the sound cues and had only received 100 sound cues during their first bout of N3 sleep, they would be matched to a previous REMD participant who received fewer sound cues, perhaps only 200, during their night. Each REMD participant was matched with only one NREMI participant and vice versa. When an appropriate match for a current participant was not available among previous and unmatched REMD participants, the current participant was assigned to the REMD group. In all cases, the number of sound cues played was attempted to be maximized within the TMR procedure, and the number of cues played was never adjusted to more closely match a participant.

Participant PSG was monitored throughout the night, and participants assigned to the REMD group were awakened by the experimenter whenever they appeared to enter

REM sleep. During online monitoring, onset of REM sleep was defined by the occurrence of rapid eye movements during sleep or roughly two minutes of low-amplitude, mixed-frequency EEG with reduced EMG activity and without spindles, K-complexes, or slow-rolling eye movements. The time since the start of recording and the time since the last experimental awakening were recorded for each experimental awakening. The pattern of experimental awakenings obtained from a REMD participant was used to guide those of a NREMI participant, considering, in order of importance: the total number of awakenings, the time between awakenings, the number of awakenings in a series happening close in time, and the distribution of awakenings over the night. NREMI awakenings were delivered during either online identified N2 or N1 sleep with effort not to disrupt N3 sleep. Thus, awakenings would typically be postponed until after a bout of N3 sleep and not during N2 sleep with increasing amounts of high-amplitude, low-frequency activity.

For an experimental awakening, a researcher would open the bedroom door, greet the participant, and instruct them to sit up in the bed. To prevent electrodes from falling off during this process, the researcher would approach the bed, lift the attached equipment as the participant sat upright, and place the equipment beside the participant while they sat. Participants typically awoke to the sound of the door opening. When participants did not wake to that sound, the researcher would knock on the wooden headboard to wake the participant. A 2-min timer would be started from when the participant awoke. To ensure the participant remained awake through its duration, the researcher verbally gave participants a mixture of two types of questions to solve verbally: simple arithmetic, or letter game. Arithmetic questions included counting backwards by odd numbers (e.g., “count backwards by 3 starting from 47.”), simple addition (e.g., “what is 12 plus 18?”), and simple multiplication (e.g., “what is 4 times 9?”). The letter game had the researcher

ask the participant to name examples of a category that begin with a given letter (e.g., “name countries that start with the letter S.”). Arithmetic errors were not corrected or acknowledged unless by request of the participant. If arithmetic errors were regularly made or signs of frustration or anxiety were observed, the difficulty of the questions was roughly and covertly adjusted to be simpler and a greater proportion of letter game questions were given. For the letter game, categories that overlap with items from the memory task (e.g., animals, instruments, or vehicles) were not given. After 2 min, the timer would beep, the researcher would move the EEG equipment back to the headboard while telling the participant they may lie down once again, and the researcher would then leave the room. To prevent rumination over the given questions, the researcher would complete any letter game or math problem left unanswered after the timer beeped. Lights were generally not turned on during experimental awakenings, though low-level light did enter rooms through the doorway while the door was open. On rare occasion, when necessary, experimental awakenings were used by participants as bathroom breaks or by the researcher as an opportunity to reapply electrodes that had become unfixed.

Data Analysis

Data analysis scripts for this study are available from MacDonald (2020) via the Open Science Framework.

Measures. Memory performance measures of recall percent and recall SD were acquired through methods identical to those used in Study 2 and nearly identical to those used in Study 1 with the only difference between the current study and Study 1 being the decrease in the defined recall failure criterion from 90º to 80º as was also the case for Study 2. Measures of sleep architecture, average EEG power spectra, and induced changes to EEG power spectra in response to cues were acquired through methods nearly

identical to those used in Study 2. Although they were obtained for Study 2, spindle count and spindle density measures were not obtained for the current study. Another difference in method was that sleep scoring in the current study did not require estimation of the EOG channels recommended by the American Academy of Sleep Medicine (Berry et al., 2015) and required by the automated scoring software because these EOG channels were already available from properly placed electrodes. The only other difference in method was that time-frequency bands of interest in the induced EEG power response to cues during the night were not determined solely through visual inspection. Instead, given that grand averages of response from the current study closely resembled those of Study 2, the same combinations of channels and time-frequency bands defined in Study 2 were also used in the current study. Readers may refer to Chapter 3 for details regarding the acquisition of performance and sleep measures.

Statistical analyses. Statistical analyses were conducted in R (R Core Team, 2017) using base functions and a variety of specialized packages, including *reshape2* (Wickham, 2007) for handling data structures and *ggplot2* (Wickham, 2016), *grid* (R Core Team, 2017), and *gridExtra* (Auguie, 2016) for creating figures.

***Data exclusion and adjustment*.** Methods for data exclusion and adjustment were identical to those of Study 2 and nearly identical to those of Study 1 with the only difference between the current study and Study 1 being the increase from 30 to 50 in the estimated instances of recall success required to allow the inclusion of the recall SD measure of a given performance as was also the case for Study 2. Readers may refer to Chapter 2 for details of these data exclusion and adjustment methods. As for Study 1 and Study 2, regression models fit without influential cases were the focus of follow-up analyses and evaluations of statistical significance when predicting performance

measures, and, when effects or associations are reported as significant in models fit to adjusted data, the specific exclusions and adjustments are reported.

All recall test performances (i.e., each participant for each recall test for cued items, control items, and both item types combined) were found to be greater than chance, and thus no performances were excluded for this reason. Recall SD scores were excluded if they were based on fewer than an estimated 50 instances of recall success. No recall SD scores for performance measures combining both cued and control items were excluded based on too few instances of recall success. A total of six participants had at least one of their recall SD scores excluded for performance measures involving either cued or control items alone. One of the six had both their pre-sleep and post-sleep control item recall SD scores excluded. Another one had both their cued item and control item post-sleep recall SD scores excluded, and another one had only their control item post-sleep recall SD scores excluded. Three of the six had only their post-sleep cued item recall SD scores excluded for this reason. For one of these three cases, control item post-sleep recall SD score was also excluded after this performance was identified as being only slightly above the criterion of an estimated 50 recall successes (51 recall successes) and being extremely high on recall SD (33º, +3.47 IQR); the corresponding overnight change score was also extreme (21.68º, +6.97 IQR) and also removed. Thus, no data was excluded for poor performance for analyses combining both cued and control items, but recall SD scores were excluded for 9 of 148 (6.1%) performances concerning only cued or only control items, and these exclusions came from 6 of 37 (16.2%) participants.

Statistical tests. The focus of statistical testing was to identify differences in performance between REMD and NREMI groups and identify any sleep or pre-sleep performance measures that predicted overnight changes in performance either on their

own or in interaction with cueing condition or sleep group. As was the case for Study 2, overnight changes in memory performance were examined through measures calculated by subtracting pre-sleep scores from post-sleep scores. Overnight change scores were not computed for cases in which either the pre-sleep or post-sleep recall SD scores were excluded; such cases were considered has having missing data for analyses of overnight change. Analyses not concerned with possible effects of cueing were conducted on all-item performance measures computed from the full distribution of response errors collapsing across cued and control items due to the greater confidence placed in performance estimates computed from larger distributions of response errors.

The strategies guiding the formation of regression models and statistical tests were the same as those for Study 1 and Study 2. A combination of simple (non-mixed-effect) linear regression models and mixed-effect linear regression models were used to examine how pre-sleep recall performance, sleep properties, group, and condition associated with memory performance. Simple regression models were used to examine associations between predictors, including group, and performance when each participant offered only one data point for each measure and mixed-effect regression models were used when participants had more than one data point each for a measure. Mixed-effect models were constructed using *lme4* (Bates et al., 2015). In all such models, a random intercept was allowed for each participant. Tests of associations, or "effects", in simple regression models were conducted using the *Anova* function in the *car* package (Fox & Weisberg, 2011) to perform analyses of variance *F*-tests using type II sums of squares. For mixed-effect regression models, the same function was used to perform type II Wald *F*-tests with denominator degrees of freedom approximated through the Kenward-Roger method (Kenward & Roger, 1997). All tests of interaction terms were conducted in models

containing the corresponding lower-order terms. Simple effects and slopes were estimated at the 25th and 75th percentiles of the interacting predictor variables. Continuous predictors in regression models were standardized ($m = 0$, $sd = 1$). *B* values from standardized predictors on unstandardized outcomes are reported to show the magnitude and direction of effects from linear models. Categorical variables of test time, condition, and group were effect coded such that coefficients indicated the effect of post-sleep test relative to pre-sleep test, the effect of cued condition compared to control condition, and the effect of REMD group compared to NREMI group.

Thus, the most critical tests were designed as follows. The tests of the effect of group on overnight change in either recall percent or recall SD were conducted within a simple regression model predicting the overnight change in the all-item performance measure. The tests of condition and condition × group interactions were conducted within mixed-effect regression models using both cue items and control item measures and including group, condition, and the condition × group interaction as terms predicting overnight change in the performance measure. Tests of the pre-sleep performance measures, sleep measures, or induced power measures as predictors of memory change over the night either on their own or in interaction with group or condition were conducted within four regression models for both overnight change in recall percent and overnight change in recall SD. Tests of each measure as a predictor on their own were conducted in simple linear regression models containing only that measure as a predictor and used all-item performance data. Group and its interaction with the measure were added to these models to test whether sleep group moderated the association. Tests of each measure as a moderator of an effect of cueing were conducted in mixed-effect regression models which included condition and its interaction with the measure and were

constructed using separate overnight change measures for cued items and control items. Finally, group and all associated interaction terms were added to this model to test whether any moderation of condition was further moderated by sleep group. As was the case for Study 2, for measures of average EEG power in a given sleep stage, analyses were not concerned with the average EEG power measure on its own but the interaction between the EEG power measure and the duration of the associated sleep stage (e.g., N3 duration × N3 delta power). Thus, for each of these tests, the measure examined as a potential predictor was this interaction.

The following additional tests were also conducted. A one-sample *t*-test against 0 was used to test whether participants could discriminate between cued and non-cued sounds in the morning. Pearson's correlation coefficient was calculated to quantify the association between memory performance measures within the same test. Two-sample *t*-tests were used to compare group means (REMD vs. NREMI) on measures of sleep; degrees of freedom were adjusted using the Welch method (Welch, 1951) when Levene's test indicated the groups had unequal variances.

Statistical significance was defined through *p* value less than alpha of .05 for tests concerning hypotheses and measures of considerable practical or theoretical interest. More specifically, this was the only significance criterion when testing the effect of test time, condition, or group and when testing the association between memory performance and the following sleep measures: N3 duration, R duration, and the R duration × R theta power interaction at channel P4, with the latter measure examined only in the NREMI group. Other theoretically interesting measures were considered as sets of related measures. One set included the N3 duration × N3 delta power interaction for channels P3, Pz, P4, PO7, Oz, and PO8 and the full scalp measure; another set included the N2

duration × N2 sigma power interaction for the full scalp measure and each channel, and the N3 duration × N3 sigma power interaction for the full scalp measure and each channel. For measures of the induced power response to cues during the night, each quantified time-frequency band response of interest had its own defined set consisting of the different EEG channels (including the maximum channel measures) for which the time-frequency band response was quantified. Within these defined sets of measures, false discovery rate was controlled for using the Benjamini–Hochberg method (Benjamini & Hochberg, 1995) to limit false discovery rate to 0.10. Thus, for these measures, statistical significance was defined as having a *p* value less than .05 and the critical *Q* determined to limit false discovery to 0.10. Tests on recall percent were always considered in a different set than tests on recall SD. Alpha was lowered to .01 for tests of other sleep measures, including total sleep duration, N1 duration, N2 duration, the percent of total sleep duration spent in each sleep stage, arousal count, arousal density, and average EEG power at other channels and in other frequency of stages N2, N3, and R sleep. For these average EEG power measures, the Benjamini–Hochberg method was applied to limit false discovery rate to .05 within each similarly constructed set with a separate set for each frequency band within stage N2 sleep and stage N3 sleep and each frequency band within stage R sleep. Brief explorations of results of marginal significance, defined as those meeting the criterion of limit false discovery but exceeding α by no more than 0.04, are reported. Average EEG power from stage R sleep was only examined as a predictor within the NREMI group because the R sleep in the REMD group was considered unreliable based on the small number of epochs and possible changes in frequency distributions due to the REMD manipulation.

Results

Sleep

Descriptive statistics for sleep architecture variables and select power spectral variables separated are reported in Table 4.1 for the NREMI group and Table 4.2 for the

Table 4.1
Descriptive statistics for sleep architecture and select EEG power measures for the NREMI group

Measure		n	*m*	*sd*	min	max	Q_1	Q_2	Q_3
Awakenings		18	17.22	4.72	9.00	24.00	14.00	18.00	21.00
Sleep dur.		18	382.39	42.42	292.50	449.50	346.50	387.25	420.00
N1 dur.		18	58.67	27.42	26.00	127.00	38.00	49.75	75.00
N2 dur.	*	18	185.22	47.73	86.50	272.00	155.00	186.50	208.00
N3 dur.		18	66.22	21.44	27.50	104.00	54.00	63.25	81.50
R dur.	*	18	72.28	23.92	34.00	110.50	52.00	72.75	93.00
N1 %		18	15.62	7.83	6.18	36.92	10.25	12.73	17.86
N2 %	*	18	48.08	9.62	25.15	67.08	44.29	47.50	54.74
N3 %		18	17.54	6.13	7.25	30.01	14.73	15.55	23.45
R %	*	18	18.76	5.64	9.22	28.44	15.01	17.70	24.51
Arousals		18	86.50	26.32	41.00	122.00	63.00	85.00	113.00
Arousal den.		18	0.23	0.08	0.10	0.36	0.18	0.21	0.31
N2 delta		18	-10.31	0.15	-10.54	-10.01	-10.43	-10.27	-10.23
N2 theta		18	-11.13	0.19	-11.43	-10.81	-11.31	-11.10	-11.00
N2 alpha		18	-11.60	0.17	-11.92	-11.37	-11.73	-11.56	-11.48
N2 sigma		18	-11.62	0.20	-11.95	-11.29	-11.78	-11.57	-11.47
N2 beta		18	-12.61	0.15	-12.83	-12.30	-12.73	-12.66	-12.49
N3 delta		18	-9.58	0.20	-9.97	-9.21	-9.76	-9.56	-9.49
N3 theta		18	-10.88	0.19	-11.17	-10.58	-11.06	-10.87	-10.70
N3 alpha		18	-11.55	0.23	-11.92	-11.14	-11.77	-11.53	-11.41
N3 sigma		18	-11.78	0.22	-12.12	-11.44	-11.94	-11.77	-11.61
N3 beta		18	-12.77	0.17	-13.00	-12.39	-12.91	-12.81	-12.63
R delta		18	-10.81	0.15	-11.06	-10.47	-10.90	-10.84	-10.72
R theta		18	-11.36	0.14	-11.61	-11.15	-11.45	-11.36	-11.24
R alpha		18	-11.75	0.17	-12.00	-11.42	-11.89	-11.75	-11.68
R sigma		18	-12.26	0.15	-12.52	-11.92	-12.37	-12.24	-12.19
R beta		18	-12.66	0.15	-12.89	-12.36	-12.79	-12.66	-12.52
Ind. delta		18	3.60	2.45	-0.23	7.41	1.58	3.25	6.06
Ind. theta		18	4.19	2.41	0.32	8.39	2.25	3.64	5.55
Ind. sigma		18	4.52	2.76	-0.09	10.19	2.68	4.77	6.07

* measures for which the REMD and NREMI groups differed significantly ($\alpha = .05$)
Note. Awakenings refers to those induced as part of the NREMI procedure. Duration (dur.) reported in minutes. Density (den.) reported as count per minute of sleep. Induced (ind.) power measured at central channels as z-scores relative to pre-stimulus interval.

Table 4.2
Descriptive statistics for sleep architecture and select EEG power measures for the REMD group

Measure		n	*m*	*sd*	min	max	Q_1	Q_2	Q_3
Awakenings		19	17.58	4.54	9.00	24.00	14.00	18.00	22.00
Sleep dur.		19	373.84	32.90	286.00	429.00	351.50	384.00	393.00
N1 dur.		19	58.71	18.64	34.50	90.00	39.00	60.50	70.50
N2 dur.	*	19	230.47	28.67	185.00	292.50	208.00	227.50	253.00
N3 dur.		19	77.11	27.04	3.50	109.00	58.00	84.50	99.50
R dur.	*	19	7.55	4.09	2.50	17.50	4.00	6.00	11.00
N1 %		19	15.73	4.77	8.28	23.06	11.89	15.78	19.55
N2 %	*	19	61.89	7.79	50.34	84.09	56.91	60.70	65.52
N3 %		19	20.35	6.93	1.22	28.83	15.05	20.57	25.91
R %	*	19	2.03	1.11	0.71	4.76	1.06	1.56	2.85
Arousals		19	100.68	26.57	62.00	155.00	79.00	96.00	123.00
Arousal den.		19	0.27	0.08	0.16	0.42	0.19	0.26	0.32
N2 delta		19	-10.23	0.09	-10.54	-10.12	-10.27	-10.21	-10.18
N2 theta		19	-11.16	0.14	-11.46	-10.95	-11.27	-11.10	-11.06
N2 alpha		19	-11.57	0.18	-11.96	-11.34	-11.64	-11.52	-11.46
N2 sigma		19	-11.65	0.25	-12.21	-11.21	-11.82	-11.64	-11.48
N2 beta		19	-12.68	0.14	-13.03	-12.47	-12.71	-12.67	-12.61
N3 delta		19	-9.55	0.19	-10.12	-9.22	-9.66	-9.52	-9.43
N3 theta		19	-10.88	0.15	-11.31	-10.65	-10.96	-10.86	-10.81
N3 alpha		19	-11.47	0.25	-11.97	-10.93	-11.60	-11.43	-11.35
N3 sigma		19	-11.79	0.23	-12.32	-11.43	-11.94	-11.72	-11.62
N3 beta		19	-12.81	0.14	-13.14	-12.60	-12.84	-12.79	-12.73
Ind. delta		19	3.20	1.55	1.19	6.39	2.07	2.96	3.99
Ind. theta		19	3.49	1.59	1.29	5.99	1.90	3.56	5.39
Ind. sigma		19	4.96	2.69	1.78	10.16	2.68	3.89	6.74

* measures for which the REMD and NREMI groups differed significantly ($\alpha = .05$)
Note. Awakenings refers to those induced as part of the REMD procedure. Duration (dur.) reported in minutes. Density (den.) reported as count per minute of sleep. Induced (ind.) power measured at central channels as *z*-scores relative to pre-stimulus interval.

REMD group. Correlations among the number of cues played, the number of experimental awakenings, and select sleep measures are depicted in Figure 4.1 for the full sample, Figure 4.2 for the NREMI group, and Figure 4.3 for the REMD group.

The REMD procedure was successful in limiting average duration of stage R sleep to 7.55 min in the REMD group, much less than the average 72.28 min in the NREMI group, $t(18) = -11.32$, $p < .001$. Conversely, the NREMI group had significantly

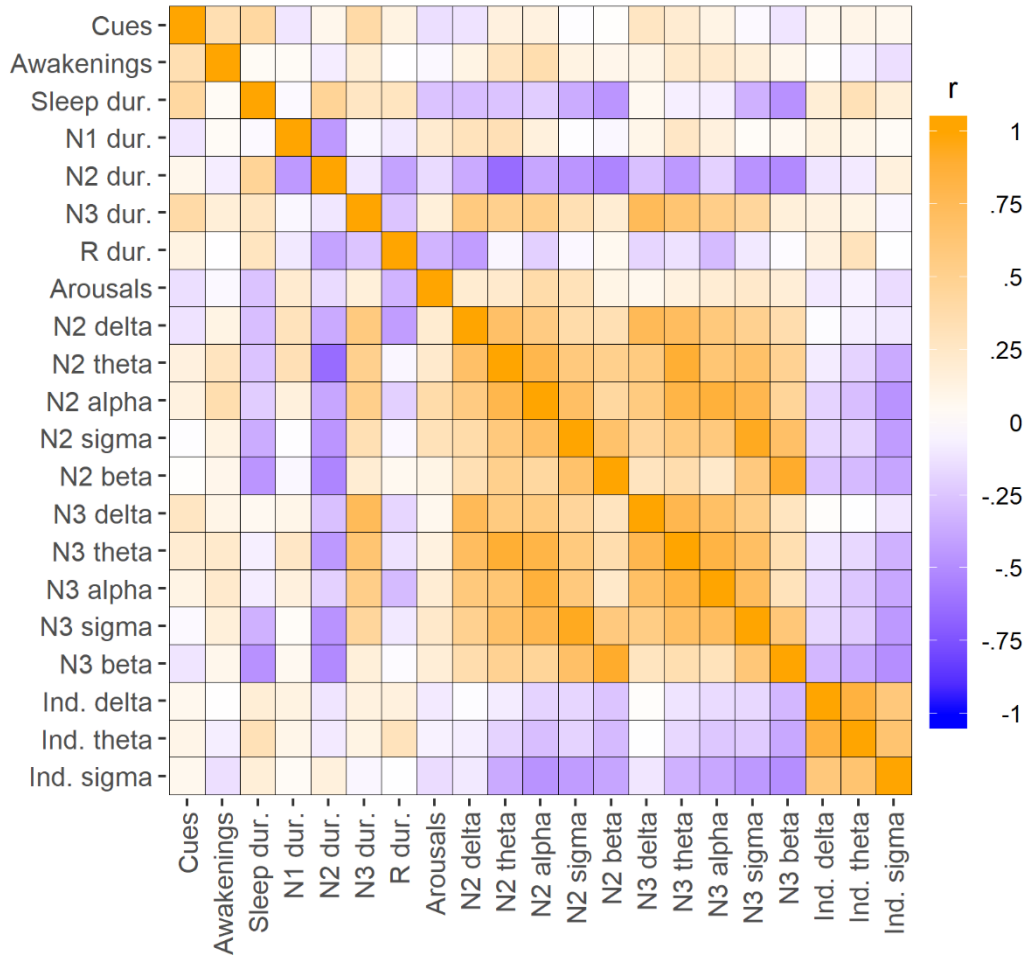

Figure 4.1. Full sample Pearson correlation matrix of cue count and sleep measures with *r* values depicted on a coloured scale. Dur. = duration. Ind. = induced. Cues indicates the number of sound cues played during the night. Awakenings refers to those induced as part of the REMD and NREMI procedures.

less stage N2 sleep, $t(35) = 3.52$, $p = .001$. The groups likewise differed on the percent of total sleep spent in stage R, $t(18) = -12.36$, $p < .001$, and stage N2, $t(35) = 4.81$, $p < .001$. On average, participants received 17.41 ($sd = 4.57$) experimental awakenings, and the number of these awakenings did not differ between groups, $t(35) = 0.23$, $p = .816$.

The NREMI and REMD groups did not differ significantly in total sleep duration,

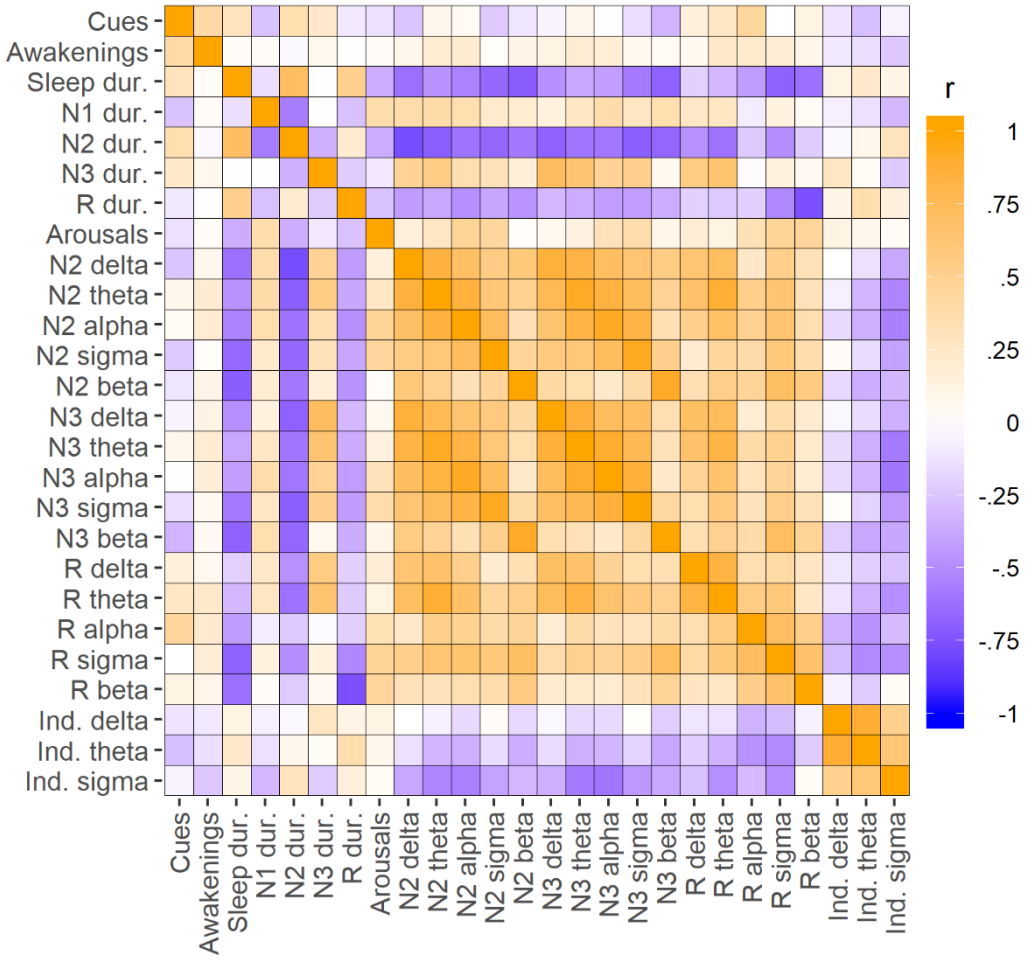

Figure 4.2. NREMI group Pearson correlation matrix of cue count and sleep measures with *r* values depicted on a coloured scale. Dur. = duration. Ind. = induced. Cues indicates the number of sound cues played during the night. Awakenings refers to those induced as part of the NREMI procedure.

t(35) = -0.69, *p* = .497; N1 duration, *t*(35) = 0.01, *p* = .995; N3 duration, *t*(35) = 1.35, *p* = .185; the percent of sleep time spent in stage N1 sleep, *t*(35) = 0.05, *p* = .959, or stage N3 sleep, *t*(35) = 1.3, *p* = .200; arousal count, *t*(35) = 1.63, *p* = .112, or density, *t*(35) = 1.52, *p* = .137; average delta power in epochs of stage N2, *t*(35) = 1.89, *p* = .067, or stage N3, *t*(35) = 0.50, *p* = .622; average theta power in epochs of stage N2, *t*(35) = -0.61, *p* = .548,

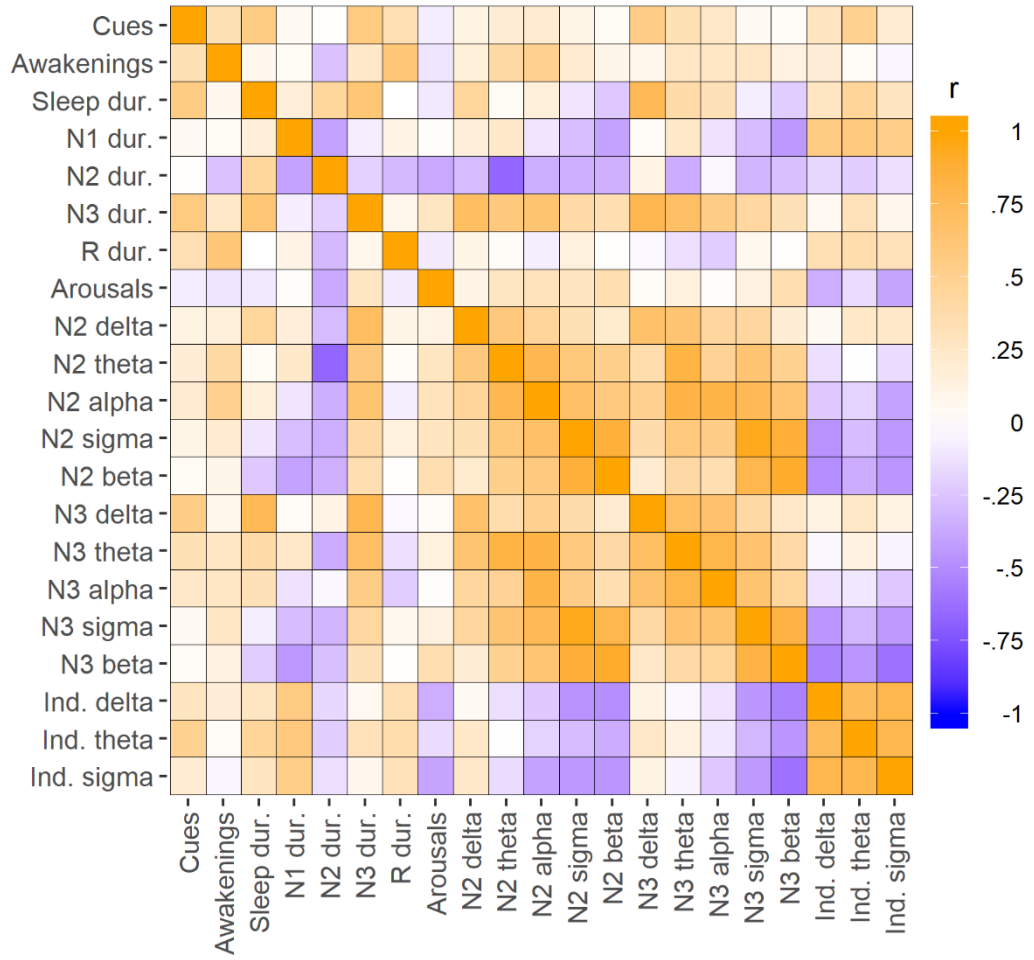

Figure 4.3. REMD group Pearson correlation matrix of cue count and sleep measures with *r* values depicted on a coloured scale. Dur. = duration. Ind. = induced. Cues indicates the number of sound cues played during the night. Awakenings refers to those induced as part of the REMD procedure.

or stage N3, $t(35) = -0.04$, $p = .972$; average alpha power in epochs of stage N2, $t(35) = 0.35$, $p = .730$, or stage N3, $t(35) = 1.07$, $p = .293$; average sigma power in epochs of stage N2, $t(35) = -0.42$, $p = .675$, or stage N3, $t(35) = -0.20$, $p = .846$; or average beta power in epochs of stage N2, $t(35) = -1.44$, $p = .157$, or stage N3, $t(35) = -0.70$, $p = .491$. The models for group comparisons of N1 duration and N1 percent each contained an

influential case from the NREMI group with the maximum N1 duration and percent; winsorizing this case on both measures did result in significant group differences. The models for group comparisons of N3 duration and average N3 delta power each contained an influential case from the REMD group with the full-sample minimum N3 duration and the group minimum N3 delta power; winsorizing this case on N3 delta power did not result in a significant group difference for N3 delta power, but excluding this case for N3 duration resulted in the REMD group (m = 81.19 min, sd = 20.93) having significantly more N3 sleep than the NREMI group (m = 66.22, sd = 21.44), $t(34) = 2.12$, $p = .041$. Still, the two groups were reasonably well matched apart from the disruptions of stage R sleep and stage N2 sleep and a tendency for NREMI participants to have less stage N3 sleep, save for one REMD participant with very little stage N3 sleep.

Response to Cueing

On average 386.38 (sd = 111.41, minimum = 142, maximum = 500, Q_1 = 323, Q_2 = 408, Q_3 = 500) sound cues were played to participants during the night. Of the 37 participants in the sample, only 11 received all 500 sound cues, and 9 participants received less than 300 sound cues, meaning not all the 100 cues were played at least three times. Of all sound cues played, 93.2% were played in epochs later marked as stage N3 sleep, 6.6% were played in epochs later scored as stage N2 sleep, 0.1% were played in epochs later marked as stage N1 sleep, and 0.1% were played in epochs later scored as wake with the latter two categories the result of arousals immediately following the cue. No participants reported hearing sounds when probed with the question of, “other than all the awakenings, did you notice anything strange or unusual last night?” Participants were unable to reliably indicate which sounds were played during the night, with the mean d' score of -0.03 (sd = 0.20, minimum = -0.52, maximum = 0.36), $t(36) = -0.76$, p = .453.

REMD and NREMI groups did not differ in the number of sound cues played, $t(35) = -0.9$, $p = .377$. For the NREMI group, an average of 403.28 ($sd = 98.52$, minimum = 202, maximum = 500, $Q_1 = 333$, $Q_2 = 412.5$, $Q_3 = 500$) sound cues were played with six participants receiving all 500 sound cues and three participants receiving less than 300 sound cues. For the REMD group, an average of 370.37 ($sd = 122.89$, minimum = 142, maximum = 500, $Q_1 = 265$, $Q_2 = 408$, $Q_3 = 500$) sound cues were played with five participants receiving all 500 sound cues and six participants receiving less than 300 sound cues. Figure 4.4 depicts the effective matching of REMD and NREMI participants on both the number of experimental awakenings and the number of sound cues played during the night by plotting participants on these axes and connecting them with their matched participant from the other group. As shown in the figure, most REMD and

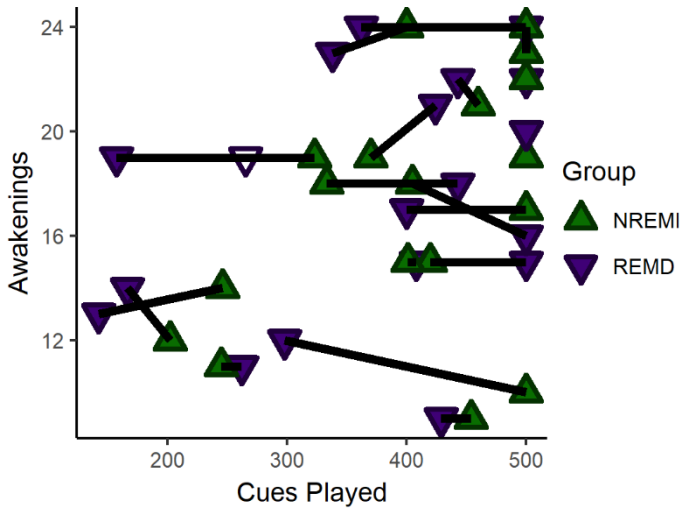

Figure 4.4. Plot of the number of cues played and number of awakenings for each participant with lines connecting each participant to their matched participant from the other group. The empty triangle indicates a participant for whom their matched pair was excluded.

NREMI participants were very well matched on the number of awakenings and reasonably well matched with few of exceptions on the number of cues played.

The grand averages for the induced EEG power response to the sound cues during sleep are depicted in Figure 4.5. Visual inspection of the grand and individual averages identified a response pattern to similar the one observed in Study 2: an increase in delta

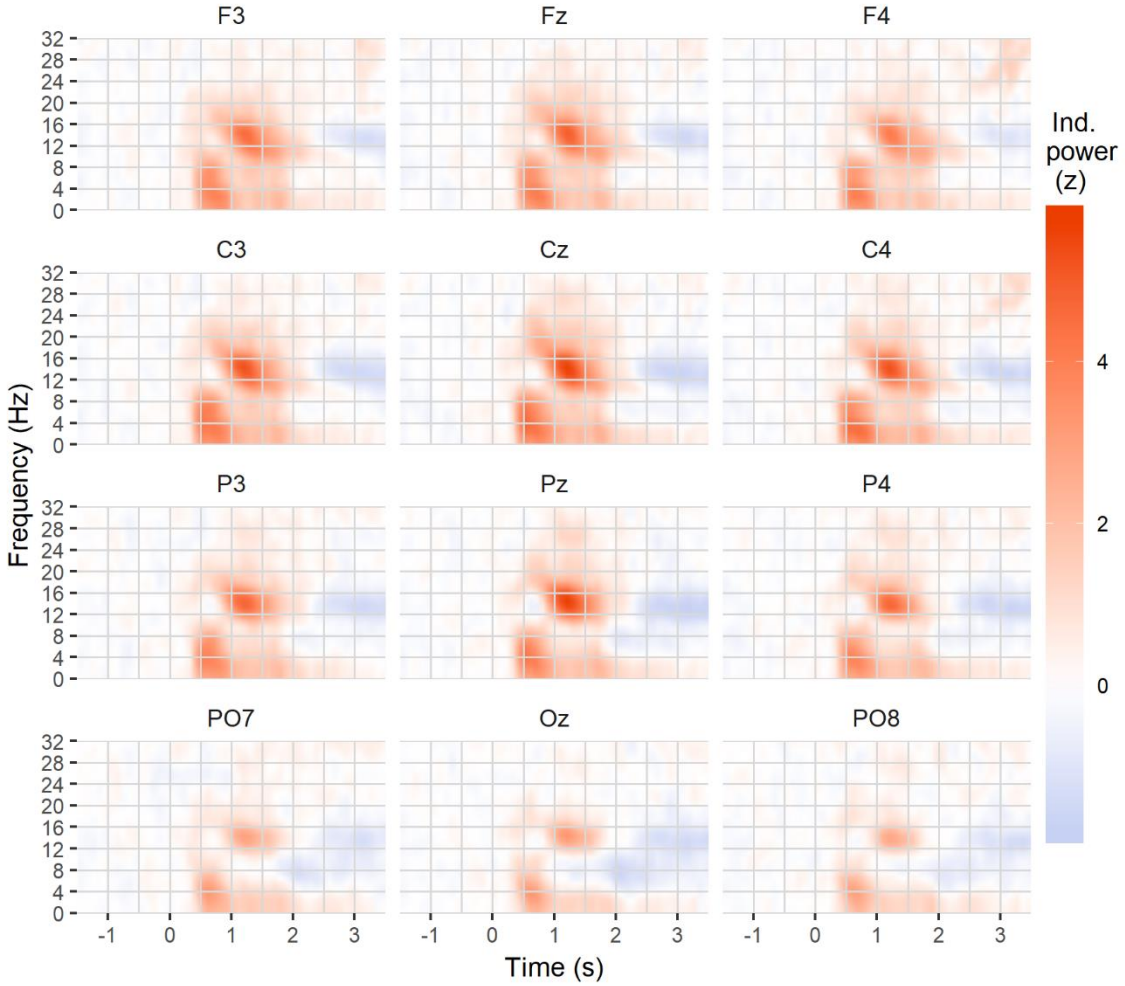

Figure 4.5. Average induced electroencephalographic power response to cues across all 37 included participants at 12 scalp channels. Induced power values with a resolution of 0.1 s and 1 Hz were calculated as z-scores relative to the -1.5–0.0 s pre-stimulus interval and are represented on a colour scale.

power 0.5–1.5 s post cue, an increase in theta power 0.5–1.0 s post cue, and an increase in sigma power 1.0–1.75 s post cue. As in Study 2, power in these time-frequency bands of interest was quantified for each participant for channels C3, Cz, and C4, along with a "Cmax" measure capturing each participant's maximum induced power value over these channels. Descriptive statistics for these maximum channel measures of induced power responses are reported in Table 4.1 and Table 4.2 for the NREMI and REMD groups, respectively. The NREMI and REMD groups did not differ significantly induced delta power, $t(29) = -0.6$, $p = .556$, induced theta power, $t(35) = -1.05$, $p = .301$, or induced sigma power, $t(35) = 0.48$, $p = .632$.

Learning Performance

The average duration of the learning period was 87.85 min (*sd* = 17.38, minimum = 62.12, maximum = 139.63, $Q_1 = 74.18$, $Q_2 = 85.32$, $Q_3 = 95.57$). On average, participants learned 99.41% of the items (*sd* = 1.72). One participant learned only 180 (90%) of the items; all other participants learned at least 195 (97.5%) of the items, and 25 (68%) participants learned all 200 items to criterion.

Pre-Sleep and Post-Sleep Recall Performance

Descriptive statistics for pre-sleep and post-sleep recall measures are reported in Table 4.3 for the NREMI group and Table 4.4 for the REMD group. The distributions of response errors for all learned items, separated by recall test and sleep group are shown in Figure 4.6. The average duration of the pre-sleep recall test was 17.28 min (*sd* = 4.02, minimum = 11.15, maximum = 26.93, $Q_1 = 14.52$, $Q_2 = 16.30$, $Q_3 = 19.98$). The average duration of the post-sleep recall test was 18.20 min (*sd* = 5.64, minimum = 11.07, maximum = 32.80, $Q_1 = 14.40$, $Q_2 = 16.38$, $Q_3 = 21.18$).

At the pre-sleep test, participants across both groups, on average, recalled 79.03%

Table 4.3
Descriptive statistics for memory performance in recall tests for the NREMI group

Items	Measure	n	*m*	*sd*	Min	max	Q_1	Q_2	Q_3
All	Pre-sleep recall %	18	78.98	12.96	52.61	98.10	70.06	80.46	89.95
	Pre-sleep recall SD	18	11.30	2.52	7.53	17.41	9.50	10.79	13.16
	Pre-sleep confidence	18	2.20	0.20	1.81	2.50	2.04	2.25	2.40
	Post-sleep recall %	18	70.96	14.53	49.83	94.48	59.91	72.42	83.62
	Post-sleep recall SD	18	12.17	3.15	6.70	19.50	9.99	11.62	13.69
	Post-sleep confidence	18	2.14	0.24	1.69	2.55	1.97	2.21	2.31
	Overnight Δ recall %	18	-8.02	6.58	-25.27	-0.90	-10.85	-4.97	-3.62
	Overnight Δ recall SD	18	0.86	2.08	-1.52	6.33	-0.83	0.48	1.86
	Overnight Δ confidence	18	-0.06	0.12	-0.27	0.15	-0.13	-0.06	-0.01
Control	Pre-sleep recall %	18	78.10	13.61	46.78	97.19	70.06	79.10	88.15
	Pre-sleep recall SD	17	10.96	2.77	6.18	15.97	9.24	10.51	11.95
	Pre-sleep confidence	18	2.20	0.21	1.80	2.53	2.03	2.24	2.40
	Post-sleep recall %	18	71.45	15.36	44.92	95.38	59.20	75.12	82.72
	Post-sleep recall SD	16	11.65	3.10	5.34	17.69	9.51	11.64	14.04
	Post-sleep confidence	18	2.14	0.24	1.67	2.55	1.93	2.22	2.25
	Overnight Δ recall %	18	-6.65	7.94	-25.41	5.43	-9.05	-6.33	-1.81
	Overnight Δ recall SD	16	0.66	2.33	-4.01	5.75	-0.72	0.45	1.73
	Overnight Δ confidence	18	-0.06	0.12	-0.27	0.14	-0.17	-0.05	0.03
Cued	Pre-sleep recall %	18	79.85	12.58	57.39	99.00	70.06	80.91	89.95
	Pre-sleep recall SD	18	11.32	2.91	7.10	19.08	9.38	11.02	12.98
	Pre-sleep confidence	18	2.20	0.20	1.82	2.47	2.03	2.25	2.40
	Post-sleep recall %	18	70.75	14.64	46.73	94.11	57.39	67.03	84.53
	Post-sleep recall SD	17	11.64	3.17	6.96	19.45	9.95	11.17	12.30
	Post-sleep confidence	18	2.14	0.25	1.70	2.55	1.98	2.19	2.30
	Overnight Δ recall %	18	-9.10	8.20	-25.14	5.43	-12.66	-7.24	-3.62
	Overnight Δ recall SD	17	0.20	1.83	-3.04	4.82	-0.52	0.38	0.72
	Overnight Δ confidence	18	-0.06	0.13	-0.27	0.18	-0.13	-0.07	0.01

Note. All-item recall percent and confidence differed significantly ($\alpha = .05$) from pre- to post-sleep. No significant effects of condition (cued vs. control) or group (REMD vs. NREMI) were detected. Recall SD measured in degrees of error. Confidence rated on a 1–3 scale.

Table 4.1
Descriptive statistics for memory performance in recall tests for the REMD group

Items	Measure	n	*m*	*sd*	min	max	Q_1	Q_2	Q_3
All	Pre-sleep recall %	19	79.08	10.57	56.49	89.95	69.00	81.81	88.15
	Pre-sleep recall SD	19	12.29	2.81	8.48	18.19	10.44	11.76	14.71
	Pre-sleep confidence	19	2.16	0.28	1.66	2.68	1.96	2.19	2.34
	Post-sleep recall %	19	71.12	13.97	43.82	89.05	60.82	74.37	84.53
	Post-sleep recall SD	19	12.42	2.83	8.85	19.46	10.54	11.44	14.16
	Post-sleep confidence	19	2.10	0.30	1.42	2.66	1.94	2.11	2.25
	Overnight Δ recall %	19	-7.96	5.09	-17.19	0.90	-11.84	-8.18	-3.62
	Overnight Δ recall SD	19	0.13	1.98	-3.61	2.65	-1.15	0.40	2.16
	Overnight Δ confidence	19	-0.07	0.14	-0.24	0.27	-0.17	-0.09	0.02
Control	Pre-sleep recall %	19	79.76	10.50	59.20	91.76	71.86	81.85	88.39
	Pre-sleep recall SD	19	12.77	3.46	7.54	19.86	9.12	12.80	14.99
	Pre-sleep confidence	19	2.16	0.27	1.65	2.67	1.98	2.19	2.32
	Post-sleep recall %	19	72.04	13.80	48.10	89.95	59.20	73.67	82.91
	Post-sleep recall SD	17	12.66	3.69	5.85	21.17	10.93	11.85	15.03
	Post-sleep confidence	19	2.08	0.29	1.42	2.64	1.90	2.08	2.28
	Overnight Δ recall %	19	-7.71	5.74	-19.90	3.62	-12.79	-7.24	-3.62
	Overnight Δ recall SD	17	-0.06	2.44	-6.11	3.82	-0.87	0.04	1.77
	Overnight Δ confidence	19	-0.08	0.15	-0.23	0.34	-0.20	-0.09	0.00
Cued	Pre-sleep recall %	19	78.66	11.14	53.77	89.95	66.11	82.72	86.53
	Pre-sleep recall SD	19	11.97	3.35	7.65	20.95	9.47	10.83	13.39
	Pre-sleep confidence	19	2.17	0.28	1.66	2.69	1.94	2.20	2.36
	Post-sleep recall %	19	70.49	15.14	35.68	89.95	61.01	73.67	80.91
	Post-sleep recall SD	16	11.36	2.02	8.42	16.05	9.61	11.24	12.80
	Post-sleep confidence	19	2.11	0.31	1.41	2.69	1.91	2.13	2.27
	Overnight Δ recall %	19	-8.17	5.94	-18.10	3.43	-12.79	-9.05	-3.81
	Overnight Δ recall SD	16	0.45	1.84	-2.22	3.45	-1.35	0.38	1.94
	Overnight Δ confidence	19	-0.05	0.14	-0.26	0.21	-0.14	-0.08	0.04

Note. All-item recall percent and confidence differed significantly ($\alpha = .05$) from pre- to post-sleep. No significant effects of condition (cued vs. control) or group (REMD vs. NREMI) were detected. Recall SD measured in degrees of error. Confidence rated on a 1–3 scale.

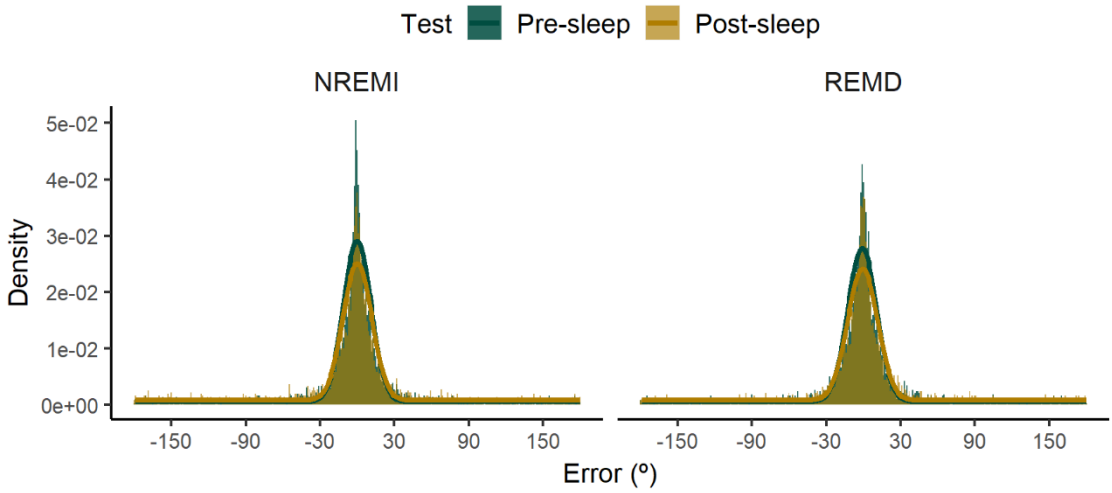

Figure 4.6. Distributions of responses measured as error in angular distance from target location. Solid lines indicate density of a mixed distribution composed of a uniform distribution of guesses and a normal distribution of successful recall with parameters fit to the corresponding distribution of errors. Data separated by sleep group and coloured by time of recall test.

(*sd* = 11.63) of learned item locations with a recall SD of 11.81° (*sd* = 2.68) and a confidence rating of 2.18 (*sd* = 0.24). There was a negative correlation between pre-sleep recall percent and pre-sleep recall SD, $r(35) = -.38$, $p = .022$, indicating that those who recalled a greater percentage of learned item locations also had higher fidelity of recall. Pre-sleep recall confidence was positively correlated with pre-sleep recall percent, $r(35) = .52$, $p = .001$, and negatively correlated with pre-sleep recall SD, $r(34) = -.44$, $p = .008$, indicating those with higher recall confidence recalled more item locations and recalled them with greater accuracy. The latter correlation was calculated in an adjusted sample excluding one influential case without extreme scores on these measures. Still with both groups combined, pre-sleep control item performance and pre-sleep cued item

performance were not significantly different on recall percent, $F(1, 36) = 0.20$, $p = .657$, recall SD, $F(1, 35) = 0.28$, p .599, or confidence, $F(1, 36) = 1.34$, $p = .254$.

The NREMI group and REMD group were not significantly different at pre-sleep recall test on recall percent, $F(1, 35) < 0.01$, $p = .979$, recall SD, $F(1, 35) = 1.25$, $p = .271$, or confidence, $F(1, 35) = 0.22$, $p = .641$. A non-significant pre-sleep recall SD × group interaction effect on pre-sleep recall percent, $F(1, 32) = 0.20$, $p = .662$, indicated that the relationship between pre-sleep recall SD and pre-sleep recall percent did not differ significantly between groups. Relationships between pre-sleep confidence and other performance measures also did not differ significantly between groups for pre-sleep recall percent, $F(1, 33) = 0.16$, $p = .694$, or pre-sleep recall SD, $F(1, 33) = 1.03$, $p = .317$. Tests of condition × group interactions indicated that differences, or lack thereof, between control items and cued items at pre-sleep recall test did not vary significantly by group for recall SD, $F(1, 34) = 1.06$, $p = .310$, or confidence, $F(1, 35) = 1.22$, $p = .278$. The condition × group interaction was significant for pre-sleep recall percent, $F(1, 35) = 5.55$, $p = .024$. Although the interaction was significant, cue and control item pre-sleep recall percent did not differ within the REMD group, $F(1, 18) = 1.87$, $p = .189$, or within the NREMI group, $F(1, 17) = 3.72$, $p = .071$. Thus, NREMI and REMD groups had similar performance before sleep and had similar relationships between performance measures, and control item performance and cued item performance were generally well matched within NREMI and REMD groups.

At the post-sleep test, participants across both groups, on average, recalled 71.04% ($sd = 14.04$) of learned item locations with a recall SD of 12.30º ($sd = 2.95$) and a confidence rating of 2.12 ($sd = 0.27$). There was still a negative correlation between recall percent and recall SD after sleep, $r(35) = -.35$, $p = .034$. Likewise, there was still a

positive correlation between recall confidence and recall percent, $r(34) = .51$, $p = .001$, and a negative correlation between recall percent and recall SD, $r(34) = -.35$, $p = .039$. These correlations with confidence were calculated in an adjusted sample with an influential case with extremely low post-sleep confidence (-1.82 IQR) removed. Control item performance was not significantly different from cued item performance after sleep for recall percent, $F(1, 36) = 0.86$, $p = .360$, recall SD, $F(1, 32) = 1.19$, $p = .283$, or confidence, $F(1, 36) = 1.83$, $p = .185$.

The NREMI group and REMD group were not significantly different at post-sleep recall test on recall percent, $F(1, 35) < 0.01$, $p = .972$, recall SD, $F(1, 35) = 0.07$, $p = .800$, or confidence, $F(1, 35) = 0.21$, $p = .653$. The relationship between recall SD and recall percent did not differ between sleep groups, as indicated by a non-significant post-sleep recall SD × group interaction on post-sleep recall percent in the unadjusted sample, $F(1, 33) = 0.73$, $p = .399$. Likewise, a non-significant post-sleep recall percent × group interaction, $F(1, 33) = 0.21$, $p = .649$, and a non-significant post-sleep recall SD × group interaction, $F(1, 33) < 0.01$, $p = 994$, indicated there were no group differences in the association between post-sleep confidence and either post-sleep recall percent or post-sleep recall SD. Tests of condition × group interactions indicated there were no group differences in the differences between control items and cued items at post-sleep recall test for recall percent, $F(1, 35) = 0.12$, $p = .730$, recall SD, $F(1, 31) = 0.71$, $p = .405$, or confidence, $F(1, 35) = 1.10$, $p = .301$. Thus, the NREMI and REMD groups did not differ in performance at post-sleep recall test.

Overnight Change in Recall Performance

The focus of analyses and predictions was on changes in memory performance measures over the night from the pre-sleep recall test to the post-sleep recall test as these

measures were thought to be more sensitive to effects of sleep processes. Descriptive statistics for overnight change scores are reported in Table 4.3 for the NREMI group and Table 4.4 for the REMD group.

For the all-item performance measures in both NREMI and REMD groups combined, the percent of learned item locations recalled was significantly lower post-sleep compared to pre-sleep, $F(1, 36) = 70.63$, $p < .001$, with an average 7.99-point (se = 0.95) decrease in recall percent over the night. Recall SD did not change significantly from pre-sleep test to post-sleep test, $F(1, 36) = 2.11$, $p = .155$, $B = 0.49$, $se = 0.33$. Participants were less confident at the post-sleep test relative to the pre-sleep test, $F(1, 36) = 9.49$, $p = .004$, with an average rating decrease of 0.06 (se = 0.02). With and only with exclusion of a non-extreme influential case, there was a significant correlation between overnight change in recall percent and overnight change in recall SD, $r(34) = .38$, $p = .021$. Overnight change in recall confidence was significantly, positively correlated with overnight changes in recall percent with and only with exclusion of an influential case with extreme change score for recall percent (-1.95 IQR), $r(34) = .44$, $p = .007$, but was not correlated with overnight changes in recall SD, $r(35) = .15$, $p = .371$, with the latter correlation calculated after exclusion of two influential cases, one with an extreme change score for recall SD (+1.67 IQR). Altogether, these results indicate that participants who had the greatest declines in ability to recall item locations tended to have greater decreases in average confidence, but greater fidelity of recall on those recalled successfully.

The following sections will first examine the effect of sleep group and cueing condition on overnight change in recall performance then examine multiple pre-sleep performance and sleep measures as potential predictors of overnight changes in

performance.

Sleep group and cueing condition. Against all predictions regarding experimental manipulations, there was no evidence for any overall main effect of sleep group or cueing condition nor any evidence for a condition × group interaction effect on overnight changes in recall percent, recall SD, or confidence. In regression models predicting overnight changes, there was no significant main effect of group for recall percent, $F(1, 34) = 0.31$, $p = .579$, $B = -0.96$, $se = 1.71$ (in an adjusted sample excluding one influential case), for recall SD, $F(1, 35) = 1.21$, $p = .280$, $B = -0.73$, $se = 0.67$, or for confidence $F(1, 35) = 0.01$, $p = .941$, $B = 0.00$, $se = 0.04$; there was no significant main effect of condition for recall percent, $F(1, 36) = 1.22$, $p = .277$, $B = -1.42$, $se = 1.29$, for recall SD, $F(1, 33) = < 0.01$, $p = .977$, $B = 0.01$, $se = 0.49$, or for confidence, $F(1, 36) = 0.92$, $p = .343$, $B = 0.01$, $se = 0.01$; and there was no significant condition × group interaction for recall percent, $F(1, 35) = 0.58$, $p = .450$, $B = 1.99$, $se = 2.60$, for recall SD, $F(1, 32) = 1.08$, $p = .306$, $B = 1.00$, $se = 0.96$, or for confidence, $F(1, 35) = 0.45$, $p = .508$, $B = 0.02$, $se = 0.03$. This absence of effect was confirmed in regression models including recall test (pre-sleep and post-sleep) as a factor rather than examining the overnight change scores.

Pre-sleep performance. Measures of pre-sleep recall percent, pre-sleep recall SD, and pre-sleep confidence were examined as potential predictors of overnight change in performance.

Recall percent. Overnight change in the percent of learned item locations recalled was not significantly predicted by pre-sleep recall percent, $F(1, 35) = 1.67$, $p = .205$, $B = 1.23$, $se = 0.95$, pre-sleep recall SD, $F(1, 35) = 3.94$, $p = .055$, $B = -1.84$, $se = 0.93$, or pre-sleep confidence, $F(1, 35) = 0.38$, $p = .539$, $B = 0.60$, $se = 0.97$. The marginally

significant association between pre-sleep recall SD and overnight change in recall percent was characterized by an estimated 1.84-point decrement to recall percent for every 2.68º (1 SD) increase in pre-sleep recall SD. This tendency for overnight decrements in recall percent to be greater for those with lower fidelity recall before sleep is consistent with significant findings of Study 2.

There was no evidence that sleep group moderated associations between overnight change in recall percent and any of pre-sleep recall percent, $F(1, 33) = 2.42$, $p = .129$, $B = 3.00$, $se = 1.93$, pre-sleep recall SD, $F(1, 33) = 0.08$, $p = .775$, $B = -0.56$, $se = 1.95$, or pre-sleep confidence, $F(1, 33) = 0.04$, $p = .852$, $B = 0.40$, $se = 2.13$. There was also no evidence that a cueing effect was moderated over both groups by pre-sleep recall percent, $F(1, 35) = 0.20$, $p = .656$, $B = 0.59$, $se = 1.31$, by pre-sleep recall SD, $F(1, 35) = 0.27$, $p = .607$, $B = 0.68$, $se = 1.31$, or by pre-sleep confidence, $F(1, 35) < 0.01$, $p = .970$, $B = -0.05$, $se = 1.32$.

Tests of the pre-sleep recall measure × condition × group interactions indicated that the pre-sleep measure × condition interaction effects on overnight change in recall percent varied by sleep group for pre-sleep recall SD, $F(1, 32) = 4.40$, $p = .044$, $B = -5.71$, $se = 2.72$ (in an adjusted sample excluding one non-extreme influential case), and for pre-sleep confidence, $F(1, 33) = 7.35$, $p = .011$, $B = 7.03$, $se = 2.59$, but not for pre-sleep recall percent, $F(1, 33) = 0.21$, $p = .647$, $B = 1.26$, $se = 2.72$. Pre-sleep recall SD × condition and pre-sleep confidence × condition interactions were tested separately within each group.

A pre-sleep recall SD × condition interaction effect on overnight change in recall percent was observed in the NREMI group, $B = 4.33$, $se = 2.05$, $t = 2.12$, but not in the REMD group, $B = -0.35$, $se = 1.46$, $t = -0.24$, for whom cueing had no clear effect on

overnight retention of recall percent relative to control at mean pre-sleep recall SD, *B* = 0.46, *se* = 1.45, *t* = -0.32 (Figure 4.7). In the NREMI group, there was no clear effect of

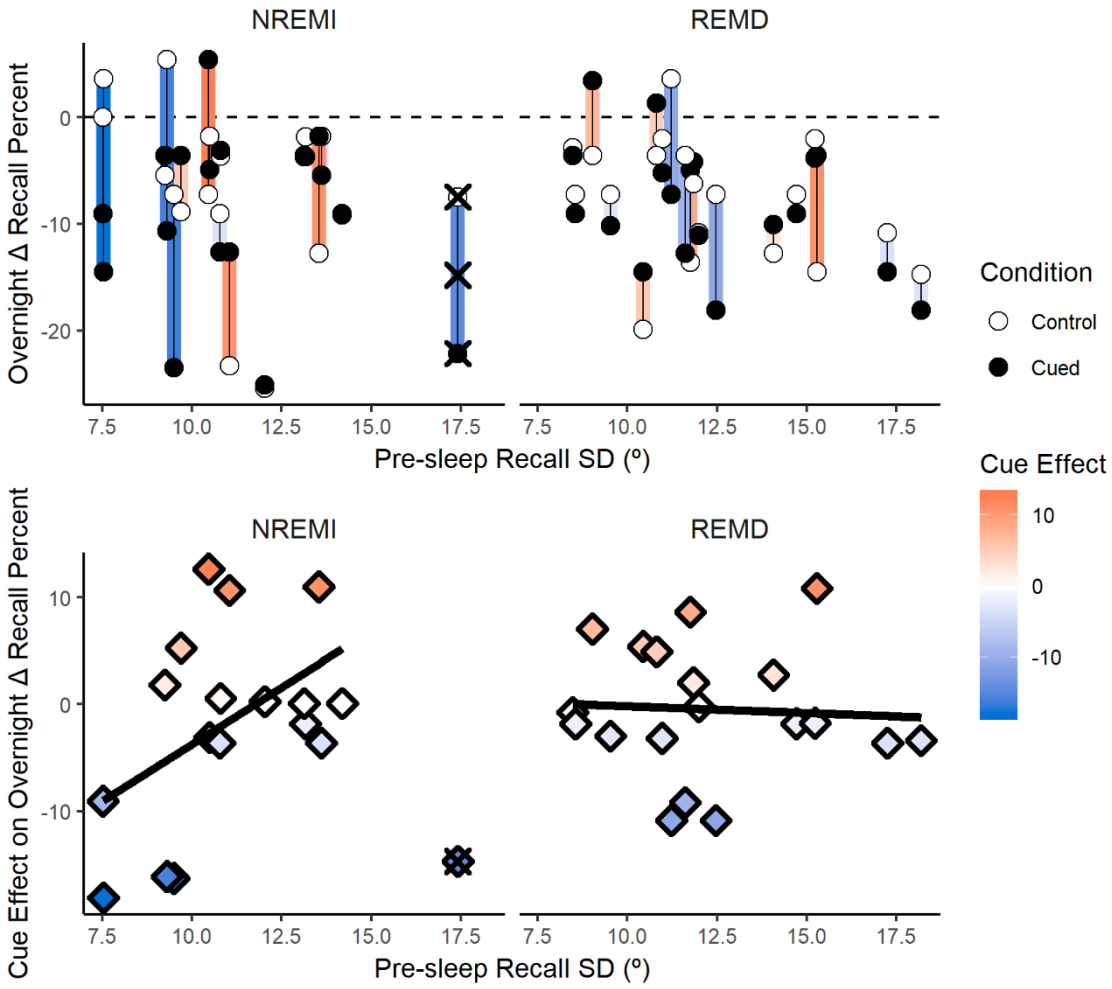

Figure 4.7. Association between pre-sleep recall SD and the cueing effect (cued - control) on overnight change (post-sleep - pre-sleep) in recall percent separated by sleep group (NREMI vs. REMD). Vertical lines in top panels connect overnight retention scores for cued and control items within each participant. Cueing effect depicted in colour and on y-axis in bottom panels. The influential case excluded from linear models in figure and report is marked by the × symbol.

condition at high pre-sleep recall SD, $B = 2.93$, $se = 2.98$, $t = 0.98$, but, at low pre-sleep recall SD, estimates indicated an effect of condition to decrease recall percent for cued items by 4.80 points ($se = 2.45$) relative to control items, $t = -1.93$.

A pre-sleep confidence × condition interaction effect on overnight retention of recall percent was evident in the REMD group, $B = 2.71$, $se = 1.31$, $t = 2.07$, and the NREMI group, $B = -3.87$, $se = 2.11$, $t = -1.83$ (Figure 4.8). Estimated simple effects of cueing indicated no clear effect at relatively low pre-sleep confidence in the NREMI group, $B = 0.77$, $se = 2.72$, $t = 0.28$, no clear effect at relatively high pre-sleep confidence in the REMD group, $B = 1.27$, $se = 1.54$, $t = 0.82$, slight evidence for a 2.51-point ($se = 1.63$) decrease at relatively low pre-sleep confidence in the REMD group, $t = -1.54$, and a large 6.41-point ($se = 3.00$) decrease at relatively high pre-sleep confidence in the NREMI group, $t = -2.14$. This interaction suggests that cueing decreased recall of item locations for NREMI participants with high pre-sleep average confidence and REMD participants with low pre-sleep average confidence.

Recall SD. Overnight change in recall SD was not significantly predicted by pre-sleep recall percent, $F(1, 33) = 0.15$, $p = .702$, $B = 0.11$, $se = 0.29$ (with two influential cases removed), by pre-sleep recall SD, $F(1, 35) = 2.15$, $p = .152$, $B = -0.49$, $se = 0.33$, or by pre-sleep confidence, $F(1, 34) = 0.78$, $p = .385$, $B = -0.29$, $se = 0.33$ (with one influential case removed).

Sleep group did not moderate an association between overnight change in recall SD and either pre-sleep recall SD, $F(1, 33) = 0.51$, $p = .480$, $B = -0.50$, $e = 0.69$, or pre-sleep confidence, $F(1, 31) = 2.16$, $p = .151$, $B = 1.02$, $se = 0.69$ (with two influential cases removed), but sleep group did significantly moderate a relationship between overnight change in recall SD and pre-sleep recall percent, $F(1, 32) = 5.21$, $p = .029$, $B = 1.32$, $se =$

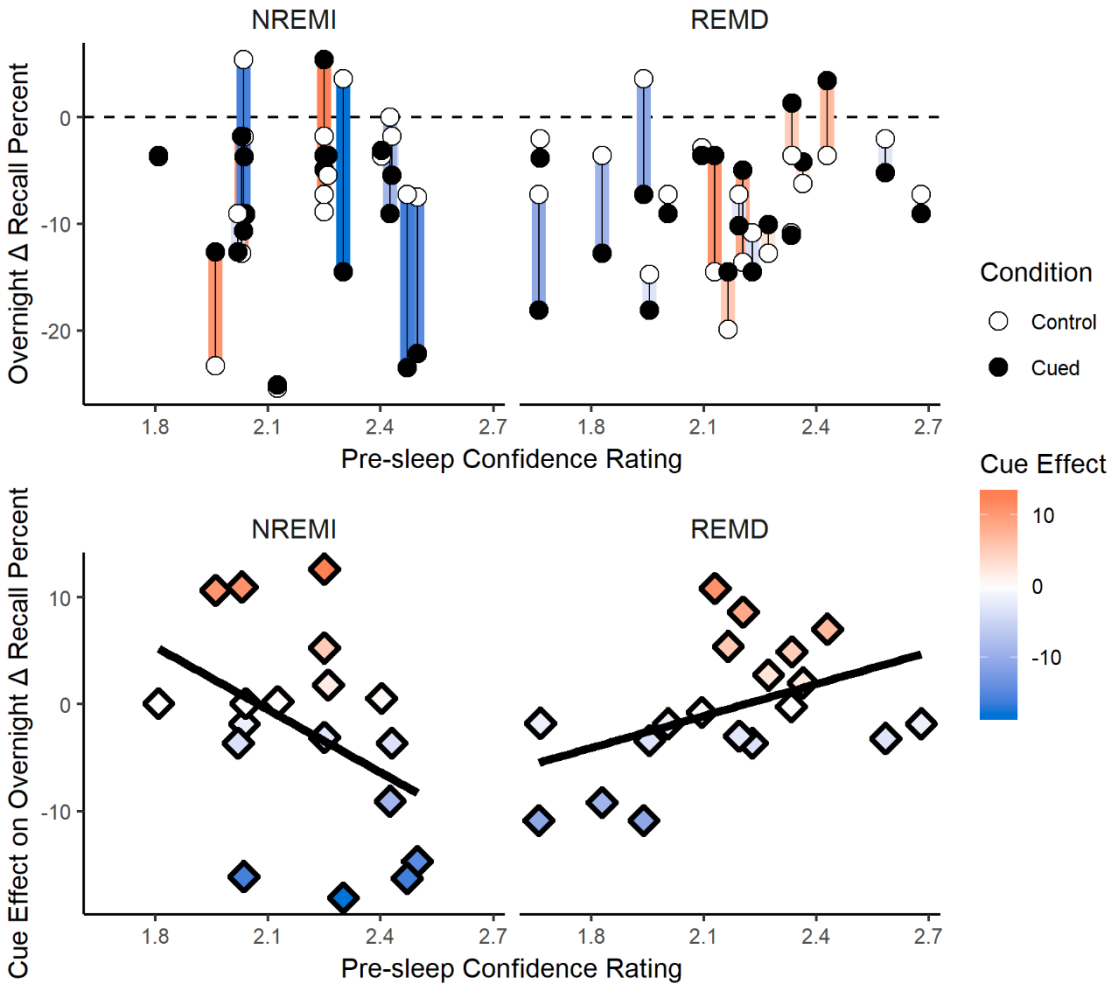

Figure 4.8. Association between pre-sleep recall confidence and the cueing effect (cued - control) on overnight change (post-sleep - pre-sleep) in recall percent separated by sleep group (NREMI vs. REMD). Confidence measured on a 1–3 scale. Vertical lines in top panels connect overnight retention scores for cued and control items within each participant. Cueing effect depicted in colour and on y-axis in bottom panels.

0.58, as calculated in an adjusted sample excluding an influential case from the NREMI group with an extreme overnight change sore (+1.67 IQR). Follow-up analyses indicated no clear association between pre-sleep recall percent and overnight change in recall SD in the NREMI group, B = -0.32, se = 0.41, t = -0.78; however, in the REMD group, a 10.57-

point (1 SD) increase in pre-sleep recall percent was associated with a 0.99º (*se* = 0.42) increase in overnight change score for recall SD, *t* = 2.38, indicating that recalling fewer item locations before sleep predicted less declines or gains in recall fidelity over the night for REMD participants (Figure 4.9).

There was no evidence that a cueing effect on overnight change in recall SD was moderated by pre-sleep recall SD, $F(1, 33) = 1.32$, $p = .259$, $B = 0.62$, $se = 0.54$, or by pre-sleep confidence, $F(1, 31) = 0.29$, $p = .594$, $B = 0.26$, $se = 0.49$. A model to test pre-sleep recall percent as moderator could not be determined before excluding too many cases, but there was no evidence of such moderation in the unadjusted sample or in an adjusted sample with four influential cases excluded. There was also no evidence that sleep group moderated interactions of condition and pre-sleep performance for pre-sleep recall percent, $F(1, 39) = 0.35$, $p = .560$, $B = 0.91$, $se = 1.53$ (with three influential cases

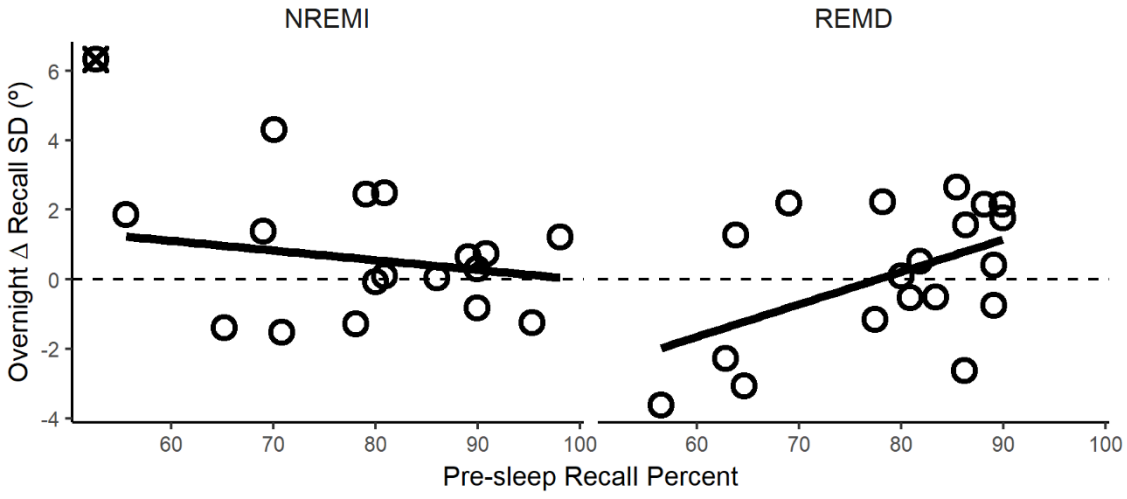

Figure 4.9. Overnight change in recall SD predicted by pre-sleep recall percent separated by sleep group. The influential case excluded from linear models in figure and report is marked by the × symbol.

excluded), for pre-sleep recall SD, $F(1, 31) = 0.09$, $p = .765$, $B = 0.33$, $se = 1.09$, or for pre-sleep confidence, $F(1, 30) = 0.76$, $p = .390$, $B = -0.87$, $se = 1.00$ (with one influential cases excluded).

Sleep measures. Measures of the duration of total sleep and each sleep stage, the percent of total sleep time spent in each sleep stage, the number and density of arousals, and average EEG power within each sleep stage for each frequency band were examined

Table 4.5
Tests of sleep measures as predictors of overnight change in recall performance on their own for both recall percent and recall SD

	As Predictor of Recall Percent						As Predictor of Recall SD					
Measure	n	*F*	df_r	*p*	*B*	*se*	n	*F*	df_r	*p*	*B*	*se*
Sleep dur.	37	0.28	35	.603	0.51	0.97	36^{r1}	2.13	34	.154	-0.44	0.30
N1 dur.	37	1.50	35	.229	1.17	0.96	37	1.14	35	.292	0.36	0.34
N2 dur.	34^{r3}	0.09	32	.767	-0.24	0.80	36^{a1r1}	9.34	34	.004	-0.84	0.27
N3 dur.	35^{r2}	1.30	33	.263	-0.92	0.81	37	0.31	35	.582	0.19	0.34
R dur.	37	0.13	35	.717	0.36	0.98	37	0.66	35	.421	0.28	0.34
Arousals	37	2.97	35	.094	-1.62	0.94	36^{r1}	0.59	34	.450	0.23	0.31
Ind. delta	37	0.82	35	.372	-0.87	0.97	37	0.78	35	.384	0.30	0.34
Ind. theta	36^{r1}	0.00	34	.995	-0.01	0.95	36^{r1}	0.01	34	.910	-0.04	0.34
Ind. sigma	36^{r1}	0.02	34	.884	-0.13	0.87	37	0.96	35	.333	-0.33	0.34
N2 delta	35^{r2}	1.50	31	.230	-0.90	0.73	35^{r2}	0.77	31	.386	-0.26	0.30
N2 theta							36^{r1}	0.38	32	.543	-0.24	0.39
N2 alpha	35^{r2}	0.76	31	.391	-0.86	0.99	36^{r1}	0.23	32	.635	-0.24	0.50
N2 sigma	35^{r2}	0.93	31	.343	-0.93	0.97						
N2 beta	34^{r3}	0.43	30	.518	0.72	1.10	36^{r1}	0.47	32	.497	-0.28	0.41
N3 delta	35^{r2}	1.42	31	.242	1.00	0.84	37	0.81	33	.373	-0.23	0.26
N3 theta	35^{r2}	0.43	31	.519	0.64	0.98	37	0.00	33	.946	-0.02	0.30
N3 alpha	35^{r2}	0.00	31	.983	0.02	0.91	37	0.38	33	.541	-0.21	0.34
N3 sigma	35^{r2}	2.32	31	.138	1.40	0.92	37	0.00	33	.951	0.02	0.38
N3 beta							36^{r1}	0.00	32	.988	0.01	0.48
R delta												
R theta												
R alpha	17r1	0.15	13	.707	-0.99	2.58	18	0.31	14	.585	0.35	0.64
R sigma												

Note. Dur. = duration. Ind. = induced. For a given test, a number after superscript "a" indicates the number of influential cases that had predictor values winsorized, and a number after superscript "r" indicates the number of influential cases removed from the sample. Empty cells are shown for tests for which a model without influential cases could not be fit.

as potential predictors of overnight change in performance.

Recall percent. A selection of the *F*-tests used to examine whether sleep measures were predictors of overnight change in recall percent on their own are reported in Table 4.5. A selection of the *F*-tests used to examine whether sleep measures were predictors of overnight change in recall percent in interactions with cueing condition and sleep group is reported in Table 4.6.

N3 duration was not a significant predictor of overnight change in recall percent. R duration was not a significant predictor on its own, but the R duration × group interaction was significant, $F(1, 33) = 5.41$, $p = .026$, $B = 27.72$, $se = 11.92$. Within the NREMI group, a 23.92-min (1 SD) increase in R duration had little association with overnight retention of recall percent, $B = 0.67$, $se = 1.64$, $t = 0.41$, but, within the REMD group, a 4.09-min (1 SD) increase in R duration was associated with an estimated 3.20-point ($se = 0.96$) increase in overnight retention of recall percent, $t = 3.33$ (Figure 4.10).

There was no evidence that R duration moderated an effect of cueing either across both groups or in interaction with sleep group.

As in the previous studies, there was evidence of N3 delta power as a predictor of overnight change in recall percent. The N3 duration × N3 delta power interaction was significant at channel Pz, $F(1, 30) = 6.88$, $p = .014$, $B = 2.21$, $se = 0.84$, channel P4, $F(1, 31) = 5.20$, $p = .030$, $B = 2.02$, $se = 0.89$, channel Oz, $F(1, 31) = 4.28$, $p = .047$, $B = 1.89$, $se = 0.91$, and channel PO8, $F(1, 31) = 7.67$, $p = .009$, $B = 2.55$, $se = 2.55$. This interaction was marginally significant at P3 and PO7 but had a *p* value greater than .164 at all other channels and the full scalp measure, pointing to a localization of effect to the same electrodes as the previous studies. These results came from adjusted samples excluding one influential case with an extreme overnight change in recall percent (-1.94

Table 4.6

Tests of sleep measures as predictors of overnight change in recall percent in interactions with condition and group

	Measure × Condition						Measure × Group						Measure × Condition × Group				
Measure	n	*F*	df_r	*p*	*B*	*se*	n	*F*	df_r	*p*	*B*	*se*	n	*F*	df_r	*p*	*B*	*se*
Sleep dur.	37	0.99	35	.327	1.29	1.30	37	1.85	33	.183	-2.74	2.01	37	0.54	33	.466	-2.00	2.71
N1 dur.	37	1.13	35	.295	1.38	1.30	37	0.91	33	.346	1.98	2.07	37	0.78	33	.384	-2.46	2.79
N2 dur.	36[r1]	1.12	34	.298	1.36	1.29	35[a1r2]	4.61	31	.040	-4.13	1.92	36[a1r1]	0.13	32	.722	1.27	3.54
N3 dur.	36[r1]	0.04	34	.843	0.27	1.35	35[r2]	0.00	31	.957	-0.10	1.82	36[r1]	0.13	32	.725	-1.06	3.00
R dur.	37	0.06	35	.805	-0.33	1.32	37	5.41	33	.026	27.72	11.92	36[r1]	0.37	32	.550	9.60	15.88
Arousals	37	0.00	35	.968	-0.05	1.32	37	1.13	33	.296	2.09	1.97	37	0.46	33	.503	1.88	2.77
Ind. delta	37	2.03	35	.163	1.83	1.28	37	2.53	33	.121	3.35	2.11	37	0.04	33	.849	-0.55	2.86
Ind. theta	37	1.59	35	.215	1.63	1.29	37	5.00	33	.032	4.55	2.04	37	0.18	33	.678	-1.20	2.85
Ind. sigma	36[r1]	0.58	34	.450	1.03	1.34	35[r2]	11.53	31	.002	5.60	1.65	37	0.06	33	.806	0.66	2.68
N2 delta	36[r1]	0.08	32	.778	-0.38	1.33	35[r2]	3.31	27	.080	-8.80	4.84	35[r2]	0.01	27	.929	-0.77	8.63
N2 theta	34[r3]	1.83	30	.186	1.74	1.29	34[r3]	1.16	26	.291	-3.50	3.25	35[r2]	0.66	27	.423	3.84	4.72
N2 alpha	36[r1]	0.41	32	.527	1.03	1.61	35[r2]	0.39	27	.537	-1.99	3.18						
N2 sigma	37	0.00	33	.999	0.00	1.67	35[r2]	6.02	27	.021	-5.45	2.22	35[r2]	0.38	27	.541	-2.74	4.42
N2 beta	34[r3]	0.16	30	.691	0.76	1.89	35[r2]	2.88	27	.101	-5.38	3.17	34[r3]	0.20	26	.658	5.33	11.89
N3 delta	35[r2]	0.81	31	.375	1.24	1.37												
N3 theta	35[r2]	1.36	31	.252	1.62	1.39												
N3 alpha	37	0.96	33	.334	1.19	1.21							34[r3]	2.02	26	.167	-8.00	5.63
N3 sigma	37	3.08	33	.089	2.37	1.35	35[r2]	20.45	27	.000	8.40	1.86	35[r2]	0.60	27	.446	-2.63	3.40
N3 beta	35[r2]	1.75	31	.196	2.67	2.02	34[r3]	15.97	26	.000	10.23	2.56						
R theta																		
R alpha	18	1.94	14	.185	3.63	2.60												
R sigma	18	2.81	14	.116	4.08	2.43												
R delta																		
R beta	17[r1]	4.81	13	.047	4.26	1.94												

Note. For a given test, a number after superscript "a" indicates the number of influential cases that had predictor values winsorized, and a number after superscript "r" indicates the number of influential cases removed from the sample. Empty cells are shown for tests for which a model without influential cases could not be fit.

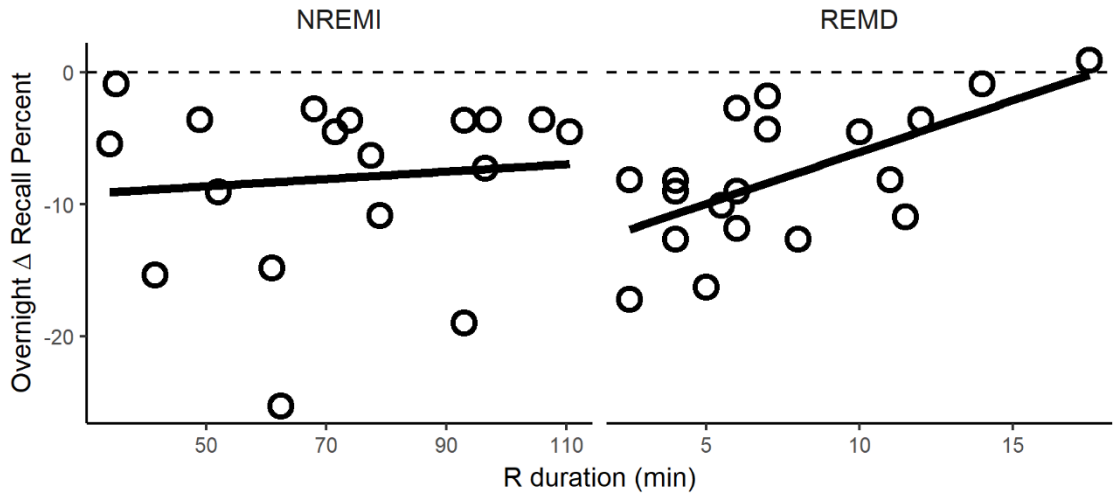

Figure 4.10. Overnight change in recall percent predicted R duration separated by sleep group. Due to the large difference in R duration between groups, the scale of the x-axis varies by panel.

IQR) and a second influential case with extremely low N3 delta power at several sites (full scalp: -1.95 IQR) and very little N3 sleep (3.5 min). A third influential case without extreme scores was excluded for the test at P3 and Pz. Further investigation was performed on a composite measure of average N3 delta power over all parietal and occipital channels. The N3 duration × N3 delta power interaction was significant with this composite measure in an adjusted sample excluding the same three influential cases, $F(1, 30) = 8.80$, $p = .006$, $B = 2.53$, $se = 0.85$ (Figure 4.11). N3 duration was not associated with overnight change in recall percent at relatively low N3 delta power, $B = 0.26$, $se = 1.08$, $t = 0.24$, but, at relatively high N3 delta power, an increase of 21.19 min (1 SD) of N3 sleep was associated with an estimated 3.09-point ($se = 1.18$) increase in the overnight retention of recall percent, $t = 2.61$.

There was also evidence that a relationship between N3 delta power and overnight

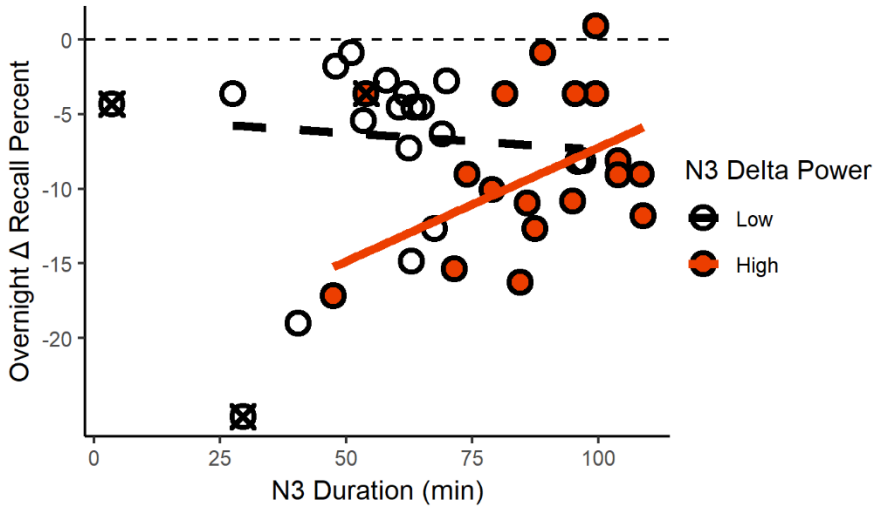

Figure 4.11. Overnight change in recall percent (post-sleep - pre-sleep) predicted by N3 duration on a median split of average delta (1–3.5 Hz) electroencephalographic power in stage N3 epochs over channels P3, Pz, P4, PO7, Oz, and PO8. Influential cases excluded from linear models in figure and report are marked by the × symbol.

retention in recall percent was moderated by sleep group. The N3 duration × N3 delta power × group interaction was significant at channel PO7, $F(1, 27) = 13.87$, $p = .001$, $B = 8.08$, $se = 2.17$, in an adjusted sample excluding one NREMI participant without extreme values on these measures and the REMD participant with only 3.5 min of N3 sleep. Regression models without influential cases could not be determined without excluding too many participants for all frontal channels, Cz, all parietal channels, Oz, PO8, and the full scalp measure. At PO7, the interaction was unexpectedly characterized by greater overnight retention of recall percent associating with greater N3 duration among NREMI participants with low delta power, $B = 6.86$, $se = 2.00$, $t = 3.43$. However, this unexpected interaction appeared to be driven by the two NREMI participants with the greatest overnight decline in recall percent (-27.98 and -19.00) also having low N3 delta power at

PO7 and low N3 duration. Repeating this analysis without the case with the most extreme decline (-1.95 IQR) resulted in a nonsignificant N3 duration × N3 delta power × group interaction, $F(1, 25) = 0.28$, $p = .603$, $B = 2.64$, $se = 5.00$, at PO7.

There was no evidence that a cueing effect was moderated by N3 delta power and no evidence that a N3 duration × N3 delta power × condition interaction was moderated by sleep group.

N2 and N3 sigma power were examined as predictors of overnight retention in recall percent. There was no evidence of either the N2 duration × N2 sigma power interaction or N3 duration × N3 sigma power interaction predicting overnight change in recall percent on their own, in interaction with condition, or in interaction with both condition and group. There was, however, evidence of sleep group moderating associations between sigma power and overnight change in all-item recall percent for both stage N2 sleep and stage N3 sleep.

The N2 duration × N2 sigma power × group interaction was significant for all frontal and central channels, P4, and the full scalp measure, all in an adjusted sample excluding one influential case from the NREMI group with an extreme decline in recall percent over the night (-1.95 IQR) and a second influential case from the NREMI group with extremely low N2 duration (86.5 min, -2.07 IQR). The effect was largest by *B* value at channel F3, $F(1, 27) = 9.72$, $p = .004$, $B = -7.49$, $se = 2.40$, channel F4$(1, 27) = 7.79$, $p = .010$, $B = -7.09$, $se = 2.54$, and channel C3, $F(1, 27) = 8.78$, $p = .006$, $B = -6.96$, $se = 2.35$. The N3 duration × N3 sigma power × group interaction was significant at channels F4, Fz, C4, and PO7, and the full scalp measure, all in adjusted samples excluding two (PO7 and full scalp) or three (F4, Fz, and C4) influential cases, one of which had extremely high N3 sigma power at F4 (+1.55 IQR). The effect was largest by *B* value at

channel C4, $F(1, 26) = 21.08$, $p < .001$, $B = 11.82$, $se = 2.58$. Though there was some indication of a similar interaction at other channels, models without influential cases could not be determined without excluding too many cases. Notably, the interactions for N3 sigma power were opposite in sign to those for N2 sigma power.

After exploration of these opposing and complex interactions, effort was made to simplify the pattern of results. Given little rationale for a functional distinction between sigma activity in N2 sleep and sigma activity in N3 sleep, and the high correlation between N2 and N3 sigma power (e.g., $r = .92$ at F3), a composite measure averaging sigma power over both stage N2 and stage N3 epochs (N2/3) and channels F3, Fz, F4, C3, Cz, and C4 was created. The inversion of the interaction between N2 and N3 sleep was likely linked to negative correlations between N2 and N3 sleep duration in both the NREMI group ($r = -.34$) and the REMD group ($r = -.20$). Thus, to simplify, a N3:N2 ratio of sleep stage durations was calculated to capture the depth of NREM sleep (a higher ratio reflects more N3 sleep relative to N2 sleep and overall deeper NREM sleep). The N3:N2 ratio × N2/3 sigma power × group interaction was significant in an adjusted sample excluding three influential cases, $F(1, 26) = 8.57$, $p = .007$, $B = 5.18$, $se = 1.77$. The influential cases included two NREMI participants, one with an extreme overnight decline in recall percent (-1.95 IQR) and one with an extremely high N3:N2 ratio (+2.31 IQR), and one REMD participant without extreme scores on these measures. There was a strong N3:N2 ratio × N2/3 sigma power interaction in the REMD group, $B = 4.65$, $se = 0.89$, $t = 5.21$, but no clear interaction in NREMI group, $B = -0.61$, $se = 1.56$, $t = -0.39$ (Figure 4.12). A 1 SD increase in average N2/3 sigma power was not strongly associated with overnight change in recall percent for the NREMI group at mean N3:N2 ratio (≈2:5),

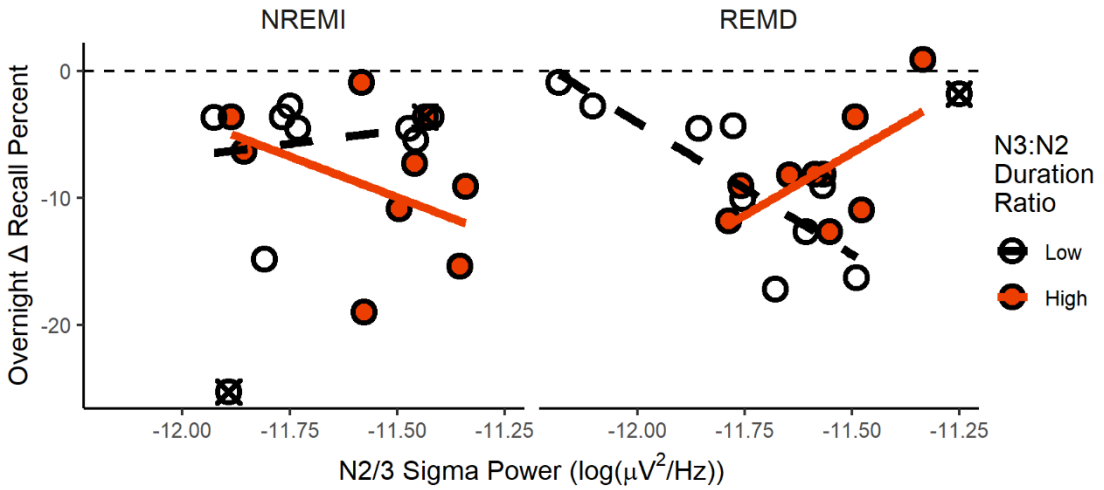

Figure 4.12. Overnight change in recall percent (post-sleep - pre-sleep) predicted by the average sigma (12–15.5 Hz) electroencephalographic power in stage N2 and N3 epochs over channels F3, Fz, F4, C3, Cz, and C4 on a median split of the ratio of N3 duration to N2 duration and separated by sleep group. Influential cases excluded from linear models in figure and report are marked by the × symbol.

B = -1.14, *se* = 1.56, *t* = -0.73, but was associated with a 4.09-point (*se* = 0.92) decrease in retention of recall percent for REMD participants with a low N3:N2 ratio (1:4), *t* = -4.47, and a 3.44-point (*se* = 1.39) increase in retention of recall percent for REMD participants with a high N3:N2 ratio (≈1:2), t = 2.48. This pattern indicates that greater sigma power in stages N2 and N3 sleep predicted less retention of locations for REMD participants with relatively light NREM sleep and predicted more retention of locations for REMD participants with relatively deep NREM sleep while having little predictive value within the NREMI group.

None of total sleep duration, N1 duration, N2 duration, arousal count, or arousal density were found to be significant predictors of overnight change in recall percent.

Results for measures of the percent of total sleep time spent in each stage matched those of the sleep stage duration measures.

Average delta, theta, alpha, and beta power of N2 sleep and average theta, alpha, and beta power of N3 sleep were all examined as potential predictors of overnight change in recall percent. For none of these measures was there evidence of the sleep stage duration × EEG power interaction predicting overnight change in recall percent on its own. There was also no evidence of interactions with cueing condition in either stage duration × EEG power × condition interactions or stage duration × EEG power × condition × group interactions. There were some significant stage duration × EEG power × group interactions, specifically for average theta, alpha, and beta power in stage N3 sleep.

Both the N3 duration × N3 theta power × group interactions and the N3 duration × N3 alpha power × group interactions appeared unstable and dependent on two NREMD participants: one with relatively low N3 duration (29.5 min) and an extreme overnight decline in recall percent (25.27 points; -1.95 IQR) and another less extreme case with 40.5 min of stage N3 sleep and an overnight decline of 19.00 points. The interactions were unstable in that regression models could not be fit for 50% of the measures of average EEG power and multiple and varied influential cases were identified across different channel measures. For models in which these two notable NREMI participants remained and other influential cases were excluded, the interactions in question were typically significant. When the typical analysis strategy was completed without the most extreme of these cases included, there was no evidence of significant N3 duration × EEG power × group interactions for either theta or alpha power.

There was evidence of a N3 duration × N3 beta power × group interaction

predicting overnight change in recall percent at nearly every EEG channel and the full scalp measure. This interaction was significant in an unadjusted sample without influential cases at channel Oz, $F(1, 29) = 13.67$, $p = .001$, $B = 7.71$, $se = 2.08$. This interaction was also significant at channel PO7 and with the full scalp measure in an adjusted sample excluding two non-extreme influential cases from the REMD group and one non-extreme influential case from the NREMI group and at channels Cz, P3, and Pz with only the two REMD participants excluded. Models without influential cases could not be determined for F3, Fz, F4, C3, C4, and P4. Follow-up analyses were conducted using the Oz measure because it used the full sample. There were N3 duration × N3 beta power interactions in both the REMD group, $B = 3.14$, $se = 1.49$, $t = 2.10$, and the NREMI group, $B = -4.20$, $se = 1.37$, $t = -3.05$ (Figure 4.13). Simple slope estimation of the effect of a 1 SD increase in N3 duration on overnight change in recall percent indicated a small 2.34-point ($se = 1.26$) increase for NREMI participants with high N3 beta power, $t = 1.86$, a large 8.19-point ($se = 2.30$) increase for NREMI participants with low N3 beta power, $t = 3.57$, a negligible 0.94-point ($se = 1.45$) increase for REMD participants with high N3 beta power, $t = 0.65$, and a small 2.59-point ($se = 1.44$) decrease for REMD participants with low N3 beta power, $t = -1.79$. While more stable than the corresponding interactions for theta and alpha power, it should be noted that the large effect of N3 duration estimated for NREMI participants with low N3 beta power is driven largely by the previously noted participant with an extreme overnight decline in recall percent. When analyses were repeated without this participant excluded, the overall N3 duration × N3 beta power × group interaction was not significant for the few measures (PO7, Oz, and PO8) where a model without influential cases could be determined.

Average delta, theta, alpha, sigma, and beta EEG power measures of R sleep were

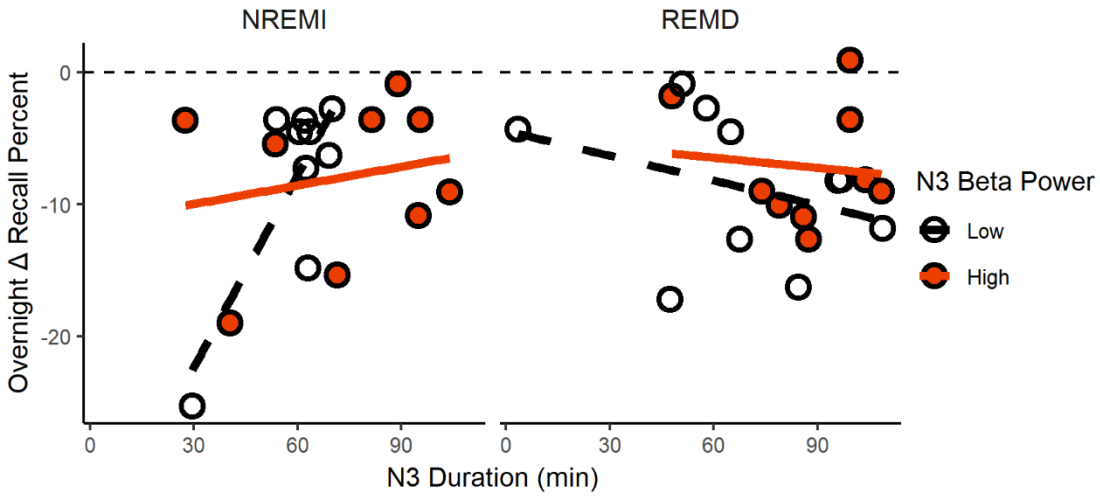

Figure 4.13. Overnight change in recall percent (post-sleep - pre-sleep) predicted N3 duration on a median split of average beta (8–11.5 Hz) electroencephalographic power in stage N3 epochs at channel Oz and separated by sleep group.

examined as potential predictors of overnight change in recall percent within the NREMI group. No R duration × EEG power interactions or R duration × EEG power × condition interactions met criteria for statistical significance in predicting overnight change in recall percent for any frequency band. In part due to the small sample size within the NREMI group, regression models without influential cases could not be determined without excluding too many cases for many of these tests. A model without influential cases was obtained for only 32% of tests of EEG measures as a predictor, for only 46% of tests of these EEG measures as moderators of condition, and only 5% of tests of R delta and R theta power.

Recall SD. A selection of the *F*-tests used to examine whether sleep measures were predictors of overnight change in recall SD on their own are reported in Table 4.5. A selection of the *F*-tests used to examine whether sleep measures were predictors of

Table 4.7
Tests of sleep measures as predictors of overnight change in recall SD in interactions with condition and group

	Measure × Condition						Measure × Group						Measure × Condition × Group					
Measure	n	*F*	df_r	*p*	*B*	*se*	n	*F*	df_r	*p*	*B*	*se*	n	*F*	df_r	*p*	*B*	*se*
Sleep dur.	35	1.11	33	.300	0.53	0.50	35[r2]	2.95	31	.096	1.03	0.60	34[r1]	3.64	30	.066	-1.90	0.99
N1 dur.	35	0.27	30	.604	0.25	0.48	37	0.55	33	.465	0.53	0.72	35	1.05	28	.315	-1.08	1.06
N2 dur.	35[a1]	2.44	34	.128	0.85	0.54							32[r3]	0.62	28	.437	-1.10	1.39
N3 dur.	35	0.04	31	.844	-0.10	0.48	37	0.09	33	.772	-0.21	0.73	34[r1]	1.41	28	.245	-1.09	0.91
R dur.	35	0.26	33	.611	-0.26	0.50	37	5.21	33	.029	9.52	4.17	35	1.05	28	.314	6.34	6.18
Arousals	35	0.16	31	.693	0.20	0.49	37	0.35	33	.559	0.42	0.72	35	0.13	29	.718	-0.38	1.04
Ind. delta	34[r1]	2.95	30	.096	0.69	0.40	37	0.79	33	.381	0.67	0.75	34[r1]	1.57	27	.221	-1.07	0.85
Ind. theta	32[r3]	0.27	29	.605	0.21	0.40	36[r1]	4.80	32	.036	1.48	0.67	34[r1]	3.55	27	.070	-1.62	0.86
Ind. sigma	35	0.02	31	.897	0.06	0.48	37	5.12	33	.030	1.45	0.64	35	0.80	29	.378	-0.85	0.95
N2 delta	32[r3]	1.53	29	.225	-0.73	0.58												
N2 theta	33[r2]	0.11	26	.744	0.19	0.56	34[r3]	5.78	26	.024	-1.94	0.81						
N2 alpha	34[r1]	0.06	28	.801	0.20	0.78							34[r1]	1.23	27	.277	3.23	2.89
N2 sigma																		
N2 beta																		
N3 delta	32[r3]	8.01	25	.009	-1.00	0.35	35[r2]	1.34	27	.257	0.82	0.71						
N3 theta	33[r2]	9.74	27	.004	-1.53	0.49												
N3 alpha	35	2.67	29	.113	-0.77	0.47	35[r2]	2.79	27	.107	1.66	0.99						
N3 sigma	34[r1]	0.01	28	.919	0.05	0.49	35[r2]	8.15	27	.008	2.25	0.79						
N3 beta	33[r2]	1.94	26	.175	0.70	0.50	35[r2]	9.94	27	.004	3.41	1.08						
R theta																		
R alpha	17[r1]	0.04	11	.845	0.21	1.07												
R sigma	18	0.20	12	.663	0.36	0.80												
R delta	17[r1]	2.06	11	.179	-2.12	1.47												
R beta	16[r2]	1.30	11	.280	0.89	0.78												

Note. For a given test, a number after superscript "a" indicates the number of influential cases that had predictor values winsorized, and a number after superscript "r" indicates the number of influential cases removed from the sample. Empty cells are shown for tests for which a model without influential cases could not be fit.

overnight change in recall SD in interactions with cueing condition and sleep group is reported in Table 4.7.

N3 duration was not a significant predictor of overnight recall percent. There was also no evidence that R duration predicted overnight change in recall SD on its own or moderated an effect of cueing either across both groups or in interaction with sleep group. However, as was the case overnight change in recall percent, the R duration × group interaction was significant, $F(1, 33) = 5.21$, $p = .029$, $B = 9.52$, $se = 4.17$. In separate models for each group, R duration had little association with overnight change in recall SD in the NREMI group, $B = -0.35$, $se = 0.51$, $t = -0.69$, but a 4.09-min (1 SD) increase in R duration in the REMD group was associated with an estimated 1.00º ($se = 0.41$) increase in overnight change score for recall SD, indicating a greater decline in fidelity of recall over the night, $t = 2.41$ (Figure 4.14).

The N3 duration × N3 delta power interaction was not significant in predicting overnight changes in recall SD on its own or in any interaction with sleep group, but there was evidence that it moderated an effect of condition. The N3 duration × N3 delta power × condition interaction was significant for the full scalp measure and at channels C3, C4, P3, Pz, P4, PO7, and Oz, all in various adjusted samples excluding or winsorizing one-to-three influential cases. The effect was of similar size in each of these models (B values from -1.29 to -1.00) except C3. For consistency with previous analyses, a composite channel averaging delta power over parietal and occipital channels was used for subsequent analyses. With this measure, the interaction was significant in the adjusted sample with only one (the participant with only 3.5 min of N3 sleep) influential case excluded, $F(1, 28) = 7.57$, $p = .010$, $B = -1.18$, $se = 0.43$ (Figure 4.15). For low N3 delta power, simple slope estimation indicated no clear N3 duration × condition interaction, B

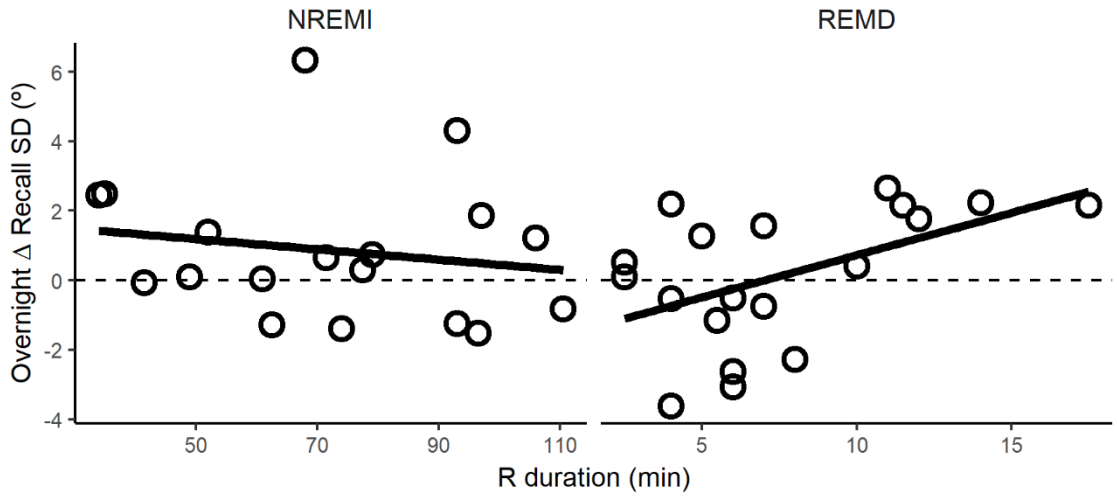

Figure 4.14. Overnight change in recall SD predicted by R duration separated by sleep group. Due to the large difference in R duration between groups, the scale of the x-axis varies by panel.

= 0.51, *se* = 0.67, *t* = 0.76, and no clear effect of cueing on overnight change in recall SD at mean N3 duration, *B* = 0.62, *se* = 0.72, *t* = 0.86. For high N3 delta power, the interaction was more notable, *B* = -1.16, *se* = 0.74, *t* = -1.55: cueing relative to control was estimated to amplify overnight increases in recall SD by 1.85º (*se* = 1.09) at low N3 duration, *t* = 1.70, while having no clear effect at high N3 duration, *B* = -0.03, *se* = 0.63, *t* = -0.05. Thus, an effect of cueing to decrease fidelity of recall was apparent specifically for those with low N3 duration but high N3 delta power.

For sigma power in N2 sleep and N3 sleep, there was a similar pattern for predicting overnight change in recall SD as there was for predicting overnight change in recall percent. There was no indication that either N2 or N3 sigma power were significant predictors of overnight change in recall SD on their own or in any interaction with condition. The N3 duration × N3 sigma power × group interaction was significant for the

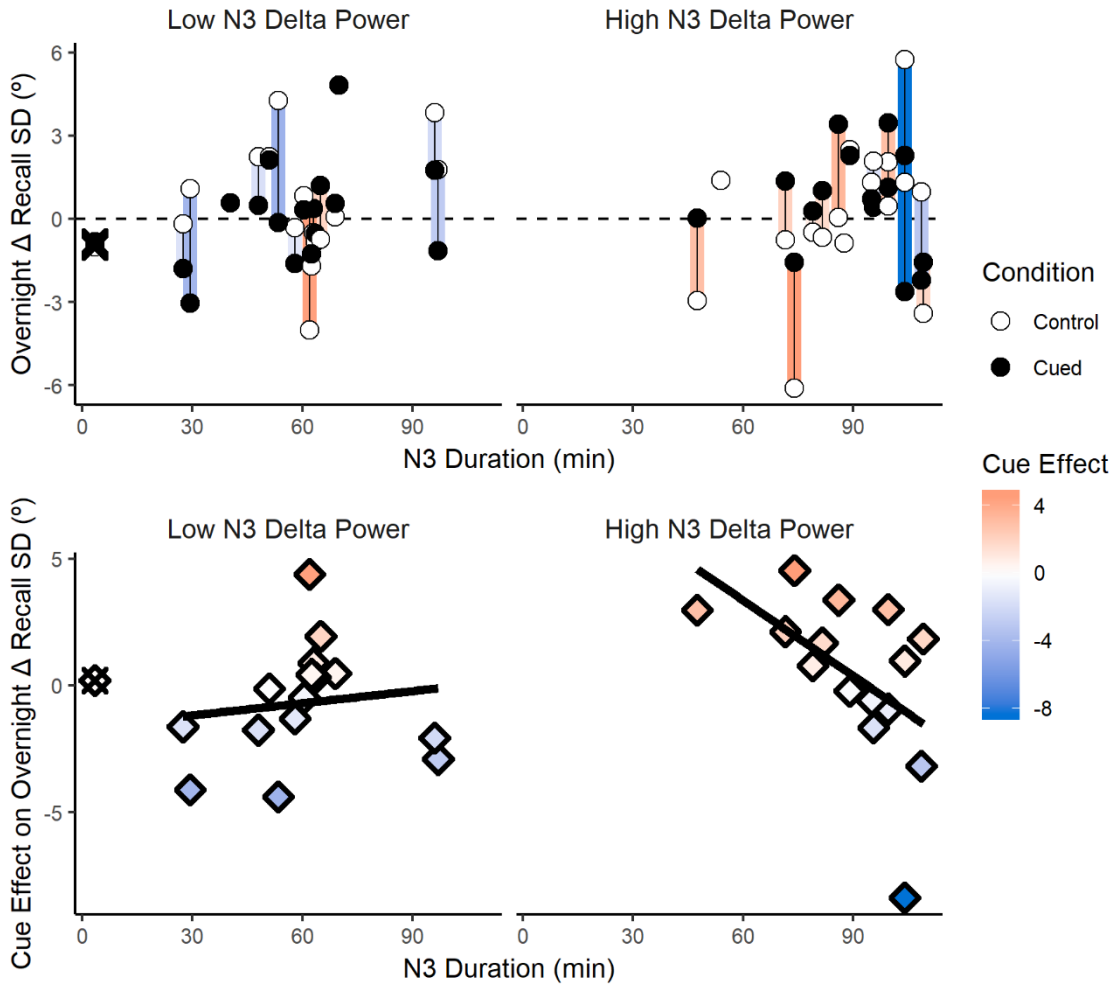

Figure 4.15. N3 duration and average delta (1–3.5 Hz) electroencephalographic power in stage N3 epochs moderating the effect of cueing (cued - control) on overnight change (post-sleep - pre-sleep) in recall SD. Relationship between N3 duration and cueing effect shown on a median split of delta power averaged over channels P3, Pz, P4, PO7, Oz, and PO8. Vertical lines in top panels connect overnight change scores for cued and control items within participants. Cueing effect depicted in colour and on y-axis in bottom panels. The influential case excluded from linear models in figure and report is marked by the × symbol.

full scalp measure and at channels F3, Fz, F4, C3, Cz, C4, P3, Pz, and P4, all in various

adjusted samples excluding one-to-three influential cases. The effect was of similar size in each of these models (*B* values from 1.96 to 2.60) and largest by *B* value at Cz, $F(1, 27) = 6.81$, $p = .015$, $B = 2.60$, $se = 1.00$, where an influential case with extremely low N3 duration (3.5 min, -1.58 IQR) and a separate non-extreme influential case were excluded. As was the case for overnight change in recall percent, there was a tendency for N2 duration × N2 sigma power × group interactions to be in the opposite direction (e.g., for F4 in the unadjusted sample: $F(1, 29) = 2.23$, $p = .146$, $B = -1.90$, $se = 1.27$), though models without influential cases could not be determined for these stage N2 interactions. For consistency with analyses for recall percent, further investigation used the ratio of N3 duration to N2 duration instead of either individual stage duration to capture the depth of NREM sleep and used the N2/3 sigma power measure of average sigma power over F3, Fz, F4, C3, Cz, and C4 in stage N2 and N3 epochs. The N3:N2 ratio × N2/3 sigma power × group interaction was significant in predicting overnight changes in recall SD in an adjusted sample excluding one non-extreme influential case from the REMD group, $F(1, 28) = 6.58$, $p = .016$, $B = 2.27$, $se = 0.89$ (Figure 4.16). In the NREMI group, there was little evidence of a N3:N2 ratio × N2/3 sigma power interaction, $B = -0.87$, $se = 0.64$, $t = -1.37$, and a 1 SD increase in average N2/3 sigma power was not associated with overnight change in recall SD at mean N3:N2 ratio (2:5), $B = 0.15$, $se = 0.61$, $t = 0.25$. In the REMD group, there was a N3:N2 ratio × N2/3 sigma power interaction, $B = 1.24$, $se = 0.54$, $t = 2.30$, characterized by a 1 SD increase in average N2/3 sigma power predicting a negligible -0.62º shift in recall SD at low N3:N2 ratio (1:4), $t = -1.12$, and a 1.39º ($se = 0.84$) shift in recall SD at high N3:N2 ratio (≈1:2), $t = 1.65$. Thus, there was tendency for reliable overnight declines in recall fidelity among REMD participants with both

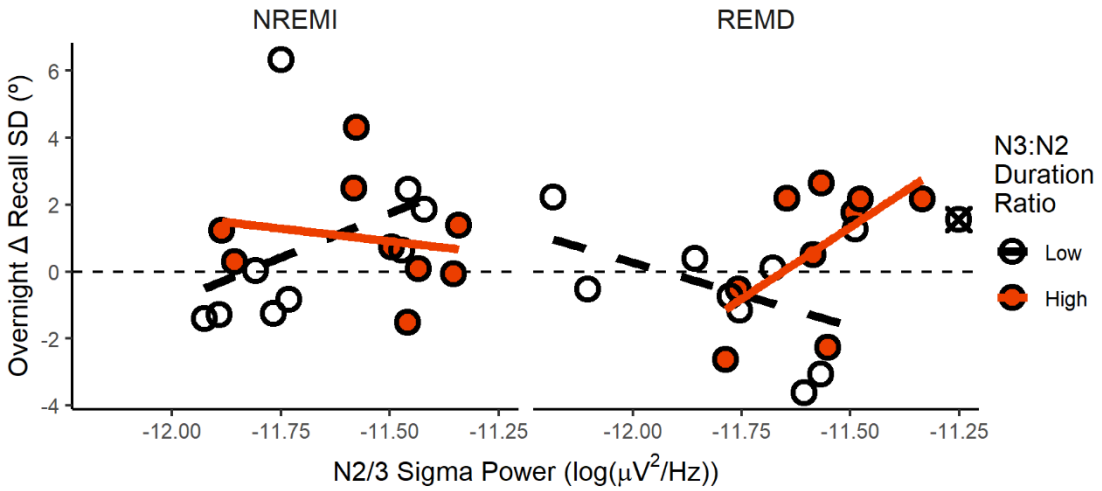

Figure 4.16. Overnight change in recall SD (post-sleep - pre-sleep) predicted by average sigma (12–15.5 Hz) electroencephalographic power in stage N2 and N3 epochs over channels F3, Fz, F4, C3, Cz, and C4 on a median split of the ratio of N3 duration to N2 duration and separated by sleep group. The influential case excluded from linear models in figure and report is marked by the × symbol.

relatively deep and sigma-rich NREM sleep.

The remaining sleep architecture measures, including total sleep duration, N1 duration, N2 duration, arousal count, and arousal density, were examined as predictors of overnight change in recall SD. Of these measures, there were no significant interactions with group or condition, and only N2 duration was a significant predictor of overnight change in recall SD. Results for measures of the percent of total sleep time spent in each stage matched those of the sleep stage duration measures.

N2 duration was a significant predictor in an adjusted sample with one influential case with an extreme overnight change score for recall SD (+1.67 IQR) excluded and another influential case with extremely low N2 duration (86.5 min, -2.07 IQR) winsorized

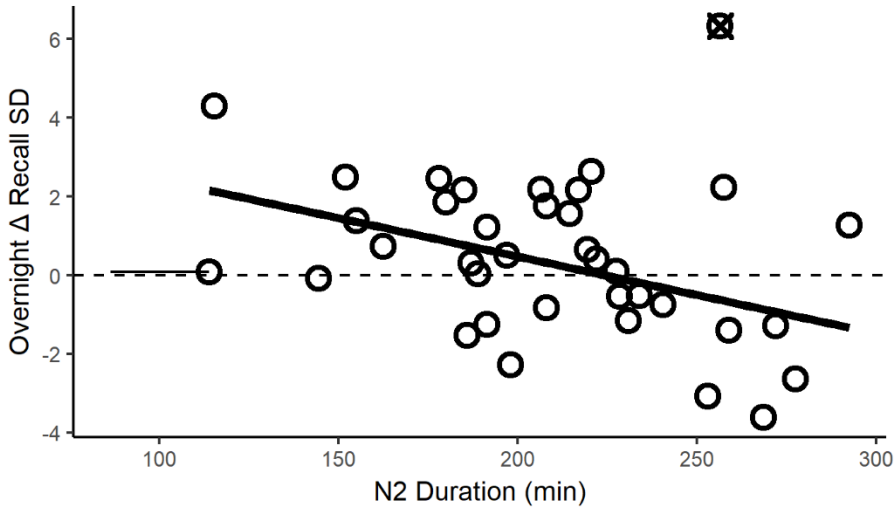

Figure 4.17. Overnight change in recall SD predicted by N2 duration. Solid horizonal line connects an influential case that has been winsorized to its original value. The influential case excluded from linear models in figure and report is marked by the × symbol.

on that measure, $F(1, 34) = 9.34$, $p = .004$; a 42.82 (1 SD) increase in N2 duration was associated with a 0.84º (*se* = 0.27) decrease in overnight change score for recall SD, indicating less declines or gains in recall fidelity with greater duration of stage N2 sleep (Figure 4.17).

Average delta, theta, alpha, and beta power of N2 sleep and theta, alpha, and beta power of N3 sleep were all examined as potential predictors of overnight change in recall SD. For none of these measures was there evidence of the sleep stage duration × EEG power interaction predicting overnight change in recall SD on its own. There was also no evidence of interactions with group. There was, however, evidence that EEG power measures from stages N2 and N3 and sleep stage durations moderated a cueing effect on overnight change in recall SD. The significant stage duration × EEG power × condition

interactions were only observed for N3 theta power and N2 beta power.

The N3 duration × N3 theta power × condition interaction was significant in predicting overnight change in recall SD at channels Fz, F4, C3, P3, and Pz, and for the full scalp measure. The participant with low N3 sleep duration (3.5 min) and extremely low N3 theta power at multiple sites (-1.50 IQR for full scalp) was influential and excluded for each of these models. An influential case influential case with extremely low N3 theta power at Fz (-1.50 IQR) was also excluded from the Fz and F4 models, and an additional non-extreme influential case was excluded from the Fz and full scalp models. The interaction was strongest by *B* value in the full scalp model, $F(1, 27) = 9.74$, $p = .004$, $B = -1.53$, $se = 0.49$. There was some evidence of a N3 duration × condition interaction at low N3 theta power, $B = 1.14$, $se = 0.67$, $t = 1.71$, but less so at high N3 theta power, $B = -1.11$, $se = 0.73$, $t = -1.52$ (Figure 4.18). Simple slope estimation indicated that, at low N3 theta power, cueing relative to control increased the overnight change in recall SD by 2.86º ($se = 1.24$) at high N3 duration, $t = 2.31$, while having a smaller increase of 0.90º ($se = 0.61$) at low N3 duration, $t = 1.48$. Thus, cueing appeared to contribute to declines in recall fidelity for those with low N3 theta power, particularly in the presence of high N3 duration. At high N3 theta power, there was no clear effect of cueing at low, mean, or high N3 duration (at mean: $B = 0.15$, $se = 0.73$, $t = 0.21$).

The N2 duration × N2 beta power × condition interaction was significant in predicting overnight change in recall SD at channels C3, P3, and Pz in adjusted samples excluding an influential case with extremely low N2 duration (86.5 min, -2.07 IQR) and a non-extreme influential case. The adjusted sample for Pz also had a third influential case with extremely low N2 beta power at multiple channels including Pz (-2.28 IQR) winsorized on this value, and the adjusted sample for C3 excluded this influential case.

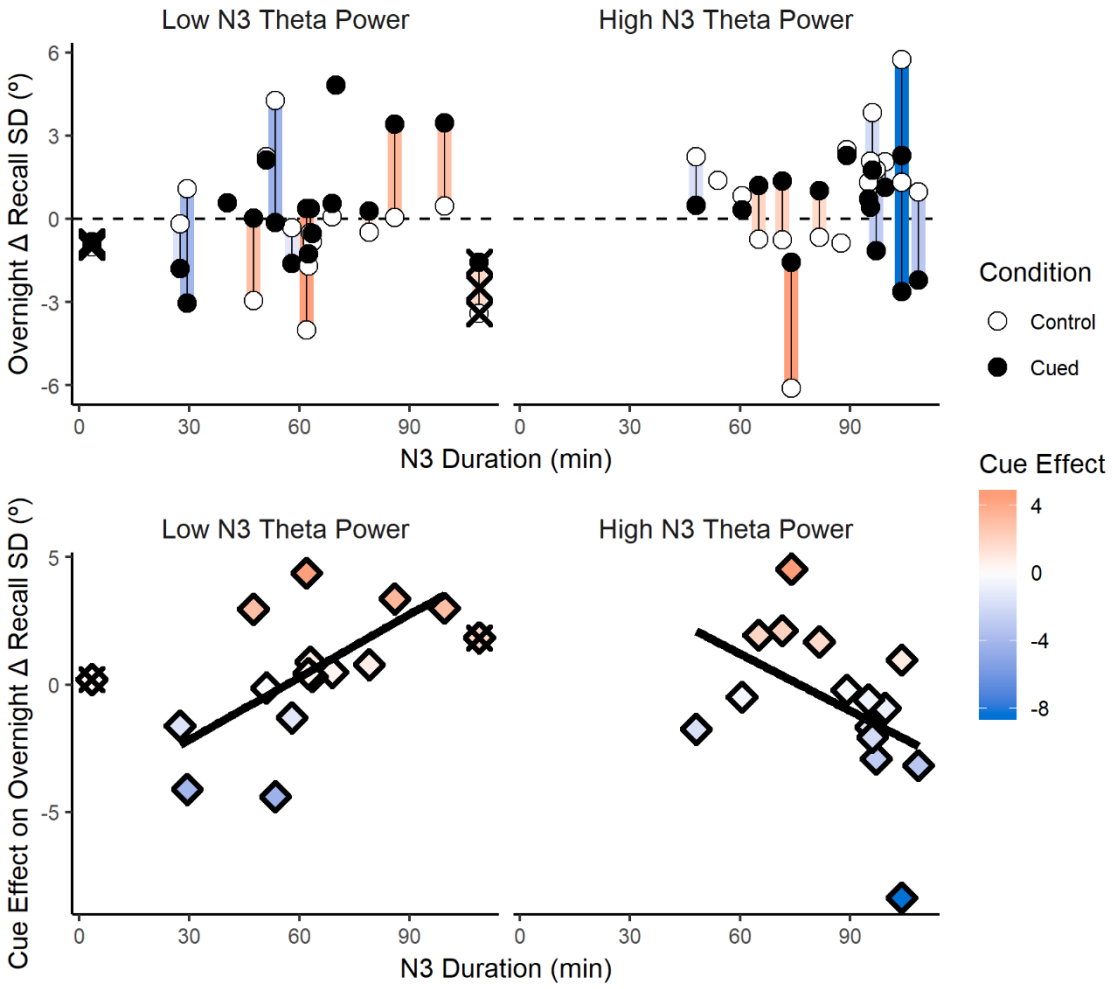

Figure 4.18. N3 duration and average theta (4–7.5 Hz) electroencephalographic power in stage N3 epochs moderating the effect of cueing (cued - control) on overnight change (post-sleep - pre-sleep) in recall SD. Relationship between N3 duration and cueing effect shown on a median split of N3 theta power. Vertical lines in top panels connect overnight retention scores for cued and control items within each participant. Cueing effect depicted in colour and on y-axis in bottom panels. The influential case excluded from linear models in figure and report is marked by the × symbol.

The interaction was strongest by *B* value at C3, $F(1, 26) = 13.48$, $p = .001$, $B = 1.97$, $se = 0.54$, but similar at all three channels. The interaction nearly met criteria for statistical

significance at sites Fz, Cz, C4, P4, Oz, and PO8 though models without influential cases could not be determined for Cz, C4, and P4. A composite channel averaging N2 beta power over all channels except F3 and F4 (where high beta power was often an artifact of sleeping on one's side) was computed, and the interaction was significant with this measure in an adjusted sample excluding the same three influential cases, $F(1, 26) = 9.34$, $p = .005$, $B = 1.56$, $se = 0.51$. There was no clear N2 duration × condition interaction estimated at low N2 beta power, $B = 0.86$, $se = 0.64$, $t = 1.34$, but a notable N2 duration × condition interaction at high N2 beta power, $B = 2.52$, $se = 0.71$, $t = 3.54$ (Figure 4.19). Simple slope estimation indicated that, at low N2 beta power, cueing had little effect on overnight change in recall SD at mean N2 duration, $B = 0.72$, $se = 0.54$, $t = 1.32$, but, at high N2 beta power, cueing relative to control was estimated to substantially increase recall SD by 2.15º ($se = 0.89$) at high N2 duration, $t = 2.43$, while showing some signs of decreasing recall SD by 0.90º ($se = 0.55$) at low N2 duration, $t = -1.62$. This pattern suggests that cueing reliably decreased recall fidelity for cued items among participants with large amounts of stage N2 sleep with relatively high beta power.

Average delta, theta, alpha, sigma, and beta EEG power measures of R sleep were examined as potential predictors of overnight change in recall SD within the NREMI group. No R duration × EEG power interactions or R duration × EEG power × condition interactions met criteria for statistical significance in predicting overnight change in recall SD for any frequency band.

Cue-induced power changes. The following tests examined whether the identified changes in delta, theta, and sigma EEG power induced by sound cues during the night predicted overnight change in memory performance. *F*-tests used to examine whether the maximum site measure for each band predicted either recall percent or recall

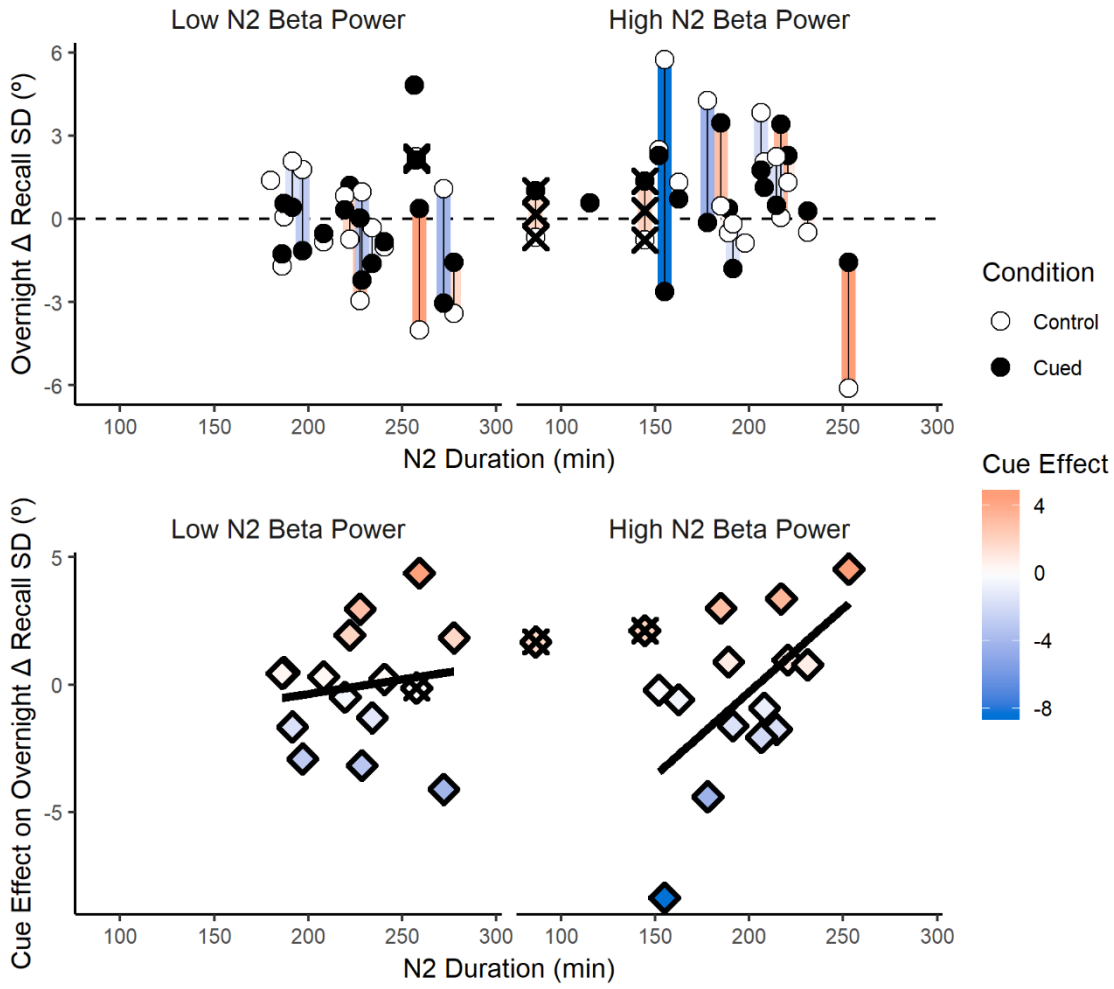

Figure 4.19. N2 duration and average beta (16–29.5 Hz) electroencephalographic power in stage N2 epochs moderating the effect of cueing (cued - control) on overnight change (post-sleep - pre-sleep) in recall SD. Relationship between N2 duration and cueing effect shown on a median split of N2 beta power averaged over channels Fz, C3, Cz, C4, P3, Pz, P4, PO7, Oz, and PO8. Vertical lines in top panels connect overnight retention scores for cued and control items within each participant. Cueing effect depicted in colour and on y-axis in bottom panels. Influential cases excluded from linear models in figure and report are marked by the × symbol.

SD on their own are reported in Table 4.5. *F*-tests used to examine these measures

predicted performance as interactions with cueing condition or sleep group are reported in Table 4.6 for recall percent and Table 4.7 for recall SD.

Recall percent. None of induced delta, theta, or sigma power were significant predictors on their own, nor did any of these measures significantly interact with condition to predict overnight change in recall percent. Induced theta and sigma power, but not induced delta power, did significantly interact with group to predict overnight change in recall percent.

The induced theta power × group interaction was significant at channel C4, $F(1, 33) = 4.95$, $p = .033$, $B = 4.31$, $se = 1.94$, for Cmax, $F(1, 33) = 5.00$, $p = .032$, $B = 4.55$, $se = 2.04$, and at channel C3 in an adjusted sample excluding a non-extreme influential case from the NREMI group, $F(1, 32) = 6.08$, $p = .019$, $B = 4.89$, $se = 1.98$. Further investigation indicated that slight and opposing associations in the NREMI and REMD groups (Figure 4.20a). For Cmax, it was estimated that a within-group 1 SD increase in induced theta power was associated with a 2.49-point ($se = 1.52$) decrease in overnight retention of recall percent for the NREMI group, $t = -1.63$, and a 1.92-point ($se = 1.14$) increase in overnight retention of recall percent for the REMD group, $t = 1.68$.

The induced sigma power × group interaction was significant at channel C3, $F(1, 31) = 7.76$, $p = .009$, $B = 4.59$, $se = 1.65$, at channel Cz, $F(1, 31) = 7.08$, $p = .012$, $B = 4.82$, $se = 1.81$, at channel C4, $F(1, 31) = 11.89$, $p = .002$, $B = 5.38$, $se = 1.59$, and for Cmax, $F(1, 31) = 11.53$, $p = .002$, $B = 5.60$, $se = 1.65$. These results came from adjusted samples with an influential case from the NREMI group with an extreme overnight change in recall percent (-1.95 IQR) excluded from all samples, a non-extreme influential case from the NREMI group excluded from samples for Cz and Cmax, and a non-extreme influential case from the REMD group excluded from samples for C3 and C4. Further

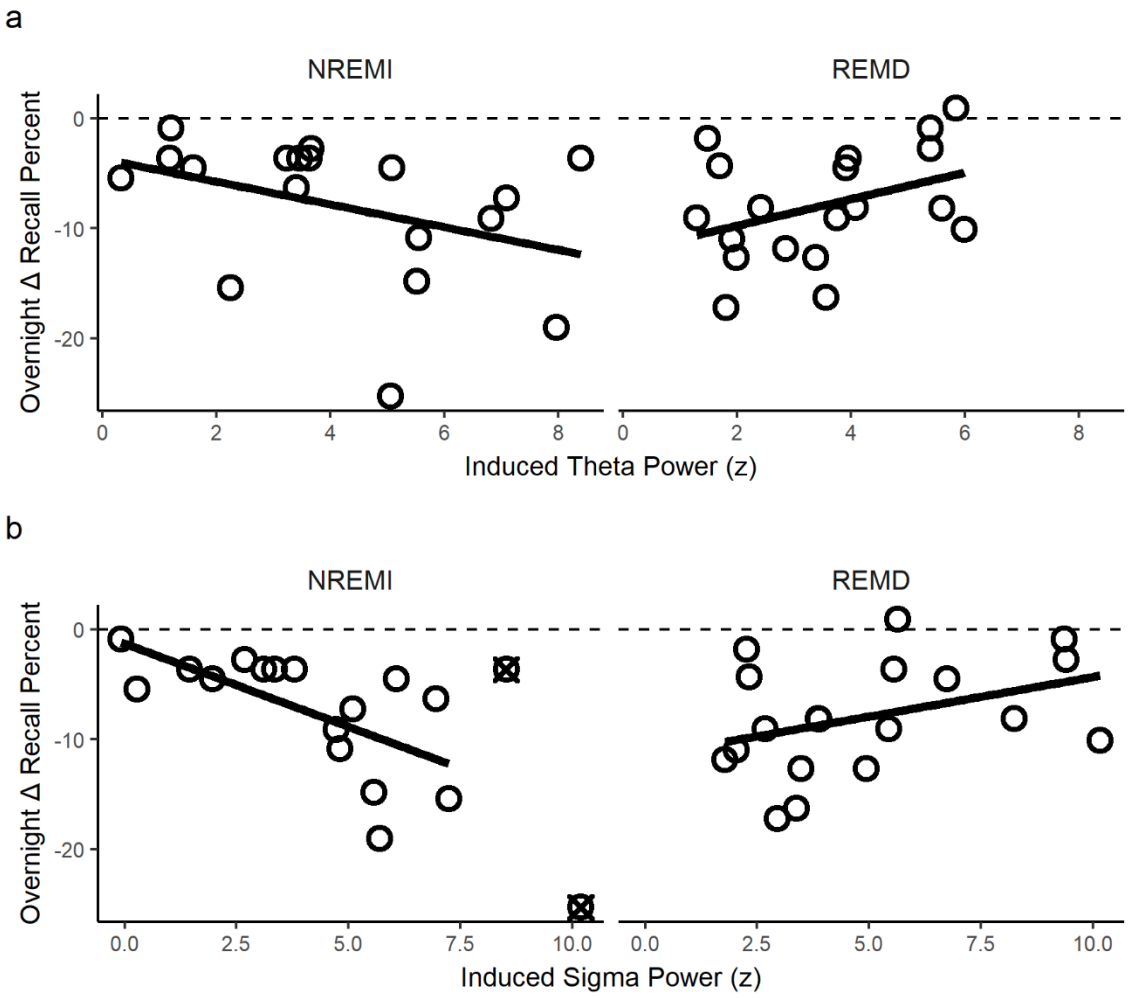

Figure 4.20. Overnight change in recall percent predicted by maximum cue-induced increases in electroencephalographic power over central channels separated by group. **a.** Overnight change in recall percent predicted by induced theta (4–7.5 Hz) power. **b.** Overnight change in recall percent predicted by induced sigma (12–15.5 Hz) power. Influential cases excluded from linear models in figure and report are marked by the × symbol.

investigation indicated a pattern for induced sigma power pattern like that for induced theta power, but with a stronger effect in the NREMI group. For Cmax, it was estimated that a within-group 1 SD increase in induced sigma power was associated with a 3.39-

point (*se* = 1.06) decrease in overnight retention of recall percent for the NREMI group, *t* = -3.18, and a 1.93-point (*se* = 1.14) increase in overnight retention of recall percent for the REMD group, *t* = 1.69 (Figure 4.20b).

Recall SD. For induced power measures, results from models predicting overnight change in recall SD were like those for models predicting overnight change in recall percent. None of induced delta, theta, or sigma power were significant predictors on their own, and they did not interact significantly with condition to predict overnight change in recall SD. Induced theta and sigma power, but not induced delta power, significantly interacted with group to predict overnight change in recall SD.

The induced theta power × group interaction was significant at channel C3 in an adjusted sample with three non-extreme influential cases (one from the NREMI group) excluded, $F(1, 30) = 12.76$, $p = .001$, $B = 2.65$, $se = 0.74$, and for Cmax in an adjusted sample excluding the same influential case from the NREMI group, $F(1, 32) = 4.80$, $p = .036$, $B = 1.48$, $se = 0.67$. Further investigation with the Cmax measure indicated the interaction was characterized by slight but opposing associations in each group (Figure 4.21a). It was estimated that a within-group 1 SD increase in induced theta power was associated with a 0.70º (*se* = 0.47) decrease in overnight decline of recall fidelity in the NREMI group, *t* = -1.49, and a 0.72º (*se* = 0.45) increase in overnight decline of recall fidelity in the REMD group, *t* = 1.62.

The induced sigma power × group interaction was significant at channel C4, $F(1, 33) = 4.22$, $p = .049$, $B = 1.34$, $se = 0.65$, and for Cmax, $F(1, 33) = 5.12$, $p = .030$, $B = 1.45$, $se = 0.64$. Further investigation with the Cmax measure indicated that induced sigma power had little association with overnight change in recall SD in the REMD group and had a negative association with overnight change in recall SD in the NREMI group

(Figure 4.21b). It was estimated that a within-group 1 SD increase in induced sigma power was associated with a 1.06º (*se* = 0.45) decrease in overnight decline of recall fidelity in the NREMI group, *t* = -2.36, and was, again, not clearly associated with overnight change in recall SD in the REMD group, *B* = 0.42, *se* = 0.47, *t* = 0.89.

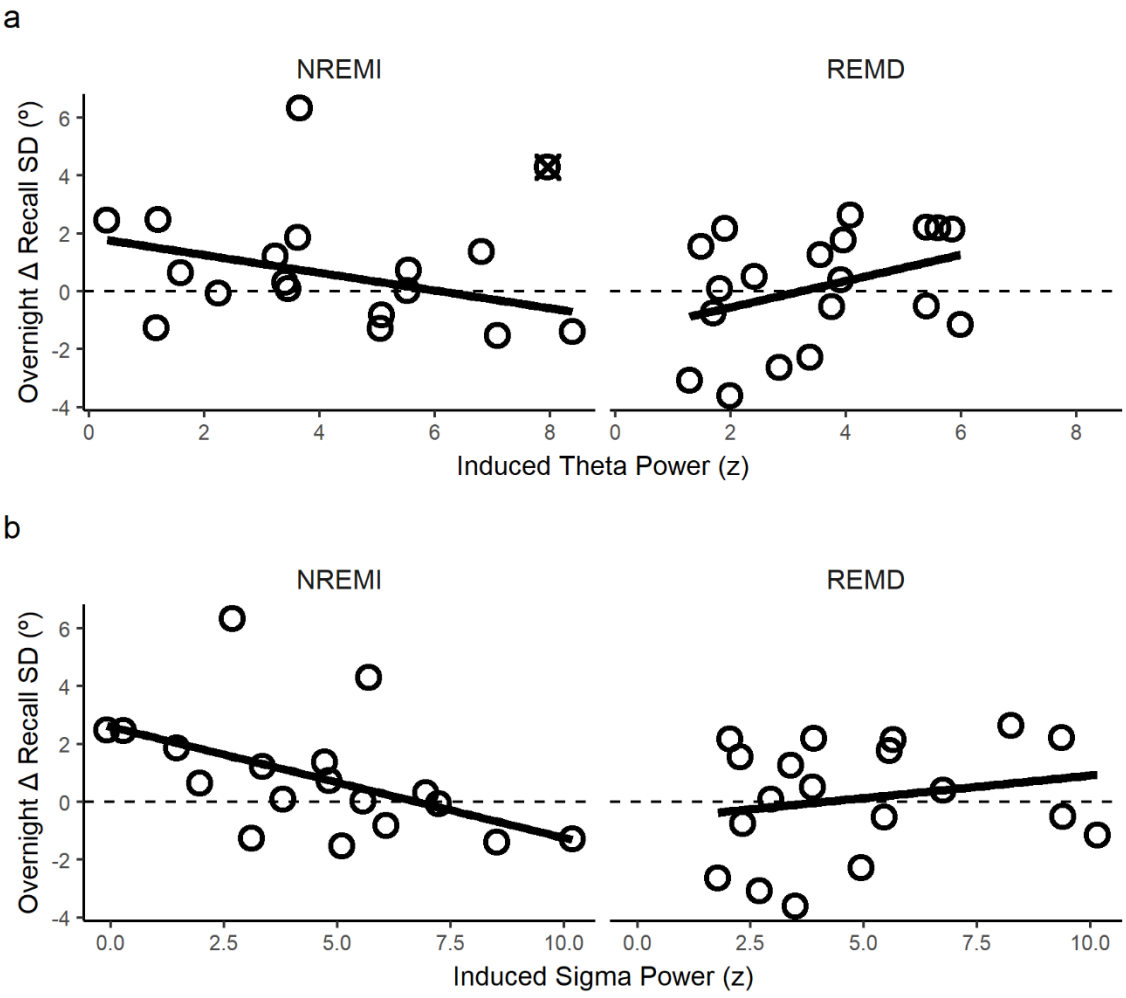

Figure 4.21. Overnight change in recall SD predicted by maximum cue-induced increases in electroencephalographic power over central channels separated by group. **a.** Overnight change in recall SD predicted by induced theta (4–7.5 Hz) power. Influential cases excluded from linear models in figure and report are marked by the × symbol. **b.** Overnight change in recall SD predicted by induced sigma (12–15.5 Hz) power.

Discussion

Study 3 was designed to test the SR2 hypothesis proposal that REM sleep processes refine newly acquired memories, supporting the fidelity at which they can be retrieved. With a substantial reduction of REM sleep in the REMD group, it was predicted that, compared to the NREMI group with intact REM sleep, the REMD group would show greater overnight declines in the fidelity at which they recalled item locations. NREM sleep factors of slow wave activity or items having been cued during NREM slow wave sleep were predicted to be linked to greater overnight retention of approximate item locations. However, within the REMD group, the declines in fidelity were predicted to be greater for individuals with more NREM slow wave activity during the night as it was hypothesized that the general NREM sleep processes supporting memory accessibility would result in worse fidelity if subsequent REM sleep refinement was limited or prevented.

Against prediction, there was no overall effect of the REMD–NREMI sleep manipulation: no group effects on either recall percent, recall SD, or recall confidence. It was also not the case that relatively greater declines in fidelity were observed in the REMD participants with greater NREM slow wave activity. Some results may, however, offer some support for the notion that REM sleep maintains the fidelity of memory representations and perhaps counters general amplification processes of NREM sleep. For REMD participants with comparatively deep NREM sleep, as indicating by a high N3 sleep duration to N2 sleep duration ratio, high sigma power in N2 and N3 sleep was associated with greater overnight retention of approximate item locations (Figure 4.12) but greater overnight declines in fidelity of recall (Figure 4.16). Deeper NREM sleep with greater sigma EEG power or spindle activity being associated with accessibility of item

location memories is consistent with the SR2 hypothesis and other hypotheses linking memory consolidation to the concert of slow waves and spindles in NREM sleep (Diekelmann & Born, 2010; Poe et al., 2010; Rasch & Born, 2013). Finding that these same features may also impair fidelity of recall if REM sleep is prevented suggests that REM sleep may have a role in maintaining memory fidelity. That is, NREM sleep may generally strengthen memory traces at a potential cost to fidelity that is observed if REM sleep is prevented but countered by refinement processes if REM sleep occurs. However, it is difficult to explain why memory accessibility benefits from sigma-rich, deep NREM sleep might only be observed in the REMD group and not in the NREMI group. Another question is why typical predictors of memory accessibility, slow wave activity of NREM sleep for example, did not similarly predict declines in fidelity in the REMD group. A separate result offering support for the proposal of REM sleep refinement was the finding that, within the REMD group only, having more item locations recalled at pre-sleep test was predictive of greater declines in recall fidelity over the night (Figure 4.9). It may be that those with the greater number of locations stored in memory would be most reliant on REM sleep refinement to maintain fidelity.

A curious negative association in the REMD group between duration of R sleep and greater retention of item location but greater declines in fidelity of recall (Figure 4.10) appears to contradict the proposal of REM sleep refinement. However, given that there were no overall group differences between NREMI and REMD groups despite their substantially different amounts of R sleep, it is unlikely that this correlation in the REMD group reflects a causal association of REM sleep degrading fidelity of recall. Instead, the correlation may reflect greater declines in fidelity among those who had a greater pressure for REM sleep or a subtle REM sleep deficit. Young adults under sleep

restriction show a dramatic pressure for REM sleep (Carskadon & Dement, 1981). Although participants in the current study were not sleep restricted, some participants may have carried subtle sleep deficits, especially REM-sleep deficits, at the onset of the study or as a result the "first night effect" (Agnew et al., 1966) from sleeping in the laboratory the previous night. These participants may have learned the item locations less deeply as a result of their subtle sleep deficit and then proceeded to both show declines in recall fidelity and have a high pressure for REM sleep. This association would be less apparent in the NREMI group as sleep architecture differences resulting from a REM sleep deficit might not manifest when REM sleep is allowed throughout the 8-hr sleep period (Webb & Agnew, 1975).

Another indication that REMD may have had subtle effects on memory performance was a pattern of some relationships with overnight changes in performance existing exclusively in the NREMI group and being seemingly disrupted by REMD. This pattern was observed with N3 sleep duration and N3 beta power (Figure 4.13). In the NREMI group, greater N3 sleep duration was predictive of better retention of memory accessibility particularly if that N3 sleep was low in beta power. In the REMD group, N3 duration had little association with overnight change in ability to recall approximate item locations regardless of the level of N3 beta power. If anything, greater N3 sleep with low beta power in the REMD group predicted a greater decline in item location recall. Given that greater beta EEG power during sleep is characteristic of arousals (American Sleep Disorders Association, 1992; Berry et al., 2015), has been associated with sleep maintenance insomnia (Merica et al., 1998) and the subjective perception of insomnia (Krystal et al., 2002), and is generally considered a sign of hyperactivity or hyperarousal (Riemann et al., 2010), this pattern in the NREMI group again reflects a tendency for

deep N3 sleep to be associated with retention of approximate item locations. The disruption in this relationship by the REMD manipulation points to a REM sleep having a complementary role in preserving memory representations. It was also the case that relationships of induced sigma and induced theta power predicting overnight changes in both the percentage of items recalled and the fidelity of recall were observed only in the NREMI group. The nature of these relationships is not fully understood and is discussed further on, but their absence in the REMD groups provides some indication that REMD may have had subtle effects on the interactions of sleep and memory.

As in Study 2, results of the current study deviated from past findings of memory benefits from TMR cueing during slow wave sleep (e.g., Rasch et al., 2007; Rudoy et al., 2009) as performance for cued items was not generally different than performance for control items. The results of the current study were only partially consistent with those of Creery et al. (2015) showing TMR benefits for high learners and for items with less accuracy. Although pre-sleep recall fidelity significantly moderated the cueing effect, it was, even more so than in the previous study, largely driven by reduced location retrieval for cued items relative to control items at high pre-sleep recall fidelity. Notably, this pattern was observed specifically in the NREMI group with intact REM sleep (Figure 4.7) and backed by a similar pattern of TMR decrements for those with high confidence in recall before sleep (Figure 4.8). No clear effect of cueing was observed for any level of pre-sleep fidelity or confidence in the REMD group. This TMR decrement for participants with REM sleep and high pre-sleep recall fidelity again raises the possibility that REM sleep had negative impacts on memory accessibility of cued items. If that was the case, the lack of clear TMR decrements and slight TMR benefits at low pre-sleep recall fidelity may indicate that typical TMR benefits like those observed by Creery et al.

(2015) counteracted these potential TMR decrements to memory accessibility.

Cueing also appeared to decrease recall fidelity for multiple groupings of participants, including those with low N3 duration but high N3 delta power (Figure 4.15), those with high N3 duration but low N3 theta power (Figure 4.18), and those with high N2 duration and high N2 beta power (Figure 4.19). It is speculative, but low N3 duration despite high delta power, little theta power in N3 sleep, or high N2 duration with high N2 beta power may all indicate cases in which deep NREM sleep was disrupted, potentially by sound cues or experimental awakenings, leading to sleep with higher levels of arousal on average. Under this assumption, TMR may be seen to decrease recall fidelity for cued items in the context of relatively higher arousal. It has been shown that odour based TMR during wakefulness destabilizes visuospatial memories, making them more susceptible to interference (Diekelmann et al., 2011). The results observed here may reflect a similar effect: cueing during more fragile sleep with signs of high arousal, albeit not wakefulness, could destabilize the memories and decrease the fidelity at which they are later recalled.

Regarding changes to the EEG power spectrum induced by the cues played during NREM slow wave sleep, it is notable that the results were like those of Study 2, albeit with some differences. The pattern of induced power responses, increased delta, theta, and sigma power at varied post-cue time intervals was nearly identical to the pattern observed in Study 2 (Figure 3.3 and Figure 4.5). Furthermore, the extent to which individuals showed induced sigma power post-cue predicted worse overnight retention of approximate item locations in both Study 2 (Figure 3.12) and in the NREMI group of the current study (Figure 4.20). Although, induced delta power was not significantly associated with overnight changes in memory performance in the current study, induced theta power was associated with worse overnight retention of approximate item locations

in the NREMI group of the current study (Figure 4.21). Again, the pattern of induced power changes may reflect a tendency in light sleepers to respond to cues with K-complex and spindle responses that are normally sleep-protective (Bastien et al., 2000; Dang-Vu et al., 2010; Dang-Vu et al., 2011; Jahnke et al. 2012) but ultimately fail, allowing disruptions of sleep-dependent memory processing and subsequent declines in retrieval of item locations. The accompanying association between greater induced sigma and theta power and better retained or overnight increases in recall fidelity in the NREMI group suggests that it was memories with low-fidelity representations that were no longer accessible for retrieval after the night. Moreover, the absence of this pattern in the REMD group suggests that REM sleep may be involved in the loss of these low-fidelity memories in the NREMI group, an effect that would be consistent with findings of greater REM duration predicting more forgetting of low-value items (Oudiette et al., 2013).

Beyond the experimental sleep and cueing manipulations, the predicted association between greater NREM slow wave activity and greater retention of approximate item locations was observed. With a significant N3 duration × N3 delta power interaction predicting overnight change in recall percent, it was again found that retention of approximate item locations was highest for those with large amounts of delta-rich N3 sleep (Figure 4.11). As in the previous studies, the N3 delta power measures for which this interaction was most apparent were those from parietal and occipital channels. The robustness of this pattern is supportive of the SR2 hypothesis proposal that NREM slow wave sleep benefits the accessibility of memories for retrieval.

The current study also yielded some unexpected results. First, retention of recall fidelity was unexpectedly high overall as participants on average did not show the expected significant decline in recall fidelity over the night with some participants even

having greater fidelity of recall in the morning than they did before sleep. Overnight declines in recall of item locations tended to be greater for participants with lower pre-sleep recall fidelity and for participants showing general gains in fidelity over the night. Together, these correlations suggest many participants likely lost accessibility to (i.e. failed to recall) item locations that they had recalled with only low accuracy before sleep, leading to a comparative increase in overall fidelity after sleep given that it is estimated for only the items recalled successfully. A second unexpected finding was that this overnight retention or increase in recall fidelity was associated with greater N2 sleep duration (Figure 4.17). This finding was especially unexpected as participants in REMD group tended to have greater N2 sleep duration than participants in the NREMI group and it was the REMD group that was predicted to show the greatest declines in recall fidelity. Interpretation of this correlation found over the full sample is made difficult by the fact that the factors that contributed to the amount of N2 sleep participants obtained likely varied within individuals and groups, and might include, for example, the number of experimental awakenings, the ease in returning to sleep after awakening, the pressure for either N3 or R sleep, or the tendency to show arousals to TMR cues. While this association may reflect a culmination of other associations in various subsamples, it may of course be the case that N2 sleep has an active role in maintaining fidelity of memory representations that is underappreciated by the SR2 hypothesis as proposed.

The absence of general REMD-induced reductions in recall fidelity and signs of improved fidelity among REMD participants with the greatest amounts of N2 sleep may be explained by the concept of covert REM sleep and the nature of N2-to-R sleep, a term to capture the state of NREM sleep that precedes unambiguous REM sleep by a few moments and may itself be ambiguous with respect to scoring as NREM or REM sleep.

Poe et al. (Poe, 2017; Poe et al., 2010) argued that it is not only REM sleep that satisfies the conditions required for the depotentiation of weak synapses associated with hippocampal memories, but that these conditions are also met within periods of N2 sleep: Low forebrain norepinephrine is essential for such depotentiation (Katsuki et al., 1997; O'Dell et al., 2015; Thomas et al., 1996; Yang et al., 2002) and norepinephrine-providing neurons of the locus coeruleus of the rat are inactive in the second before spindles and the 20–30 s before transitions to paradoxical sleep (Aston-Jones & Bloom, 1981) or, in terms of a human analogue, N2-to-R sleep. From a different perspective, Nielson (2000) has argued that REM sleep processes can occur covertly during polysomnographically-scored NREM sleep. Notably, the probability of this covert REM sleep is thought to be increased during REMD. To this point, animal P waves occur predominantly in REM sleep but are also present in NREM sleep before transitions to REM sleep (Jouvet, 1962). Moreover, the frequency of their occurrence within NREM sleep increases during REMD (Dusan-Peyrethon et al. 1967; Ferguson & Dement, 1969) and following intense learning (Datta, 2000), and post-learning increases in P wave density predict retention of two-way avoidance learning (Datta, 2000). It was proposed that the REMD group would show reduced recall fidelity due to an absence of REM sleep refinement; however, it may be that memory refinement was sufficiently executed during periods of covert REM sleep or transitional N2 sleep occurring before the protocol demanded an experimental REMD awakening. P waves may be a mechanism of this refinement process whether they occur during REM sleep or, in the case of REMD, during frequent periods of N2-to-R sleep.

There are some limitations particular to the current study regarding the implementation of the REMD and NREMI procedures. Of course, participants were not completely deprived of REM sleep either measurable in stage R epochs or immeasurable

as covert REM sleep. Dramatic reductions in stage R duration were achieved, but the quantity of REM sleep required to achieve any proposed effect on memory performance is unknown. However, previous REMD studies have indeed resulted in impairments in memory performance for tasks learned before sleep (e.g., Empson & Clarke, 1970; Smith et al., 2004; Tilley & Empson, 1978), so this is an unsatisfactory explanation to account for the lack of REMD effect on memory performance. In any case, further investigation with different methods of having an absence of REM sleep, such as short naps or pharmacological manipulations, may be considered to better understand the effect of the presence or absence of REM sleep. Some variation of method may also avoid another limitation of the current study: the reduction in N2 sleep and, to a much-lesser extent, N3 sleep in the NREMI group. Each possible control to a REMD comes with its own limitations, and, for the current study, it is important to remember that group differences may be driven either by dramatic reductions in REM sleep duration or opposing reductions of NREM sleep.

Although the current study and Study 2 shared much of the same methodology, there were two differences beyond the REMD and NREMI manipulations that may limit the extent to which their results can be directly compared. One difference was a slight change in task instruction in the item-location task. Specifically, in the current study, participants were additionally encouraged before each test to “take an extra moment” to place each item as close to its exact location as they could if they felt that they knew that location. This additional instruction was added due to concern that some participants in the previous study were not careful to be accurate in their item placements, thus devaluing the recall SD measure. Notably, the previous study saw an overnight decline in recall fidelity in the group average that was not present in the current study. This

difference in outcome might be the result of additional efforts to maintain fidelity during the post-sleep test in the current study.

A second difference to note was that participants in the current study often received fewer TMR cues than participants in Study 2. On average, only 3.9 cues were delivered during the night for each cued item in the current study, a reduction from the average 4.4 cues delivered for each cued item in Study 2. Further, in the current study, roughly half as many participants received all 500 cues (11 vs. 21) and three times as many participants received less than 300 cues (9 vs. 3). This reduction may be due to various factors including the reselection of some sound files for this the current study, but it is thought to largely be the result of the concurrent implementation of the REMD or NREMI procedures. The amount of cueing needed to obtain a TMR effect is unknown, but it is reasonable to assume there is a lower limit for which cueing may not be effective. The numbers of cues played in Study 2 are thought to be like other similar studies using sound based TMR during slow wave sleep (e.g., Creery et al., 2015, Rudoy et al., 2009) with respect to the number of times each item had its cue played during slow wave sleep. The lower number of cues played during the current study may have been a factor in the relative absence of TMR effects on overnight changes in memory performance.

Chapter 5

General Discussion

The purpose of this investigation was to develop a greater understanding the potential individual contributions of NREM and REM sleep to the benefits of sleep on the later retrieval of newly acquired memories. The SR2 hypothesis of sleep reinforcement and sleep refinement was formulated and presented, attributing benefits in memory accessibility primarily to NREM sleep reinforcement and benefits to memory fidelity primarily to REM sleep refinement. As described in Chapter 1, this hypothesis was formulated in consideration of existing hypotheses of memory processing during sleep, including those proposed by Crick and Mitchison (1983), Diekelmann and Born (2010), Giuditta (2014), Poe (2017), Poe et al. (2010), Rasch and Born (2013), and Tononi and Cirelli (2003, 2006), and the tendency for complex memory tasks to be more dependent on REM sleep than simple memory tasks (Ackermann & Rasch, 2014; Smith et al., 2004; Stickgold, 1998; Tilley et al., 1992). It was noted that the SR2 hypothesis is consistent with recent work linking NREM sleep to task-specific dendritic spine formation (Yang et al., 2014) and REM sleep to selective strengthening and weakening of newly formed dendritic spines (Li et al., 2017). Studies directly testing the SR2 hypothesis were presented in Chapter 2, Chapter 3, and Chapter 4 using memory tasks designed to distinguish between memory accessibility and memory fidelity. The current chapter will offer conclusions that can be made from these studies regarding the two components of the SR2 hypothesis and discuss the strengths and limitations of the methods shared among these studies.

NREM Sleep Reinforcement Benefits Memory Accessibility

NREM sleep properties, particularly those involved in the previously described

concert of slow waves, spindles, and memory reactivations, were proposed to be indicative of NREM sleep reinforcement. Thus, the major predictions made for this component of the SR2 hypothesis were that sleep would benefit memory accessibility relative to wake, that properties of NREM sleep, particularly NREM sleep slow wave activity, would be positively associated with retention of item features learned before sleep, and that TMR cueing during NREM slow wave sleep would result in a relative increase in retrieving the tested features from memory.

As predicted, in all three studies conducted, slow wave activity of post-learning NREM sleep was positively associated with an ability to retain and retrieve the tested features from memory. This association was expressed both with measures of 1–3.5 Hz delta EEG activity and with amounts of N3 sleep which is defined largely by the presence of this slow wave activity. In a post-learning nap, both greater N3 sleep duration and greater delta EEG power in stage N2 and N3 epochs predicted more retrieval success for learned item colours. In both studies including a full night of post-learning sleep, it was participants with both high N3 sleep duration and high delta EEG power in stage N3 epochs who showed the greatest retention of the ability to retrieve learned item locations. The consistency of this finding is echoed by the fact that, in all three studies, the relationships between the memory accessibility measure and delta EEG power of NREM sleep were most prominent over posterior scalp electrodes (e.g., PO7, Oz, and PO8)

Beyond relationships with delta EEG power, there were other associations observed that support of the notion that memory accessibility benefits from deep NREM sleep and features specific to NREM sleep. Given that, in NREM sleep, sigma EEG power captures activity of spindles and greater alpha EEG power and beta EEG power are likely markers of lighter sleep and arousal (American Sleep Disorders Association, 1992;

Asyali et al., 2007; Berry et al., 2015; McKinney et al., 2011), it is noteworthy that relatively high retention of retrieval success was observed in at least one of the two full-night studies in each of the following clusters of participants: participants with large amounts of N2 sleep with low alpha EEG power over posterior electrodes, participants with intact REM sleep and high amounts of N3 sleep with low beta EEG power, and REMD participants with relatively deep NREM sleep with high sigma EEG power. Finally, changes in EEG power spectra induced by the TMR sound cues, particularly increases in sigma power, predicted declines in memory retrieval in both TMR studies. If this neurological response to sound cues reflects a disruption to the NREM slow wave sleep in which they were delivered, these findings are also in line with NREM slow wave sleep benefiting memory accessibly.

Of course, positive correlations between NREM sleep slow wave activity and memory performance are not novel and have been observed regularly in previous research (Diekelmann et al., 2012; Holz et al., 2012; Huber et al., 2004; Landsness et al., 2009; Ruch et al., 2012; Schabus et al., 2005; Schönauer et al., 2017; Tamaki et al., 2013). Indeed, the proposed role for NREM sleep in reinforcing memory representations is largely aligned with more established models such as the active system consolidation hypothesis (Diekelmann & Born, 2010; Rasch & Born, 2013) which proposes that NREM sleep slow wave activity mediates transformation of memories from labile and temporary states to more stable and long-lasting states. By moving beyond a unitary measure of memory performance it was found, as predicted, that NREM sleep slow wave activity was reliably associated with memory accessibility, that is, recall percent and overnight changes in recall percent. Despite moderate correlations between measures of accessibility and fidelity and overnight changes in these measures, associations between

NREM sleep properties and memory fidelity were generally not observed. This specificity supports the proposed dissociation aligning NREM sleep specifically with benefits to memory accessibility. The only significant association between NREM sleep properties and changes in memory fidelity was a tendency for greater declines in memory fidelity observed among REMD participants with relatively deep NREM sleep and high sigma power in N2 and N3 sleep. This latter result may indicate that processes of NREM sleep reinforcement have negative impacts on fidelity when complementary REM sleep refinement is disrupted.

Although correlational findings were clearly supportive of the proposed role for NREM sleep, findings from the experimental manipulation of NREM sleep processes were not clearly supportive. Specifically, TMR during NREM slow wave sleep did not result in overall benefits to memory accessibility for cued items relative to non-cued control items in either study using this method. In fact, TMR decrements to memory accessibility were observed at least as often as evidence of TMR benefits to memory accessibility. However, if one considers the results in purely relative terms of either increased TMR benefit or a decreased TMR decrement, results from both TMR studies are consistent with past research in that they show a relatively positive TMR effect for participants with low recall fidelity before sleep (as in Creery et al., 2015) and evidence of supplementing lesser amounts of NREM slow wave sleep to produce the same effect as greater amounts of NREM slow wave sleep (as in Diekelmann et al., 2012). As discussed previously, the relative TMR benefits observed here may reflect a tendency for TMR benefits to oppose additional TMR decrements occurring in subsequent theta-rich REM sleep as such decrements were not observed in the REMD group and direct TMR benefits were observed in participants with high stage R duration but low stage R theta power.

Given that these direct TMR benefits were observed for overnight change in recall percent, that the relative benefits consistent with past research were also for this measure, and that no TMR benefits to recall fidelity were observed, the studies conducted here offer some support for the SR2 hypothesis proposal that NREM sleep reinforces memory representations to maintain their accessibility for retrieval. However, given the absence of a clear overall cueing benefit, this interpretation is far from conclusive and additional experimental research is warranted.

REM Sleep Refinement Benefits Memory Fidelity

REM sleep was proposed to be responsible for a step of memory refinement that supports the fidelity of memory representations and thus the accuracy and precision at which these memories are retrieved. The major prediction made for this component of the SR2 hypothesis was that the fidelity of successful memory retrieval would be better supported by post-learning retention periods with more REM sleep than post-learning retention periods with less REM sleep, including periods of only wakefulness, sleep periods with naturally less REM sleep, or sleep periods subjected to the REMD manipulation.

Over the three studies, there was little evidence found to support this component of the SR2 hypothesis. There was no clear benefit of REM sleep on recall SD (i.e., the standard deviation of error estimated for instances of recall success) or overnight changes in this measure. In Study 1, there were difficulties obtaining valid recall SD measures due to poor performance; although, for what was analyzed, recall SD did not differ between sleep and wake conditions nor did an effect of sleep differ by whether or not participants obtained at least 5 min of stage R sleep in their nap. In Study 2, participants showed an increase in recall SD from pre-sleep test to post-sleep test, indicating a decline in fidelity,

but this change did not relate to R sleep duration. In Study 3, despite dramatic reductions in R sleep duration from the REMD manipulation, overnight changes in recall SD did not differ between the REMD group and the NREMI group with intact REM sleep, and neither of these groups showed the overnight declines in fidelity observed in Study 2. A result of relatively deep sigma-rich NREM sleep predicting overnight increases in both recall percent and recall SD in the REMD group, but not the NREMI group, may be considered some evidence for the SR2 hypothesis proposal that REM sleep refinement works to counter decreases in fidelity that might otherwise occur as a result of the more generally applied NREM sleep reinforcement. There was also some evidence that REM sleep might decrease memory fidelity. In the REMD group of Study 3, participants who managed to obtain more R sleep had greater declines in fidelity than REMD participants with less R sleep, an unexpected correlation that was also marginally significant in Study 2. However, given the absence of similar findings of fidelity declines associating with more reliable predictors of memory accessibility measures (e.g., NREM sleep slow wave activity) and the absence of any REMD effect on recall SD, these interpretations are not particularly convincing.

In all three studies, there was also no clear evidence of any REM sleep effect on the ability to retrieve features from memory (i.e., memory accessibility). A generous interpretation of this finding might be that it at least supports the proposed dissociation of it being NREM sleep and not REM sleep that is primarily responsible for supporting memory accessibility. One might also argue the lack of identifiable effects and associations suggest post-learning REM sleep does not influence memories. However, this interpretation is at odds with the abundance of research discussed in Chapter 1 supporting a role for REM sleep in memory processing as a state of bidirectional

plasticity. It is more likely that the methods used across these three studies were not sensitive to the effects of REM sleep. Thus, the results do little to address the proposed role for REM sleep refinement in benefitting memory fidelity expect to suggest REM sleep effects on memory performance may be more subtle than expected. Further research attempting to elucidate how post-learning REM sleep impacts memory performance is warranted.

Strengths and Limitations

One aspect of this investigation that was novel in the context of research on sleep-related memory benefits was the differentiation of memory accessibility and memory fidelity using a method largely based on work in visual short-term and working memory (e.g., Suchow et al., 2013; Zhang & Luck, 2008). There remains value in analyzing memory performance with a single measure at either an aggregate or individual item level, but the differentiation of recall percent (the measure of memory accessibility) and recall SD (the inverted measure of memory fidelity) was considered a strength for testing the SR2 hypothesis. These measures were obtained by using continuous-report tasks and treating each distribution of response errors at test as a mixture of a uniform distribution of response errors associated with guessing and a normal distribution of response errors centered around the target when recall was successful. This method allowed for one important conclusion regarding the relationship of sleep and memory: memory benefits of NREM sleep, specifically the benefits of NREM sleep slow wave activity on relatively short long-term visuospatial memories, appear to come more from maintenance of memory accessibility than a maintenance or increase in memory fidelity.

Of course, conclusions relating to memory accessibility and memory fidelity are only valid to the extent that the respective calculated measures of recall percent and recall

SD truly assess these concepts. Recall percent measures were largely determined by counting the number of responses too far from the target to be considered successful and extrapolating to estimate the number of instances of successful memory retrieval. The simplicity of this approach supports its validity as a measure of the accessibility of memory representations for retrieval at test. Recall SD, a relatively novel measure for use with memories held over periods longer than a few seconds, was determined by finding the parameters of the best mixed uniform-normal distribution fit to the distribution of response errors while confining the rate of sampling from the uniform distribution closely to the approximated rate of recall success. The relative complexity and novelty of this approach and the difficulty in capturing fidelity relative to all-or-none retrieval success makes the validity of recall SD more suspect.

The validity of recall SD as a measure of error within instances of successful retrieval, and thus an inverted measure of memory fidelity cannot be determined from these studies alone. However, the ways in which measures of recall SD and overnight changes in recall SD were found to correlate with other memory performance measures offers some reassurance that recall SD reflects memory fidelity. First, recall SD was moderately negatively correlated with average reported confidence of recall for all recall tests except the post-sleep recall test of Study 2, and recall SD was moderately negatively correlated with recall percent for both recall tests of Study 3. Thus, greater estimated error within instances of recall success was associated with lower average reported confidence and a lower percentage of learned items successfully recalled. These correlations are consistent with the intuitive notions that confidence in recall is higher when memory fidelity is higher and that participants who do well to recall many learned item locations also do well to recall them with high accuracy. A moderate positive correlation between

overnight change in recall percent and overnight change in recall SD was observed in Study 2 and Study 3, indicating a tendency for relative declines in the percentage of locations recalled being associated with relative gains in the accuracy of the locations that were successfully recalled. Gains on one performance measure associating with declines on another may seem counterintuitive; however, because recall SD was an aggregative measure calculated within instances of recall success, declines in error (i.e., gains in fidelity) would be expected among individuals who, over the night, forget item locations for which they had low fidelity representations, restricting their instances of recall success in the morning to the item locations for which they had high fidelity representations.

An additional concern regarding the recall SD measure and the task from which it was collected is recall SD may not have been sensitive to real effects of sleep or, more specifically, the proposed REM sleep refinement. The extent to which recall SD might be expected to decrease by hypothetical REM sleep refinement is not clear. The significant increase in recall SD observed over the night in Study 2 was characterized by only a 1.06º increase in recall SD on average. The average increase in error from pre-nap test to post-nap test in Creery et al. (2015) was roughly 5.5% of the size of the images used in their similarly designed task. Given these findings and the fact that images occluded 15º of the response circle in the item-location task, the average overnight decline in fidelity attributable to a hypothetical prevention of refinement processes is anticipated to be no more than 2–3º of error. If, over the course of 200 item placements, participants were not careful to place items to the absolute best of their ability and instead quickly and roughly placed them anywhere within 5 and 10º of their best guess, the measurement error included in recall SD might be too large for true effects on memory fidelity to be identified. Given the focus on overnight changes in performance measures, the negative

impact of such non-peak performance would be even greater if participants were more or less likely to perform in this manner on either the pre-sleep or post-sleep tests. Although no certain evidence of such performance patterns was obtained, it is certainly plausible that at least some participants may have rushed their performance to either get to sleep or complete their participation more quickly.

Different methods of acquiring memory performance measures may have offered greater sensitivity to possible effects or otherwise altered study outcomes. As mentioned, the chosen method of analysis was inspired by work in visual short-term and working memory. Some models of visual short-term and working memory argue that the precision of memory varies across items and thus may be better accounted for by a mixture of normal distributions of varied precisions rather than a single normal distribution of fixed precision (Fougnie et al., 2012; van den Berg et al., 2012; reviewed in Ma et al., 2014). Variable precision models produce estimates of both the precision of memory and the variability in precision. True variability in precision among long-term memory representations is likely, and estimation of this variability might further inform the understanding of the sleep and memory relationship. Notably, the chosen method of analysis did not estimate the height of the uniform distribution of guessing primarily through probabilistic methods (e.g., maximum likelihood estimation) as is often the case when modelling visual working memory (see Suchow et al., 2013), but instead first extrapolated a preliminary estimate from a defined region of recall failure (error over 90º for Study 1 and error over 80º for Study 2 and Study 3). Selection of these criteria were carefully considered, but it is noted that conclusions drawn from these studies may have differed if different criteria were selected, if purely probabilistic methods were used, or if different parameters were acquired from the distributions of response errors. Furthermore,

while summary statistics of memory accessibility and memory fidelity were desired for initial testing of the SR2 hypothesis, item-level approaches to analyses could and should be considered for future investigations of the SR2 hypothesis and of sleep and memory more generally as insight may be gained from tracking the quality and fate of individual memory representations over time and sleep.

Additional strengths and limitations exist within the timing of key events in the experiments, particularly in Study 2 and Study 3. In these studies, effort was made to hold constant both the time between the pre-sleep recall test and sleep and the time between the pre-sleep recall test and post-sleep recall test. This was largely achieved with a fixed schedule in which the pre-sleep recall test began at 23:00, the sleep period began shortly after this test at 23:30, and the post-sleep recall test began at 9:00. Even for participants who did not complete the learning period before 22:55, adjustments were made such that the sleep period and post-sleep recall test respectively began shortly after and exactly 10 hr after the pre-sleep recall test. This precision in task timing may be considered a strength of the method, particularly as memory change over the sleep period and between the two tests was the focus of analyses. However, fixing the start of the pre-sleep recall test meant a variable amount of time between the end of the learning period and the start of the test. Given the variation in learning period duration, the time between the end of the learning period and the start of the pre-sleep recall test regularly varied within 5 and 35 min. This variation likely added error variation in measures of pre-sleep recall test performance and overnight change in performance.

The administration of only one post-sleep recall test occurring shortly after the sleep period in each of the three studies may also have limited the ability to find effects of sleep or experimental manipulations of sleep. Beneficial effects of proposed NREM sleep

reinforcement and REM sleep refinement may become more apparent as time passes between sleep and test. Future investigations would benefit from follow-up testing at least one full day after the critical sleep period to allow more time for memory decay and perhaps more opportunity to observe effects of sleep that counteract memory decay.

Another point to consider regarding the results of these studies is the rather liberal approach to statistical testing. The purpose of this investigation was not to definitively test whether specific predictions of the SR2 hypothesis met a conservative threshold of statistical significance but to initiate investigation into the hypothesis by testing general predictions without heavy concern for the potential of type I error. For example, N3 sleep duration, delta EEG power in N3 sleep, and NREM sleep spindles were each examined as predictors of overnight changes in memory performance with α set to .05 for each test as each test was directly relevant to the proposal that these properties of NREM sleep support the accessibility of newly acquired memories. In addition, there was thorough exploration of less relevant measures (e.g., beta EEG power in N2 sleep) in effort to gain and present a detailed account of how sleep associated with memory accessibility and memory fidelity. Actions to reduce type I error included lowering α to .01 for the less relevant measures and using the Benjamini–Hochberg method (Benjamini & Hochberg, 1995) to limit false discovery rate within sets of related measures (e.g., the 13 measures of N3 delta power); however, spurious findings remain a considerable possibility. In part to complement and counter the liberal approach to statistical testing, there was extensive effort to limit so-called researcher degrees of freedom by using a methodical analysis procedure held nearly constant over all three studies. Furthermore, the conclusions of this investigation are largely tethered to results that replicated over multiple studies or fit within a pattern of similar results.

Conclusion

This doctoral dissertation contributes a detailed review of related and supporting literature for a new hypothesis proposing that the benefits of post-learning sleep on later memory performance may be explained by NREM sleep processes of memory reinforcement and REM sleep processes of memory refinement. A series of three studies directly testing this hypothesis were reported. The results of these studies further inform an understanding of the benefits of post-learning NREM slow wave sleep in maintaining the accessibility of visuospatial memories for retrieval. However, the results offer little insight into the effects of REM sleep on memory performance, and an understanding of the relationship between REM sleep and memory performance remains elusive. The three reported studies, of course, could not conclusively test the SR2 hypothesis, and it is hoped that their results and the SR2 hypothesis stimulate additional inquires that will further our understanding of the sleep and memory relationship.

References

Abraham, W. C., & Williams, J. M. (2008). LTP maintenance and its protein synthesis-dependence. *Neurobiology of Learning and Memory*, *89*(3), 260–268. https://doi.org/10.1016/j.nlm.2007.10.001

Ackermann, S., & Rasch, B. (2014). Differential effects of non-REM and REM sleep on memory consolidation? *Current Neurology and Neuroscience Reports*, *14*(2). https://doi.org/10.1007/s11910-013-0430-8

Aeschbach, D., Cutler, A. J., & Ronda, J. M. (2008). A role for non-rapid-eye-movement sleep homeostasis in perceptual learning. *Journal of Neuroscience*, *28*(11), 2766–2772. https://doi.org/10.1523/JNEUROSCI.5548-07.2008

Agnew, H. W. J., Webb, W. B., & Williams, R. L. (1966). The first night effect: an EEG study of sleep. *Psychophysiology*, *2*(3), 263–266. https://doi.org/10.1111/j.1469-8986.1966.tb02650.x

American Sleep Disorders Association. (1992). EEG arousals: Scoring rules and examples. A preliminary report from the sleep disorders atlas task force of the American Sleep Disorders Association. *SLEEP*, *15*(2), 174–184. https://doi.org/10.1093/sleep/15.2.174

Antony, J. W., Gobel, E. W., O'Hare, J. K., Reber, P. J., & Paller, K. A. (2012). Cued memory reactivation during sleep influences skill learning. *Nature Neuroscience*, *15*(8), 1114–1116. https://doi.org/10.1038/nn.3152

Aston-Jones, G., & Bloom, F. (1981). Activity of norepinephrine-containing locus coeruleus neurons in behaving rats anticipates fluctuations in the sleep-waking cycle. *Journal of Neuroscience*, *1*(8), 876–886. https://doi.org/10.1523/JNEUROSCI.01-08-00876.1981

Asyali, M. H., Berry, R. B., Khoo, M. C. K., & Altinok, A. (2007). Determining a continuous marker for sleep depth. *Computers in Biology and Medicine*, *37*(11), 1600–1609. https://doi.org/10.1016/j.compbiomed.2007.03.001

Auguie, B. (2016). *gridExtra: Miscellaneous functions for "grid" graphics*. https://CRAN.R-project.org/package=gridExtra

Baltaci, S. B., Mogulkoc, R., & Baltaci, A. K. (2019). Molecular mechanisms of early and late LTP. *Neurochemical Research*, *44*(2), 281–296. https://doi.org/10.1007/s11064-018-2695-4

Barakat, M., Doyon, J., Debas, K., Vandewalle, G., Morin, A., Poirier, G., Martin, N., Lafortune, M., Karni, A., Ungerleider, L. G., Benali, H., & Carrier, J. (2011). Fast and slow spindle involvement in the consolidation of a new motor sequence. *Behavioural Brain Research*, *217*(1), 117–121. https://doi.org/10.1016/j.bbr.2010.10.019

Barrett, T. R., & Ekstrand, B. R. (1972). Effect of sleep on memory: III. Controlling for time-of-day effects. *Journal of Experimental Psychology*, *96*(2), 321–327. https://doi.org/10.1037/h0033625

Bastien, C. H., Ladouceur, C., & Campbell, K. B. (2000). EEG characteristics prior to and following the evoked K-Complex. *Canadian Journal of Experimental Psychology*, *54*(4), 255–265. https://doi.org/10.1037/h0087345

Bastuji, H., & García-Larrea, L. (1999). Evoked potentials as a tool for the investigation of human sleep. *Sleep Medicine Reviews*, *3*(1), 23–45. https://doi.org/10.1016/S1087-0792(99)90012-6

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

https://doi.org/10.18637/jss.v067.i01

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, *4*(6), 561–571. https://doi.org/10.1001/archpsyc.1961.01710120031004

Belal, S., Cousins, J., El-Deredy, W., Parkes, L., Schneider, J., Tsujimura, H., Zoumpoulaki, A., Perapoch, M., Santamaria, L., & Lewis, P. (2018). Identification of memory reactivation during sleep by EEG classification. *NeuroImage*, *176*, 203–214. https://doi.org/10.1016/j.neuroimage.2018.04.029

Bendor, D., & Wilson, M. A. (2012). Biasing the content of hippocampal replay during sleep. *Nature Neuroscience*, *15*(10), 1439–1444. https://doi.org/10.1038/nn.3203

Benito, E., & Barco, A. (2010). CREB's control of intrinsic and synaptic plasticity: Implications for CREB-dependent memory models. *Trends in Neurosciences*, *33*(5), 230–240. https://doi.org/10.1016/j.tins.2010.02.001

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Benson, K., & Feinberg, I. (1977). The beneficial effect of sleep in an extended Jenkins and Dallenbach paradigm. *Psychophysiology*, *14*(4), 375–384. https://doi.org/10.1111/j.1469-8986.1977.tb02967.x

Berry, R. B., Brooks, R., Gamaldo, C. E., Harding, S., Marcus, C., & Vaughn, B. (2015). *The AASM manual for the scoring of sleep and associated events: Rules, terminology, and technical specifications, version 2.2*. American Academy of Sleep Medicine.

Borbély, A. A., & Achermann, P. (1999). Sleep homeostasis and models of sleep regulation. *Journal of Biological Rhythms*, *14*(6), 559–570. https://doi.org/10.1177/074873099129000894

Boyce, R., Glasgow, S. D., Williams, S., & Adamantidis, A. (2016). Causal evidence for the role of REM sleep theta rhythm in contextual memory consolidation. *Science*, *352*(6287), 812–816. https://doi.org/10.1126/science.aad5252

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences*, *105*(38), 14325–14329. https://doi.org/10.1073/pnas.0803390105

Brady, T. F., Konkle, T., Gill, J., Oliva, A., & Alvarez, G. A. (2013). Visual long-term memory has the same limit on fidelity as visual working memory. *Psychological Science*, *24*(6), 981–990. https://doi.org/10.1177/0956797612465439

Bragin, A., Engel Jr, J., Wilson, C. L., Fried, I., & Buzsáki, G. (1999). High-frequency oscillations in human brain. *Hippocampus*, *9*(2), 137–142. https://doi.org/10.1002/(SICI)1098-1063(1999)9:2<137::AID-HIPO5>3.0.CO;2-0

Bramham, C. R. (2008). Local protein synthesis, actin dynamics, and LTP consolidation. *Current Opinion in Neurobiology*, *18*(5), 524–531. https://doi.org/10.1016/j.conb.2008.09.013

Bramham, C. R., & Srebro, B. (1989). Synaptic plasticity in the hippocampus is modulated by behavioral state. *Brain Research*, *493*(1), 74–86. https://doi.org/10.1016/0006-8993(89)91001-9

Buzsáki, G. (2002). Theta oscillations in the hippocampus. *Neuron*, *33*(3), 325–340. https://doi.org/10.1016/S0896-6273(02)00586-X

Buzsáki, G. (2015). Hippocampal sharp wave-ripple: A cognitive biomarker for episodic memory and planning. *Hippocampus*, *25*(10), 1073–1188. https://doi.org/10.1002/hipo.22488

Buzsáki, G., Leung, L. S., & Vanderwolf, C. H. (1983). Cellular bases of hippocampal EEG in the behaving rat. *Brain Research Reviews*, *6*(2), 139–171. https://doi.org/10.1016/0165-0173(83)90037-1

Cairney, S. A., Guttesen, A. á V., El Marj, N., & Staresina, B. P. (2018). Memory consolidation is linked to spindle-mediated information processing during sleep. *Current Biology*, *28*(6), 948–954.e4. https://doi.org/10.1016/j.cub.2018.01.087

Callaway, C. W., Lydic, R., Baghdoyan, H. A., & Hobson, J. A. (1987). Pontogeniculooccipital waves: Spontaneous visual system activity during rapid eye movement sleep. *Cellular and Molecular Neurobiology*, *7*(2), 105–149. https://doi.org/10.1007/BF00711551

Carskadon, M. A., & Dement, W. C. (1981). Cumulative effects of sleep restriction on daytime sleepiness. *Psychophysiology*, *18*(2), 107–113. https://doi.org/10.1111/j.1469-8986.1981.tb02921.x

Carskadon, M. A., & Herz, R. S. (2004). Minimal olfactory perception during sleep: why odor alarms will not work for humans. *SLEEP*, *27*(3), 402–405. https://doi.org/10.1093/sleep/27.3.402

Chernik, D. (1972). Effect of REM sleep deprivation on learning and recall by humans. *Perceptual and Motor Skills*, *34*(1), 283–294. https://doi.org/10.2466/pms.1972.34.1.283

Christensen, R., Pearson, L. M., & Johnson, W. (1992). Case-deletion diagnostics for mixed models. *Technometrics*, *34*(1), 38–45.

https://doi.org/10.1080/00401706.1992.10485231

Clemens, Z., Fabó, D., & Halász, P. (2005). Overnight verbal memory retention correlates with the number of sleep spindles. *Neuroscience*, *132*(2), 529–535. https://doi.org/10.1016/j.neuroscience.2005.01.011

Clemens, Z., Mölle, M., Eross, L., Barsi, P., Halász, P., & Born, J. (2007). Temporal coupling of parahippocampal ripples, sleep spindles and slow oscillations in humans. *Brain*, *130*(11), 2868–2878. https://doi.org/10.1093/brain/awm146

Cohen, D. A., Pascual-Leone, A., Press, D. Z., & Robertson, E. M. (2005). Off-line learning of motor skill memory: a double dissociation of goal and movement. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(50), 18237–18241. https://doi.org/10.1073/pnas.0506072102

Collingridge, G. L., Peineau, S., Howland, J. G., & Wang, Y. T. (2010). Long-term depression in the CNS. *Nature Reviews Neuroscience*, *11*(7), 459–473. https://doi.org/10.1038/nrn2867

Contreras, D., Destexhe, A., Sejnowski, T. J., & Steriade, M. (1996). Control of spatiotemporal coherence of a thalamic oscillation by corticothalamic feedback. *Science*, *274*(5288), 771–774. https://doi.org/10.1126/science.274.5288.771

Contreras, D., Destexhe, A., Sejnowski, T. J., & Steriade, M. (1997). Spatiotemporal patterns of spindle oscillations in cortex and thalamus. *Journal of Neuroscience*, *17*(3), 1179–1196. https://doi.org/10.1523/JNEUROSCI.17-03-01179.1997

Cook, R. D. (1977). Detection of influential observation in linear regression. *Technometrics*, *19*(1), 15–18. https://doi.org/10.1080/00401706.1977.10489493

Corsi-Cabrera, M., Rosales-Lagarde, A., del Río-Portilla, Y., Sifuentes-Ortega, R., & Alcántara-Quintero, B. (2015). Effects of selective REM sleep deprivation on

prefrontal gamma activity and executive functions. *International Journal of Psychophysiology*, *96*(2), 115–124. https://doi.org/10.1016/j.ijpsycho.2015.02.027

Cox, R., Hofman, W. F., de Boer, M., & Talamini, L. M. (2014). Local sleep spindle modulations in relation to specific memory cues. *Neuroimage*, *99*, 103–110. https://doi.org/10.1016/j.neuroimage.2014.05.028

Cox, R., Schapiro, A. C., Manoach, D. S., & Stickgold, R. (2017). Individual differences in frequency and topography of slow and fast sleep spindles. *Frontiers in Human Neuroscience*, *11*, 433. https://doi.org/10.3389/fnhum.2017.00433

Creery, J. D., Oudiette, D., Antony, J. W., & Paller, K. A. (2015). Targeted memory reactivation during sleep depends on prior learning. *SLEEP*, *38*(5), 755–763. https://doi.org/10.5665/sleep.4670

Crick, F., & Mitchison, G. (1983). The function of dream sleep. *Nature*, *304*(5922), 111–114. https://doi.org/10.1038/304111a0

Crunelli, V., & Hughes, S. W. (2010). The slow (< 1 Hz) rhythm of non-REM sleep: a dialogue between three cardinal oscillators. *Nature Neuroscience*, *13*(1), 9–17. https://doi.org/10.1038/nn.2445

Dang-Vu, T. T., Bonjean, M., Schabus, M., Boly, M., Darsaud, A., Desseilles, M., Degueldre, C., Balteau, E., Phillips, C., Luxen, A., Sejnowski, T. J., & Maquet, P. (2011). Interplay between spontaneous and induced brain activity during human non-rapid eye movement sleep. *Proceedings of the National Academy of Sciences*, *108*(37), 15438–15443. https://doi.org/10.1073/pnas.1112503108

Dang-Vu, T. T., McKinney, S. M., Buxton, O. M., Solet, J. M., & Ellenbogen, J. M. (2010). Spontaneous brain rhythms predict sleep stability in the face of noise. *Current Biology*, *20*(15), R626–R627. https://doi.org/10.1016/j.cub.2010.06.032

Datta, S. (1997). Cellular basis of pontine ponto-geniculo-occipital wave generation and modulation. *Cellular and Molecular Neurobiology*, *17*(3), 341–365. https://doi.org/10.1023/A:1026398402985

Datta, S. (2000). Avoidance task training potentiates phasic pontine-wave density in the rat: a mechanism for sleep-dependent plasticity. *Journal of Neuroscience*, *20*(22), 8607–8613. https://doi.org/10.1523/JNEUROSCI.20-22-08607.2000

Datta, S., Li, G., & Auerbach, S. (2008). Activation of phasic pontine-wave generator in the rat: a mechanism for expression of plasticity-related genes and proteins in the dorsal hippocampus and amygdala. *European Journal of Neuroscience*, *27*(7), 1876–1892. https://doi.org/10.1111/j.1460-9568.2008.06166.x

Datta, S., & O'Malley, M. W. (2013). Fear extinction memory consolidation requires potentiation of pontine-wave activity during REM sleep. *Journal of Neuroscience*, *33*(10), 4561–4569. https://doi.org/10.1523/JNEUROSCI.5525-12.2013

Datta, S., & Siwek, D. F. (1997). Excitation of the brain stem pedunculopontine tegmentum cholinergic cells induces wakefulness and REM sleep. *Journal of Neurophysiology*, *77*(6), 2975–2988. https://doi.org/10.1152/jn.1997.77.6.2975

De Koninck, J., Christ, G., Hebert, G., & Rinfret, N. (1990). Language learning efficiency, dreams and REM sleep. *Psychiatric Journal of the University of Ottawa*, *15*(2), 91–92.

De Koninck, J., Lorrain, D., Christ, G., Proulx, G., & Coulombe, D. (1989). Intensive language learning and increases in rapid eye movement sleep: evidence of a performance factor. *International Journal of Psychophysiology*, *8*(1), 43–47. https://doi.org/10.1016/0167-8760(89)90018-4

Debarnot, U., Creveaux, T., Collet, C., Doyon, J., & Guillot, A. (2009). Sleep

contribution to motor memory consolidation: a motor imagery study. *SLEEP*, *32*(12), 1559–1565. https://doi.org/10.1093/sleep/32.12.1559

Dement, W., & Kleitman, N. (1957). Cyclic variations in EEG during sleep and their relation to eye movements, body motility, and dreaming. *Electroencephalography and Clinical Neurophysiology*, *9*(4), 673–690. https://doi.org/10.1016/0013-4694(57)90088-3

Destrade, C., Hennevin, E., Leconte, P., & Soumireu-Mourat, B. (1978). Relationship between paradoxical sleep and time-dependent improvement of performance in BALB/c mice. *Neuroscience Letters*, *7*(2–3), 239–244. https://doi.org/10.1016/0304-3940(78)90175-1

Diekelmann, S., Biggel, S., Rasch, B., & Born, J. (2012). Offline consolidation of memory varies with time in slow wave sleep and can be accelerated by cuing memory reactivations. *Neurobiology of Learning and Memory*, *98*(2), 103–111. https://doi.org/10.1016/j.nlm.2012.07.002

Diekelmann, S., & Born, J. (2010). The memory function of sleep. *Nature Reviews Neuroscience*, *11*(2), 114–126. https://doi.org/10.1038/nrn2762

Diekelmann, S., Born, J., & Wagner, U. (2010). Sleep enhances false memories depending on general memory performance. *Behavioural Brain Research*, *208*(2), 425–429. https://doi.org/10.1016/j.bbr.2009.12.021

Diekelmann, S., Büchel, C., Born, J., & Rasch, B. (2011). Labile or stable: opposing consequences for memory when reactivated during waking and sleep. *Nature Neuroscience*, *14*, 381–386. https://doi.org/10.1038/nn.2744

Dresler, M., Kluge, M., Genzel, L., Schussler, P., & Steiger, A. (2010). Impaired off-line memory consolidation in depression. *European Neuropsychopharmacology*,

20(8), 553–561. https://doi.org/10.1016/j.euroneuro.2010.02.002

Drosopoulos, S., Schulze, C., Fischer, S., & Born, J. (2007). Sleep's function in the spontaneous recovery and consolidation of memories. *Journal of Experimental Psychology: General*, *136*(2), 169–183. https://doi.org/10.1037/0096-3445.136.2.169

Durán, E., Oyanedel, C. N., Niethard, N., Inostroza, M., & Born, J. (2018). Sleep stage dynamics in neocortex and hippocampus. *SLEEP*, *41*(6). zsy060. https://doi.org/10.1093/sleep/zsy060

Durrant, S. J., Cairney, S. A., & Lewis, P. A. (2013). Overnight consolidation aids the transfer of statistical knowledge from the medial temporal lobe to the striatum. *Cerebral Cortex*, *23*(10), 2467–2478. https://doi.org/10.1093/cercor/bhs244

Durrant, S. J., Cairney, S. A., McDermott, C., & Lewis, P. A. (2015). Schema-conformant memories are preferentially consolidated during REM sleep. *Neurobiology of Learning and Memory*, *122*, 41–50. https://doi.org/10.1016/j.nlm.2015.02.011

Durrant, S. J., Taylor, C., Cairney, S., & Lewis, P. A. (2011). Sleep-dependent consolidation of statistical learning. *Neuropsychologia*, *49*(5), 1322–1331. https://doi.org/10.1016/j.neuropsychologia.2011.02.015

Dusan-Peyrethon, D., Peyrethon, J., & Jouvet, M. (1967). [Quantitative study of phasic phenomena of paradoxal sleep during and after its instrumental deprivation]. *Comptes rendus des seances de la Societe de biologie et de ses filiales*, *161*(12), 2530–2533.

Ekstrand, B. R., Sullivan, M. J., Parker, D. F., & West, J. N. (1971). Spontaneous recovery and sleep. *Journal of Experimental Psychology*, *88*(1), 142–144. https://doi.org/10.1037/h0030642

Ellenbogen, J. M., Hu, P. T., Payne, J. D., Titone, D., & Walker, M. P. (2007). Human relational memory requires time and sleep. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(18), 7723–7728. https://doi.org/10.1073/pnas.0700094104

Ellenbogen, J. M., Hulbert, J. C., Stickgold, R., Dinges, D. F., & Thompson-Schill, S. L. (2006). Interfering with theories of sleep and memory: Sleep, declarative memory, and associative interference. *Current Biology*, *16*(13), 1290–1294. https://doi.org/10.1016/j.cub.2006.05.024

Empson, J., & Clarke, P. (1970). Rapid eye movements and remembering. *Nature*, *227*(5255), 287–288. https://doi.org/10.1038/227287a0

Eschenko, O., Ramadan, W., Mölle, M., Born, J., & Sara, S. J. (2008). Sustained increase in hippocampal sharp-wave ripple activity during slow-wave sleep after learning. *Learning & Memory*, *15*(4), 222–228. https://doi.org/10.1101/lm.726008

Fang, Z., Sergeeva, V., Ray, L. B., Viczko, J., Owen, A. M., & Fogel, S. M. (2017). Sleep spindles and intellectual ability: epiphenomenon or directly related? *Journal of Cognitive Neuroscience*, *29*(1), 167–182. https://doi.org/10.1162/jocn_a_01034

Fenn, K. M., Gallo, D. A., Margoliash, D., Roediger, H. L., & Nusbaum, H. C. (2009). Reduced false memory after sleep. *Learning & Memory*, *16*(9), 509–513. https://doi.org/10.1101/lm.1500808

Ferguson, J., & Dement, W. (1969). The behavioral effects of amphetamine on REM deprived rats. *Journal of Psychiatric Research*, *7*(2), 111–118. https://doi.org/10.1016/0022-3956(69)90016-8

Fernández-Mendoza, J., Lozano, B., Seijo, F., Santamarta-Liébana, E., José Ramos-Platón, M., Vela-Bueno, A., & Fernández-González, F. (2009). Evidence of

subthalamic PGO-like waves during REM sleep in humans: a deep brain polysomnographic study. *SLEEP*, *32*(9), 1117–1126. https://doi.org/10.1093/sleep/32.9.1117

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. https://doi.org/10.1016/S0166-4328(00)00177-7

Fischer, S., Hallschmid, M., Elsner, A. L., & Born, J. (2002). Sleep forms memory for finger skills. *Proceedings of the National Academy of Sciences*, *99*(18), 11987–11991. https://doi.org/10.1073/pnas.182178199

Fischer, S., Nitschke, M. F., Melchert, U. H., Erdmann, C., & Born, J. (2005). Motor memory consolidation in sleep shapes more effective neuronal representations. *Journal of Neuroscience*, *25*(49), 11248–11255. https://doi.org/10.1523/JNEUROSCI.1743-05.2005

Fogel, S. M., & Smith, C. T. (2006). Learning-dependent changes in sleep spindles and stage 2 sleep. *Journal of Sleep Research*, *15*(3), 250–255. https://doi.org/10.1111/j.1365-2869.2006.00522.x

Fogel, S., Smith, C., & Cote, K. (2007). Dissociable learning-dependent changes in REM and non-REM sleep in declarative and procedural memory systems. *Behavioural Brain Research*, *180*(1), 48–61. https://doi.org/10.1016/j.bbr.2007.02.037

Fougnie, D., Suchow, J. W., & Alvarez, G. A. (2012). Variability in the quality of visual working memory. *Nature Communications*, *3*, 1229. https://doi.org/10.1038/ncomms2237

Fox, J., & Weisberg, S. (2011). *An R companion to applied regression* (2nd ed.). Sage. http://socserv.socsci.mcmaster.ca/jfox/Books/Companion

Frey, S., & Frey, J. U. (2008). “Synaptic tagging” and “cross-tagging” and related associative reinforcement processes of functional plasticity as the cellular basis for memory formation. *Progress in Brain Research*, *169*, 117–143. https://doi.org/10.1016/S0079-6123(07)00007-6

Frey, U., Krug, M., Reymann, K. G., & Matthies, H. (1988). Anisomycin, an inhibitor of protein synthesis, blocks late phases of LTP phenomena in the hippocampal CA1 region in vitro. *Brain Research*, *452*(1–2), 57–65. https://doi.org/10.1016/0006-8993(88)90008-x

Gais, S., Albouy, G., Boly, M., Dang-Vu, T. T., Darsaud, A., Desseilles, M., Rauchs, G., Schabus, M., Sterpenich, V., Vandewalle, G., Maquet, P., & Peigneux, P. (2007). Sleep transforms the cerebral trace of declarative memories. *Proceedings of the National Academy of Sciences*, *104*(47), 18778–18783. http://www.pnas.org/content/104/47/18778.short

Gais, S., & Born, J. (2004). Declarative memory consolidation: mechanisms acting during human sleep. *Learning & Memory*, *11*(6), 679–685. https://doi.org/10.1101/lm.80504

Gais, S., Mölle, M., Helms, K., & Born, J. (2002). Learning-dependent increases in sleep spindle density. *Journal of Neuroscience*, *22*(15), 6830–6834. https://doi.org/20026697

Gais, S., Plihal, W., Wagner, U., & Born, J. (2000). Early sleep triggers memory for early visual discrimination skills. *Nature Neuroscience*, *3*(12), 1335. https://doi.org/10.1038/81881

Genzel, L., Spoormaker, V., Konrad, B., & Dresler, M. (2015). The role of rapid eye movement sleep for amygdala-related memory processing. *Neurobiology of*

Learning and Memory, *122*, 110–121. https://doi.org/10.1016/j.nlm.2015.01.008

Girardeau, G., Benchenane, K., Wiener, S. I., Buzsáki, G., & Zugaro, M. B. (2009). Selective suppression of hippocampal ripples impairs spatial memory. *Nature Neuroscience*, *12*(10), 1222–1223. https://doi.org/10.1038/nn.2384

Giuditta, A. (1985). A sequential hypothesis for the function of sleep. In W. P. Koella, E. Ruther, & H. Schulz (Eds.), *Sleep '84* (pp. 222–224). Fischer Verlag.

Giuditta, A. (2014). Sleep memory processing: the sequential hypothesis. *Frontiers in Systems Neuroscience*, *8*, 219. https://doi.org/10.3389/fnsys.2014.00219

Giuditta, A., Ambrosini, M. V., Montagnese, P., Mandile, P., Cotugno, M., Zucconi, G. G., & Vescia, S. (1995). The sequential hypothesis of the function of sleep. *Behavioural Brain Research*, *69*(1–2), 157–166. https://doi.org/10.1016/0166-4328(95)00012-I

Goerke, M., Cohrs, S., Rodenbeck, A., & Kunz, D. (2014). Differential effect of an anticholinergic antidepressant on sleep-dependent memory consolidation. *SLEEP*, *37*(5), 977–985. https://doi.org/10.5665/sleep.3674

Gorfine, T., Yeshurun, Y., & Zisapel, N. (2007). Nap and melatonin-induced changes in hippocampal activation and their role in verbal memory consolidation. *Journal of Pineal Research*, *43*(4), 336–342. https://doi.org/10.1111/j.1600-079X.2007.00482.x

Gott, J. A., Liley, D. T. J., & Hobson, J. A. (2017). Towards a functional understanding of PGO waves. *Frontiers in Human Neuroscience*, *11*, 89. https://doi.org/10.3389/fnhum.2017.00089

Gramfort, A., Luessi, M., Larson, E., Engemann, D., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L., & Hämäläinen, M. (2013). MEG and EEG

data analysis with MNE-Python. *Frontiers in Neuroscience*, *7*, 267. https://doi.org/10.3389/fnins.2013.00267

Grieser, C., Greenberg, R., & Harrison, R. H. (1972). The adaptive function of sleep: the differential effects of sleep and dreaming on recall. *Journal of Abnormal Psychology*, *80*(3), 280–286. https://doi.org/10.1037/h0033641

Groch, S., Zinke, K., Wilhelm, I., & Born, J. (2015). Dissociating the contributions of slow-wave sleep and rapid eye movement sleep to emotional item and source memory. *Neurobiology of Learning and Memory*, *122*, 122–130. https://doi.org/10.1016/j.nlm.2014.08.013

Grosmark, A. D., Mizuseki, K., Pastalkova, E., Diba, K., & Buzsáki, G. (2012). REM sleep reorganizes hippocampal excitability. *Neuron*, *75*(6), 1001–1007. https://doi.org/10.1016/j.neuron.2012.08.015

Grosvenor, A., & Lack, L. C. (1984). The effect of sleep before or after learning on memory. *SLEEP*, *7*(2), 155–167. https://doi.org/10.1093/sleep/7.2.155

Hasselmo, M. E., & McGaughy, J. (2004). High acetylcholine levels set circuit dynamics for attention and encoding and low acetylcholine levels set dynamics for consolidation. *Progress in Brain Research, 145*, 207–231. https://doi.org/10.1016/S0079-6123(03)45015-2

Hebb, D. O. (1949). *The organization of behavior* (Vol. 65). Wiley.

Hocking, J., Dzafic, I., Kazovsky, M., & Copland, D. A. (2013). NESSTI: norms for environmental sound stimuli. *PLOS ONE, 8*(9), e73382. https://doi.org/10.1371/journal.pone.0073382

Hoedlmoser, K., Birklbauer, J., Schabus, M., Eibenberger, P., Rigler, S., & Mueller, E. (2015). The impact of diurnal sleep on the consolidation of a complex gross motor

adaptation task. *Journal of Sleep Research*, *24*(1), 100–109. https://doi.org/10.1111/jsr.12207

Hoffman, K., & McNaughton, B. (2002). Coordinated reactivation of distributed memory traces in primate neocortex. *Science*, *297*(5589), 2070–2073. https://doi.org/10.1126/science.1073538

Hölscher, C., Anwyl, R., & Rowan, M. J. (1997). Stimulation on the positive phase of hippocampal theta rhythm induces long-term potentiation that can be depotentiated by stimulation on the negative phase in area CA1 in vivo. *Journal of Neuroscience*, *17*(16), 6470–6477. https://doi.org/10.1523/JNEUROSCI.17-16-06470.1997

Holz, J., Piosczyk, H., Feige, B., Spiegelhalder, K., Baglioni, C., Riemann, D., & Nissen, C. (2012). EEG sigma and slow-wave activity during NREM sleep correlate with overnight declarative and procedural memory consolidation: EEG sigma and SWA and memory consolidation. *Journal of Sleep Research*, *21*(6), 612–619. https://doi.org/10.1111/j.1365-2869.2012.01017.x

Hornung, O. P., Regen, F., Danker-Hopfe, H., Schredl, M., & Heuser, I. (2007). The relationship between REM sleep and memory consolidation in old age and effects of cholinergic medication. *Biological Psychiatry*, *61*(6), 750–757. https://doi.org/10.1016/j.biopsych.2006.08.034

Huber, R., Ghilardi, M. F., Massimini, M., Ferrarelli, F., Riedner, B. A., Peterson, M. J., & Tononi, G. (2006). Arm immobilization causes cortical plastic changes and locally decreases sleep slow wave activity. *Nature Neuroscience*, *9*(9), 1169–1176. https://doi.org/10.1038/nn1758

Huber, R., Ghilardi, M. F., Massimini, M., & Tononi, G. (2004). Local sleep and

learning. *Nature*, *430*(6995), 78–81. https://doi.org/10.1038/nature02663

Huerta, P. T., & Lisman, J. E. (1995). Bidirectional synaptic plasticity induced by a single burst during cholinergic theta oscillation in CA1 in vitro. *Neuron*, *15*(5), 1053–1063. https://doi.org/10.1016/0896-6273(95)90094-2

Hutchison, I. C., & Rathore, S. (2015). The role of REM sleep theta activity in emotional memory. *Frontiers in Psychology*, *6*, 1439. https://doi.org/10.3389/fpsyg.2015.01439

Ishihara, S. (2014). *Ishihara's tests for colour deficiency (Concise)*. Kanehara Trading Inc.

Jahnke, K., von Wegner, F., Morzelewski, A., Borisov, S., Maischein, M., Steinmetz, H., & Laufs, H. (2012). To wake or not to wake? The two-sided nature of the human K-complex. *NeuroImage*, *59*(2), 1631–1638. https://doi.org/10.1016/j.neuroimage.2011.09.013

Jasper, H. H., & Tessier, J. (1971). Acetylcholine liberation from cerebral cortex during paradoxical (REM) sleep. *Science*, *172*(3983), 601–602. https://doi.org/10.1126/science.172.3983.601

Jenkins, J. G., & Dallenbach, K. M. (1924). Obliviscence during sleep and waking. *The American Journal of Psychology*, *35*(4), 605–612. http://www.jstor.org/stable/1414040

Ji, D., & Wilson, M. (2007). Coordinated memory replay in the visual cortex and hippocampus during sleep. *Nature Neuroscience*, *10*(1), 100–107. https://doi.org/10.1038/nn1825

Johns, M. W. (1991). A new method for measuring daytime sleepiness: The Epworth sleepiness scale. *SLEEP*, *14*(6), 540–545. https://doi.org/10.1093/sleep/14.6.540

Jouvet, M. (1962). Recherches sur les structures nerveuses et les mécanismes responsables des différentes phases du sommeil physiologique. *Archives Italiennes de Biologie, 100*(2), 125–206. https://doi.org/10.4449/aib.v100i2.1761

Kametani, H., & Kawamura, H. (1990). Alterations in acetylcholine release in the rat hippocampus during sleep-wakefulness detected by intracerebral dialysis. *Life Sciences, 47*(5), 421–426. https://doi.org/10.1016/0024-3205(90)90300-G

Karashima, A., Katayama, N., & Nakao, M. (2007). Phase-locking of spontaneous and tone-elicited pontine waves to hippocampal theta waves during REM sleep in rats. *Brain Research, 1182,* 73–81. https://doi.org/10.1016/j.brainres.2007.08.060

Karashima, A., Nakamura, K., Watanabe, M., Sato, N., Nakao, M., Katayama, N., & Yamamoto, M. (2001). Synchronization between hippocampal theta waves and PGO waves during REM sleep. *Psychiatry and Clinical Neurosciences, 55*(3), 189–190. https://doi.org/10.1046/j.1440-1819.2001.00820.x

Karni, A., Tanne, D., Rubenstein, B. S., Askenasy, J., J. M., & Sagi, D. (1994). Dependence on REM sleep of overnight improvement of a perceptual skill. *Science, 265*(5172), 679–682. https://doi.org/10.1126/science.8036518

Katsuki, H., Izumi, Y., & Zorumski, C. F. (1997). Noradrenergic regulation of synaptic plasticity in the hippocampal CA1 region. *Journal of Neurophysiology, 77*(6), 3013–3020. https://doi.org/10.1152/jn.1997.77.6.3013

Kenward, M. G., & Roger, J. H. (1997). Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics, 53*(3), 983–997. https://doi.org/10.2307/2533558

Kirkwood, A., Rozas, C., Kirkwood, J., Perez, F., & Bear, M. F. (1999). Modulation of long-term synaptic depression in visual cortex by acetylcholine and

norepinephrine. *Journal of Neuroscience*, *19*(5), 1599–1609. https://doi.org/10.1523/JNEUROSCI.19-05-01599.1999

Korman, M., Doyon, J., Doljansky, J., Carrier, J., Dagan, Y., & Karni, A. (2007). Daytime sleep condenses the time course of motor memory consolidation. *Nature Neuroscience*, *10*(9), 1206–1213. https://doi.org/10.1038/nn1959

Krug, M., Lossner, B., & Ott, T. (1984). Anisomycin blocks the late phase of long-term potentiation in the dentate gyrus of freely moving rats. *Brain Research Bulletin*, *13*(1), 39–42. https://doi.org/10.1016/0361-9230(84)90005-4

Krystal, A. D., Edinger, J. D., Wohlgemuth, W. K., & Marsh, G. R. (2002). NREM sleep EEG frequency spectral correlates of sleep complaints in primary insomnia subtypes. *SLEEP*, *25*(6), 626–636. https://doi.org/10.1093/sleep/25.6.626

Lahl, O., Wispel, C., Willigens, B., & Pietrowsky, R. (2008). An ultra short episode of sleep is sufficient to promote declarative memory performance. *Journal of Sleep Research*, *17*(1), 3–10. https://doi.org/10.1111/j.1365-2869.2008.00622.x

Landmann, N., Kuhn, M., Piosczyk, H., Feige, B., Baglioni, C., Spiegelhalder, K., Frase, L., Riemann, D., Sterr, A., & Nissen, C. (2014). The reorganisation of memory during sleep. *Sleep Medicine Reviews*, *18*(6), 531–541. https://doi.org/10.1016/j.smrv.2014.03.005

Landsness, E. C., Crupi, D., Hulse, B. K., Peterson, M. J., Huber, R., Ansari, H., Coen, M., Cirelli, C., Benca, R. M., Ghilardi, M. F., & Tononi, G. (2009). Sleep-dependent improvement in visuomotor learning: a causal role for slow waves. *SLEEP*, *32*(10), 1273–1284. https://doi.org/10.1093/sleep/32.10.1273

Lansink, C. S., Goltstein, P. M., Lankelma, J. V., Joosten, R. N. J. M. A., McNaughton, B. L., & Pennartz, C. M. A. (2008). Preferential reactivation of motivationally

relevant information in the ventral striatum. *Journal of Neuroscience*, *28*(25), 6372–6382. https://doi.org/10.1523/JNEUROSCI.1054-08.2008

Lapierre, J. L., Kosenko, P. O., Lyamin, O. I., Kodama, T., Mukhametov, L. M., & Siegel, J. M. (2007). Cortical acetylcholine release is lateralized during asymmetrical slow-wave sleep in northern fur seals. *Journal of Neuroscience*, *27*(44), 11999–12006. https://doi.org/10.1523/JNEUROSCI.2968-07.2007

Lau, H., Tucker, M. A., & Fishbein, W. (2010). Daytime napping: Effects on human direct associative and relational memory. *Neurobiology of Learning and Memory*, *93*(4), 554–560. https://doi.org/10.1016/j.nlm.2010.02.003

Laventure, S., Fogel, S., Lungu, O., Albouy, G., Sévigny-Dupont, P., Vien, C., Sayour, C., Carrier, J., Benali, H., & Doyon, J. (2016). NREM2 and sleep spindles are instrumental to the consolidation of motor sequence memories. *PLOS ONE*, *14*(3), e1002429. https://doi.org/10.1371/journal.pbio.1002429

Lechner, H. A., Squire, L. R., & Byrne, J. H. (1999). 100 years of consolidation—remembering Müller and Pilzecker. *Learning & Memory*, *6*(2), 77–87. https://doi.org/10.1101/lm.6.2.77

Lee, M., Chrobak, J., Sik, A., Wiley, R., & Buzsáki, G. (1994). Hippocampal theta activity following selective lesion of the septal cholinergic system. *Neuroscience*, *62*(4), 1033–1047. https://doi.org/10.1016/0306-4522(94)90341-7

Lega, B. C., Jacobs, J., & Kahana, M. (2012). Human hippocampal theta oscillations and the formation of episodic memories. *Hippocampus*, *22*(4), 748–761. https://doi.org/10.1002/hipo.20937

Legault, G., Smith, C. T., & Beninger, R. J. (2004). Scopolamine during the paradoxical sleep window impairs radial arm maze learning in rats. *Pharmacology*,

Biochemistry, and Behavior, *79*(4), 715–721. https://doi.org/10.1016/j.pbb.2004.09.018

Legault, G., Smith, C. T., & Beninger, R. J. (2006). Post-training intra-striatal scopolamine or flupenthixol impairs radial maze learning in rats. *Behavioural Brain Research*, *170*(1), 148–155. https://doi.org/10.1016/j.bbr.2006.02.010

Lena, I., Parrot, S., Deschaux, O., Muffat-Joly, S., Sauvinet, V., Renaud, B., Suaud-Chagny, M.-F., & Gottesmann, C. (2005). Variations in extracellular levels of dopamine, noradrenaline, glutamate, and aspartate across the sleep--wake cycle in the medial prefrontal cortex and nucleus accumbens of freely moving rats. *Journal of Neuroscience Research*, *81*(6), 891–899. https://doi.org/10.1002/jnr.20602

Leonard, B. J., McNaughton, B. L., & Barnes, C. A. (1987). Suppression of hippocampal synaptic plasticity during slow-wave sleep. *Brain Research*, *425*(1), 174–177. https://doi.org/10.1016/0006-8993(87)90496-3

Lesting, J., Narayanan, R. T., Kluge, C., Sangha, S., Seidenbecher, T., & Pape, H.-C. (2011). Patterns of coupled theta activity in amygdala-hippocampal-prefrontal cortical circuits during fear extinction. *PLOS ONE*, *6*(6), e21714. https://doi.org/10.1371/journal.pone.0021714

Lewin, I., & Glaubman, H. (1975). The effect of REM deprivation: is it detrimental, beneficial, or neutral? *Psychophysiology*, *12*(3), 349–353. https://doi.org/10.1111/j.1469-8986.1975.tb01303.x

Li, W., Ma, L., Yang, G., & Gan, W.-B. (2017). REM sleep selectively prunes and maintains new synapses in development and learning. *Nature Neuroscience*, *20*(3), 427–437. https://doi.org/10.1038/nn.4479

Lim, A. S., Lozano, A. M., Moro, E., Hamani, C., Hutchison, W. D., Dostrovsky, J. O.,

Lang, A. E., Wennberg, R. A., & Murray, B. J. (2007). Characterization of REM-sleep associated ponto-geniculo-occipital waves in the human pons. *SLEEP*, *30*(7), 823–827. https://doi.org/10.1093/sleep/30.7.823

Lopes-Aguiar, C., Romcy-Pereira, R. N., Escorsim Szawka, R., Galvis-Alonso, O. Y., Anselmo-Franci, J. A., & Pereira Leite, J. (2008). Muscarinic acetylcholine neurotransmission enhances the late-phase of long-term potentiation in the hippocampal-prefrontal cortex pathway of rats in vivo: a possible involvement of monoaminergic systems. *Neuroscience*, *153*(4), 1309–1319. https://doi.org/10.1016/j.neuroscience.2008.02.040

Louie, K., & Wilson, M. A. (2001). Temporally structured replay of awake hippocampal ensemble activity during rapid eye movement sleep. *Neuron*, *29*(1), 145–156. https://doi.org/10.1016/S0896-6273(01)00186-6

Loy, A., & Hofmann, H. (2014). HLMdiag: A suite of diagnostics for hierarchical linear models in R. *Journal of Statistical Software, Articles*, *56*(5), 1–28. https://doi.org/10.18637/jss.v056.i05

Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature Neuroscience*, *17*(3), 347–356. https://doi.org/10.1038/nn.3655

MacDonald, K. J. (2020, February 22). *Testing a hypothesis of non-REM sleep reinforcement and REM sleep refinement for the benefits of post-learning sleep on memory retrieval*. https://doi.org/10.17605/OSF.IO/9XDHK

MacDonald, K. J., & Cote, K. A. (2016). Sleep physiology predicts memory retention after reactivation. *Journal of Sleep Research*, *25*(6), 655–663. https://doi.org/10.1111/jsr.12423

MacDonald, K. J., & Schirmeister, J. (2020, February 22). *Matched image and sound clip*

stimuli. https://doi.org/10.17605/OSF.IO/6YJWX

Maingret, N., Girardeau, G., Todorova, R., Goutierre, M., & Zugaro, M. (2016). Hippocampo-cortical coupling mediates memory consolidation during sleep. *Nature Neuroscience*, *19*(7), 959–964. https://doi.org/10.1038/nn.4304

Mandai, O., Guerrien, A., Sockeel, P., Dujardin, K., & Leconte, P. (1989). REM sleep modifications following a Morse code learning session in humans. *Physiology & Behavior*, *46*(4), 639–642. https://doi.org/10.1016/0031-9384(89)90344-2

Marrosu, F., Portas, C., Mascia, M. S., Casu, M. A., Fà, M., Giagheddu, M., Imperato, A., & Gessa, G. L. (1995). Microdialysis measurement of cortical and hippocampal acetylcholine release during sleep-wake cycle in freely moving cats. *Brain Research*, *671*(2), 329–332. https://doi.org/10.1016/0006-8993(94)01399-3

Mazzoni, G., Gori, S., Formicola, G., Gneri, C., Massetani, R., Murri, L., & Salzarulo, P. (1999). Word recall correlates with sleep cycles in elderly subjects. *Journal of Sleep Research*, *8*(3), 185–188. https://doi.org/10.1046/j.1365-2869.1999.00154.x

McCarley, R. W. (1981). Pontine brainstem neuronal activity and REM sleep control mechanisms. In J. Szentágothai, M. Palkovits, & J. Hámori (Eds.), *Regulatory Functions of the CNS Principles of Motion and Organization* (pp. 293–299). Pergamon. https://doi.org/10.1016/B978-0-08-026814-9.50044-2

McCarley, R. W., Winkelman, J. W., & Duffy, F. H. (1983). Human cerebral potentials associated with REM sleep rapid eye movements: links to PGO waves and waking potentials. *Brain Research*, *274*(2), 359–364. https://doi.org/10.1016/0006-8993(83)90719-9

McCormick, D. A. (1989). Cholinergic and noradrenergic modulation of thalamocortical processing. *Trends in Neurosciences*, *12*(6), 215–221.

https://doi.org/10.1016/0166-2236(89)90125-2

McKinney, S. M., Dang-Vu, T. T., Buxton, O. M., Solet, J. M., & Ellenbogen, J. M. (2011). Covert waking brain activity reveals instantaneous sleep depth. *PLOS ONE*, *6*(3), e17351. https://doi.org/10.1371/journal.pone.0017351

Mednick, S., Nakayama, K., & Stickgold, R. (2003). Sleep-dependent learning: a nap is as good as a night. *Nature Neuroscience*, *6*(7), 697–698. https://doi.org/10.1038/nn1078

Meier-Koll, A., Bussmann, B., Schmidt, C., & Neuschwander, D. (1999). Walking through a maze alters the architecture of sleep. *Perceptual and Motor Skills*, *88*(3), 1141–1159. https://doi.org/10.2466/pms.1999.88.3c.1141

Merica, H., Blois, R., & Gaillard, J. M. (1998). Spectral characteristics of sleep EEG in chronic insomnia. *European Journal of Neuroscience*, *10*(5), 1826–1834. https://doi.org/10.1046/j.1460-9568.1998.00189.x

Mölle, M., Bergmann, T. O., Marshall, L., & Born, J. (2011). Fast and slow spindles during the sleep slow oscillation: disparate coalescence and engagement in memory processing. *SLEEP*, *34*(10), 1411–1421. https://doi.org/10.5665/SLEEP.1290

Mölle, M., Eschenko, O., Gais, S., Sara, S. J., & Born, J. (2009). The influence of learning on sleep slow oscillations and associated spindles and ripples in humans and rats. *European Journal of Neuroscience*, *29*(5), 1071–1081. https://doi.org/10.1111/j.1460-9568.2009.06654.x

Mölle, M., Marshall, L., Gais, S., & Born, J. (2002). Grouping of spindle activity during slow oscillations in human non-rapid eye movement sleep. *Journal of Neuroscience*, *22*(24), 10941–10947. https://doi.org/10.1523/JNEUROSCI.22-24-

10941.2002

Mölle, M., Marshall, L., Gais, S., & Born, J. (2004). Learning increases human electroencephalographic coherence during subsequent slow sleep oscillations. *Proceedings of the National Academy of Sciences*, *101*(38), 13963–13968. https://doi.org/10.1073/pnas.0402820101

Morin, A., Doyon, J., Dostie, V., Barakat, M., Hadj Tahar, A., Korman, M., Benali, H., Karni, A., Ungerleider, L. G., & Carrier, J. (2008). Motor sequence learning increases sleep spindles and fast frequencies in post-training sleep. *SLEEP*, *31*(8), 1149–1156. https://doi.org/10.5665/sleep/31.8.1149

Moroni, F., Nobili, L., Curcio, G., De Carli, F., Fratello, F., Marzano, C., De Gennaro, L., Ferrillo, F., Cossu, M., Francione, S., Lo Russo, G., Bertini, M., & Ferrara, M. (2007). Sleep in the human hippocampus: a stereo-EEG study. *PLOS ONE*, *2*(9), e867. https://doi.org/10.1371/journal.pone.0000867

Moscovitch, M., Rosenbaum, R. S., Gilboa, A., Addis, D. R., Westmacott, R., Grady, C., McAndrews, M. P., Levine, B., Black, S., Winocur, G., & Nadel, L. (2005). Functional neuroanatomy of remote episodic, semantic and spatial memory: a unified account based on multiple trace theory. *Journal of Anatomy*, *207*(1), 35–66. https://doi.org/10.1111/j.1469-7580.2005.00421.x

Müller, G. E., & Pilzecker, A. (1900). *Experimentelle beiträge zur lehre vom gedächtniss* (Vol. 1). JA Barth.

Nádasdy, Z., Hirase, H., Czurkó, A., Csicsvari, J., & Buzsáki, G. (1999). Replay and time compression of recurring spike sequences in the hippocampus. *Journal of Neuroscience*, *19*(21), 9497–9507. https://doi.org/10.1523/JNEUROSCI.19-21-09497.1999

Nadel, L., Samsonovich, A., Ryan, L., & Moscovitch, M. (2000). Multiple trace theory of human memory: computational, neuroimaging, and neuropsychological results. *Hippocampus*, *10*(4), 352–368. https://doi.org/10.1002/1098-1063(2000)10:4%3C352::AID-HIPO2%3E3.0.CO;2-D

Nakanishi, H., Sun, Y., Nakamura, R. K., Mori, K., Ito, M., Suda, S., Namba, H., Storch, F. I., Dang, T. P., Mendelson, W., Mishkin, M., Kennedy, C., Gillin, J. C., Smith, C. B., & Sokoloff, L. (1997). Positive correlations between cerebral protein synthesis rates and deep sleep in Macaca mulatta. *The European Journal of Neuroscience*, *9*(2), 271–279. https://doi.org/10.1111/j.1460-9568.1997.tb01397.x

Nelson, J. P., McCarley, R. W., & Hobson, J. A. (1983). REM sleep burst neurons, PGO waves, and eye movement information. *Journal of Neurophysiology*, *50*(4), 784–797. https://doi.org/10.1152/jn.1983.50.4.784

Ngo, H.-V. V., Martinetz, T., Born, J., & Mölle, M. (2013). Auditory closed-loop stimulation of the sleep slow oscillation enhances memory. *Neuron*, *78*(3), 545–553. https://doi.org/10.1016/j.neuron.2013.03.006

Nielsen, T. A. (2000). A review of mentation in REM and NREM sleep: "Covert" REM sleep as a possible reconciliation of two opposing models. *Behavioral and Brain Sciences*, *23*(6), 851–866. https://doi.org/10.1017/S0140525X0000399X

Nishida, M., & Walker, M. P. (2007). Daytime naps, motor memory consolidation and regionally specific sleep spindles. *PLOS ONE*, *2*(4), e341. https://doi.org/10.1371/journal.pone.0000341

Nowacka, A., Jurkowlaniec, E., & Trojniar, W. (2002). Microinjection of procaine into the pedunculopontine tegmental nucleus suppresses hippocampal theta rhythm in urethane-anesthetized rats. *Brain Research Bulletin*, *58*(4), 377–384.

https://doi.org/10.1016/s0361-9230(02)00801-8

O'Dell, T. J., Connor, S. A., Guglietta, R., & Nguyen, P. V. (2015). β-Adrenergic receptor signaling and modulation of long-term potentiation in the mammalian hippocampus. *Learning & Memory*, *22*(9), 461–471. https://doi.org/10.1101/lm.031088.113

O'Neill, J., Senior, T. J., Allen, K., Huxter, J. R., & Csicsvari, J. (2008). Reactivation of experience-dependent cell assembly patterns in the hippocampus. *Nature Neuroscience*, *11*(2), 209–215. https://doi.org/10.1038/nn2037

Orban, P., Rauchs, G., Balteau, E., Degueldre, C., Luxen, A., Maquet, P., & Peigneux, P. (2006). Sleep after spatial learning promotes covert reorganization of brain activity. *Proceedings of the National Academy of Sciences*, *103*(18), 7124–7129. https://doi.org/10.1073/pnas.0510198103

Otani, S., Marshall, C., Tate, W., Goddard, G., & Abraham, W. (1989). Maintenance of long-term potentiation in rat dentate gyrus requires protein synthesis but not messenger RNA synthesis immediately post-tetanization. *Neuroscience*, *28*(3), 519–526. https://doi.org/10.1016/0306-4522(89)90001-8

Oudiette, D., Antony, J. W., Creery, J. D., & Paller, K. A. (2013). The role of memory reactivation during wakefulness and sleep in determining which memories endure. *Journal of Neuroscience*, *33*(15), 6672–6678. https://doi.org/10.1523/JNEUROSCI.5497-12.2013

Oudiette, D., & Paller, K. A. (2013). Upgrading the sleeping brain with targeted memory reactivation. *Trends in Cognitive Sciences*, *17*(3), 142–149. https://doi.org/10.1016/j.tics.2013.01.006

Pavlides, C., Greenstein, Y. J., Grudman, M., & Winson, J. (1988). Long-term

potentiation in the dentate gyrus is induced preferentially on the positive phase of θ-rhythm. *Brain Research, 439*(1–2), 383–387. https://doi.org/10.1016/0006-8993(88)91499-0

Payne, J. D., Schacter, D. L., Propper, R. E., Huang, L.-W., Wamsley, E. J., Tucker, M. A., Walker, M. P., & Stickgold, R. (2009). The role of sleep in false memory formation. *Neurobiology of Learning and Memory, 92*(3), 327–334. https://doi.org/10.1016/j.nlm.2009.03.007

Pearlman, C. A. (1969). Effect of rapid eye movement (dreaming) sleep deprivation on retention of avoidance learning in rats. *US Naval Submarine Medical Research Lab,* Report No. 563. http://archive.rubicon-foundation.org/8608

Peigneux, P., Laureys, S., Delbeuck, X., & Maquet, P. (2001). Sleeping brain, learning brain. The role of sleep for memory systems. *Neuroreport, 12*(18), A111–A124. https://doi.org/10.1097/00001756-200112210-00001

Peigneux, P., Laureys, S., Fuchs, S., Collette, F., Perrin, F., Reggers, J., Phillips, C., Degueldre, C., Del Fiore, G., Aerts, J., Luxen, A., & Maquet, P. (2004). Are spatial memories strengthened in the human hippocampus during slow wave sleep? *Neuron, 44*(3), 535–545. https://doi.org/10.1016/j.neuron.2004.10.007

Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods, 162*(1), 8–13. https://doi.org/10.1016/j.jneumeth.2006.11.017

Peters, K. R., Ray, L., Smith, V., & Smith, C. (2008). Changes in the density of stage 2 sleep spindles following motor learning in young and older adults. *Journal of Sleep Research, 17*(1), 23–33. https://doi.org/10.1111/j.1365-2869.2008.00634.x

Pignatelli, M., Beyeler, A., & Leinekugel, X. (2012). Neural circuits underlying the

generation of theta oscillations. *Journal of Physiology-Paris*, *106*(3–4), 81–92. https://doi.org/10.1016/j.jphysparis.2011.09.007

Pivik, R. T., Broughton, R. J., Coppola, R., Davidson, R. J., Fox, N., & Nuwer, M. R. (1993). Guidelines for the recording and quantitative analysis of electroencephalographic activity in research contexts. *Psychophysiology*, *30*(6), 547–558. https://doi.org/10.1111/j.1469-8986.1993.tb02081.x

Plihal, W., & Born, J. (1997). Effects of early and late nocturnal sleep on declarative and procedural memory. *Journal of Cognitive Neuroscience*, *9*(4), 534–547. https://doi.org/10.1162/jocn.1997.9.4.534

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

Poe, G. R. (2017). Sleep is for forgetting. *Journal of Neuroscience*, *37*(3), 464–473. https://doi.org/10.1523/JNEUROSCI.0820-16.2017

Poe, G. R., Nitz, D. A., McNaughton, B. L., & Barnes, C. A. (2000). Experience-dependent phase-reversal of hippocampal neuron firing during REM sleep. *Brain Research*, *855*(1), 176–180. https://doi.org/10.1016/S0006-8993(99)02310-0

Poe, G. R., Walsh, C. M., & Bjorness, T. E. (2010). Cognitive neuroscience of sleep. *Progress in Brain Research*, *185*, 1–19. https://doi.org/10.1016/B978-0-444-53702-7.00001-4

Popa, D., Duvarci, S., Popescu, A. T., Lena, C., & Pare, D. (2010). Coherent amygdalocortical theta promotes fear memory consolidation during paradoxical sleep. *Proceedings of the National Academy of Sciences*, *107*(14), 6516–6519. https://doi.org/10.1073/pnas.0913016107

Qin, Y.-L., McNaughton, B. L., Skaggs, W. E., & Barnes, C. A. (1997). Memory reprocessing in corticocortical and hippocampocortical neuronal ensembles. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *352*(1360), 1525–1533. https://doi.org/10.1098/rstb.1997.0139

R Core Team. (2017). *R: A Language and environment for statistical computing*. R Foundation for Statistical Computing. http://www.R-project.org/

Ramm, P., & Smith, C. T. (1990). Rates of cerebral protein synthesis are linked to slow wave sleep in the rat. *Physiology & Behavior*, *48*(5), 749–753. https://doi.org/10.1016/0031-9384(90)90220-X

Rasch, B., & Born, J. (2013). About sleep's role in memory. *Physiological Reviews*, *93*(2), 681–766. https://doi.org/10.1152/physrev.00032.2012

Rasch, B., & Born, J. (2015). In search of a role of REM sleep in memory formation. *Neurobiology of Learning and Memory*, *122*, 1–3. https://doi.org/10.1016/j.nlm.2015.04.012

Rasch, B., Büchel, C., Gais, S., & Born, J. (2007). Odor cues during slow-wave sleep prompt declarative memory consolidation. *Science*, *315*(5817), 1426–1429. https://doi.org/10.1126/science.1138581

Rasch, B., Gais, S., & Born, J. (2009). Impaired off-line consolidation of motor memories after combined blockade of cholinergic receptors during REM sleep-rich sleep. *Neuropsychopharmacology*, *34*(7), 1843–1853. https://doi.org/10.1038/npp.2009.6

Rechtschaffen, A., & Kales, A. (1968). *A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects*. U.S. Department of Health, Education and Welfare.

Redondo, R. L., & Morris, R. G. M. (2011). Making memories last: The synaptic tagging and capture hypothesis. *Nature Reviews Neuroscience*, *12*(1), 17–30. https://doi.org/10.1038/nrn2963

Ribeiro, S., Goyal, V., Mello, C. V., & Pavlides, C. (1999). Brain gene expression during REM sleep depends on prior waking experience. *Learning & Memory*, *6*(5), 500–508. https://doi.org/10.1101/lm.6.5.500

Ribeiro, S., Mello, C. V., Velho, T., Gardner, T. J., Jarvis, E. D., & Pavlides, C. (2002). Induction of hippocampal long-term potentiation during waking leads to increased extrahippocampal zif-268 expression during ensuing rapid-eye-movement sleep. *Journal of Neuroscience*, *22*(24), 10914–10923. https://doi.org/10.1523/JNEUROSCI.22-24-10914.2002

Ribeiro, S., Shi, X., Engelhard, M., Zhou, Y., Zhang, H., Gervasoni, D., Lin, S.-C., Wada, K., Lemos, N. A. M., & Nicolelis, M. A. L. (2007). Novel experience induces persistent sleep-dependent plasticity in the cortex but not in the hippocampus. *Frontiers in Neuroscience*, *1*(1), 43–55. https://doi.org/10.3389/neuro.01.1.1.003.2007

Riemann, D., Spiegelhalder, K., Feige, B., Voderholzer, U., Berger, M., Perlis, M., & Nissen, C. (2010). The hyperarousal model of insomnia: a review of the concept and its evidence. *Sleep Medicine Reviews*, *14*(1), 19–31. https://doi.org/10.1016/j.smrv.2009.04.002

Rihm, J. S., Diekelmann, S., Born, J., & Rasch, B. (2014). Reactivating memories during sleep by odors: Odor specificity and associated changes in sleep oscillations. *Journal of Cognitive Neuroscience*, *26*(8), 1806–1818. https://doi.org/10.1162/jocn_a_00579

Robertson, E. M., Pascual-Leone, A., & Miall, R. C. (2004). Current concepts in procedural consolidation. *Nature Reviews Neuroscience*, *5*(7), 576–582. https://doi.org/10.1038/nrn1426

Ruch, S., Markes, O., Duss, S. B., Oppliger, D., Reber, T. P., Koenig, T., Mathis, J., Roth, C., & Henke, K. (2012). Sleep stage II contributes to the consolidation of declarative memories. *Neuropsychologia*, *50*(10), 2389–2396. https://doi.org/10.1016/j.neuropsychologia.2012.06.008

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

Runyan, J. D., Moore, A. N., & Dash, P. K. (2019). Coordinating what we've learned about memory consolidation: Revisiting a unified theory. *Neuroscience & Biobehavioral Reviews*. https://doi.org/10.1016/j.neubiorev.2019.02.010

Sakai, K., Sastre, J., Salvert, D., Touret, M., Tohyama, M., & Jouvet, M. (1979). Tegmentoreticular projections with special reference to the muscular atonia during paradoxical sleep in the cat: An HRP study. *Brain Research*, *176*(2), 233–254. https://doi.org/10.1016/0006-8993(79)90981-8

Salzarulo, P., Lairy, G., Bancaud, J., & Munari, C. (1975). Direct depth recording of the striate cortex during REM sleep in man: Are there PGO potentials? *Electroencephalography and Clinical Neurophysiology*, *38*(2), 199–202. https://doi.org/10.1016/0013-4694(75)90231-X

Saper, C. B., Scammell, T. E., & Lu, J. (2005). Hypothalamic regulation of sleep and circadian rhythms. *Nature*, *437*(7063), 1257–1263. https://doi.org/10.1038/nature04284

Schabus, M., Dang-Vu, T. T., Albouy, G., Balteau, E., Boly, M., Carrier, J., Darsaud, A., Degueldre, C., Desseilles, M., Gais, S., & others. (2007). Hemodynamic cerebral correlates of sleep spindles during human non-rapid eye movement sleep. *Proceedings of the National Academy of Sciences*, *104*(32), 13164–13169. https://doi.org/10.1073/pnas.0703084104

Schabus, M., Gruber, G., Parapatics, S., Sauter, C., Klösch, G., Anderer, P., Klimesch, W., Saletu, B., & Zeitlhofer, J. (2004). Sleep spindles and their significance for declarative memory consolidation. *SLEEP*, *27*(8), 1479–1485. https://doi.org/10.1093/sleep/27.7.1479

Schabus, M., Hödlmoser, K., Pecherstorfer, T., & Klösch, G. (2005). Influence of midday naps on declarative memory performance and motivation. *Somnologie-Schlafforschung Und Schlafmedizin*, *9*(3), 148–153. https://doi.org/10.1111/j.1439-054X.2005.00054.x

Schönauer, M., Alizadeh, S., Jamalabadi, H., Abraham, A., Pawlizki, A., & Gais, S. (2017). Decoding material-specific memory reprocessing during sleep in humans. *Nature Communications*, *8*, 15404. https://doi.org/10.1038/ncomms15404

Schönauer, M., Geisler, T., & Gais, S. (2014). Strengthening procedural memories by reactivation in sleep. *Journal of Cognitive Neuroscience*, *26*(1), 143–153. https://doi.org/10.1162/jocn_a_00471

Schouten, D. I., Pereira, S. I. R., Tops, M., & Louzada, F. M. (2017). State of the art on targeted memory reactivation: Sleep your way to enhanced cognition. *Sleep Medicine Reviews*, *32*, 123–131. https://doi.org/10.1016/j.smrv.2016.04.002

Schreiner, T., Göldi, M., & Rasch, B. (2015). Cueing vocabulary during sleep increases theta activity during later recognition testing: Cueing during sleep and subsequent

theta activity. *Psychophysiology*, *52*(11), 1538–1543. https://doi.org/10.1111/psyp.12505

Schreiner, T., Lehmann, M., & Rasch, B. (2015). Auditory feedback blocks memory benefits of cueing during sleep. *Nature Communications*, *6*, 8729. https://doi.org/10.1038/ncomms9729

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery, and Psychiatry*, *20*(1), 11. https://doi.org/10.1136/jnnp.20.1.11

Seibt, J., & Frank, M. G. (2019). Primed to sleep: The dynamics of synaptic plasticity across brain states. *Frontiers in Systems Neuroscience*, *13*, 2. https://doi.org/10.3389/fnsys.2019.00002

Shaffery, J. P., Roffwarg, H. P., Speciale, S. G., & Marks, G. A. (1999). Ponto-geniculo-occipital-wave suppression amplifies lateral geniculate nucleus cell-size changes in monocularly deprived kittens. *Developmental Brain Research*, *114*(1), 109–119. https://doi.org/10.1016/s0165-3806(99)00027-9

Siapas, A. G., Lubenov, E. V., & Wilson, M. A. (2005). Prefrontal phase locking to hippocampal theta oscillations. *Neuron*, *46*(1), 141–151. https://doi.org/10.1016/j.neuron.2005.02.028

Siegel, J. M. (2001). The REM sleep-memory consolidation hypothesis. *Science*, *294*(5544), 1058–1063. https://doi.org/10.1126/science.1063049

Sirota, A., Csicsvari, J., Buhl, D., & Buzsáki, G. (2003). Communication between neocortex and hippocampus during sleep in rodents. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(4), 2065–2069. https://doi.org/10.1073/pnas.0437938100

Smith, C. (1985). Sleep states and learning: a review of the animal literature. *Neuroscience & Biobehavioral Reviews*, *9*(2), 157–168. https://doi.org/10.1016/0149-7634(85)90042-9

Smith, C. (1995). Sleep states and memory processes. *Behavioural Brain Research*, *69*(1–2), 137–145. https://doi.org/10.1016/0166-4328(95)00024-n

Smith, C., & Lapp, L. (1991). Increases in number of REMS and REM density in humans following an intensive learning period. *SLEEP*, *14*(4), 325–330. https://doi.org/10.1093/sleep/14.4.325

Smith, C., & MacNeill, C. (1994). Impaired motor memory for a pursuit rotor task following Stage 2 sleep loss in college students. *Journal of Sleep Research*, *3*(4), 206–213. https://doi.org/10.1111/j.1365-2869.1994.tb00133.x

Smith, C. T., Aubrey, J. B., & Peters, K. R. (2004). Different roles for REM and stage 2 sleep in motor learning: A proposed model. *Psychologica Belgica*, *44*, 81–104.

Smith, C., Young, J., & Young, W. (1980). Prolonged increases in paradoxical sleep during and after avoidance-task acquisition. *SLEEP*, *3*(1), 67–81. https://doi.org/10.1093/sleep/3.1.67

Squire, L. R., Genzel, L., Wixted, J. T., & Morris, R. G. (2015). Memory consolidation. *Cold Spring Harbor Perspectives in Biology*, *7*(8), a021766. https://doi.org/10.1101/cshperspect.a021766

Squire, L. R., Knowlton, B., & Musen, G. (1993). The structure and organization of memory. *Annual Review of Psychology*, *44*(1), 453–495. https://doi.org/10.1146/annurev.ps.44.020193.002321

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

Steriade, M. (2003). The corticothalamic system in sleep. *Frontiers in Bioscience*, *8*, d878–899. https://doi.org/10.2741/1043

Steriade, M., Deschenes, M., Domich, L., & Mulle, C. (1985). Abolition of spindle oscillations in thalamic neurons disconnected from nucleus reticularis thalami. *Journal of Neurophysiology*, *54*(6), 1473–1497. https://doi.org/10.1152/jn.1985.54.6.1473

Steriade, M., Domich, L., Oakson, G., & Deschenes, M. (1987). The deafferented reticular thalamic nucleus generates spindle rhythmicity. *Journal of Neurophysiology*, *57*(1), 260–273. https://doi.org/10.1152/jn.1987.57.1.260

Steriade, M., Nuñez, A., & Amzica, F. (1993). Intracellular analysis of relations between the slow (< 1 Hz) neocortical oscillation and other sleep rhythms of the electroencephalogram. *Journal of Neuroscience*, *13*(8), 3266–3283. https://doi.org/10.1523/JNEUROSCI.13-08-03266.1993

Sterman, M., & Hoppenbrouwers, T. (1971). The development of sleep-waking and rest-activity patterns from fetus to adult in man. In *Brain development and behavior* (pp. 203–227). Academic Press.

Sterpenich, V., Albouy, G., Boly, M., Vandewalle, G., Darsaud, A., Balteau, E., Dang-Vu, T. T., Desseilles, M., D'Argembeau, A., Gais, S., Rauchs, G., Schabus, M., Degueldre, C., Luxen, A., Collette, F., & Maquet, P. (2007). Sleep-related hippocampo-cortical interplay during emotional memory recollection. *PLoS Biology*, *5*(11), e282. https://doi.org/10.1371/journal.pbio.0050282

Stickgold, R. (1998). Sleep: off-line memory reprocessing. *Trends in Cognitive Sciences*,

2(12), 484–492. https://doi.org/10.1016/s1364-6613(98)01258-3

Stickgold, R., James, L., & Hobson, J. A. (2000). Visual discrimination learning requires sleep after training. *Nature Neuroscience*, *3*(12), 1237–1238. https://doi.org/10.1038/81756

Stickgold, R., & Walker, M. P. (2013). Sleep-dependent memory triage: evolving generalization through selective processing. *Nature Neuroscience*, *16*(2), 139–145. https://doi.org/10.1038/nn.3303

Stickgold, R., Whidbee, D., Schirmer, B., Patel, V., & Hobson, J. A. (2000). Visual discrimination task improvement: A multi-step process occurring during sleep. *Journal of Cognitive Neuroscience*, *12*(2), 246–254. https://doi.org/10.1162/089892900562075

Suchow, J. W., Brady, T. F., Fougnie, D., & Alvarez, G. A. (2013). Modeling visual working memory with the MemToolbox. *Journal of Vision*, *13*(10), 9. https://doi.org/10.1167/13.10.9

Takashima, A., Petersson, K. M., Rutters, F., Tendolkar, I., Jensen, O., Zwarts, M. J., McNaughton, B. L., & Fernandez, G. (2006). Declarative memory consolidation in humans: a prospective functional magnetic resonance imaging study. *Proceedings of the National Academy of Sciences*, *103*(3), 756–761. https://doi.org/10.1073/pnas.0507774103

Talamini, L. M., Nieuwenhuis, I. L. C., Takashima, A., & Jensen, O. (2008). Sleep directly following learning benefits consolidation of spatial associative memory. *Learning & Memory*, *15*(4), 233–237. https://doi.org/10.1101/lm.771608

Tamaki, M., Huang, T.-R., Yotsumoto, Y., Hamalainen, M., Lin, F.-H., Nanez, J. E., Watanabe, T., & Sasaki, Y. (2013). Enhanced spontaneous oscillations in the

supplementary motor area are associated with sleep-dependent offline learning of finger-tapping motor-sequence task. *Journal of Neuroscience, 33*(34), 13894–13902. https://doi.org/10.1523/JNEUROSCI.1198-13.2013

Tamaki, M., Matsuoka, T., Nittono, H., & Hori, T. (2008). Fast sleep spindle (13–15 Hz) activity correlates with sleep-dependent improvement in visuomotor performance. *SLEEP*, *31*(2), 204–211. https://doi.org/10.1093/sleep/31.2.204

Tamaki, M., Matsuoka, T., Nittono, H., & Hori, T. (2009). Activation of fast sleep spindles at the premotor cortex and parietal areas contributes to motor learning: a study using sLORETA. *Clinical Neurophysiology*, *120*(5), 878–886. https://doi.org/10.1016/j.clinph.2009.03.006

Teber, I., Köhling, R., Speckmann, E.-J., Barnekow, A., & Kremerskothen, J. (2004). Muscarinic acetylcholine receptor stimulation induces expression of the activity-regulated cytoskeleton-associated gene (ARC). *Molecular Brain Research, 121*, 131–136. https://doi.org/10.1016/j.molbrainres.2003.11.017

Tempesta, D., Socci, V., De Gennaro, L., & Ferrara, M. (2018). Sleep and emotional processing. *Sleep Medicine Reviews, 40*, 183–195. https://doi.org/10.1016/j.smrv.2017.12.005

Thomas, M. J., Moody, T. D., Makhinson, M., & O'Dell, T. J. (1996). Activity-dependent β-adrenergic modulation of low frequency stimulation induced LTP in the hippocampal CA1 region. *Neuron, 17*(3), 475–482. https://doi.org/10.1016/s0896-6273(00)80179-8

Tilley, A. J., Brown, S., Donald, M., Ferguson, S., Piccone, J., Plasto, K., & Statham, D. (1992). Human sleep and memory processes. In R. J. Broughton, & R. D. Ogilvie (Eds.), *Sleep, arousal, and performance: A tribute to Bob Wilkinson* (pp. 117–

127). Birkhäuser.

Tilley, A. J., & Empson, J. A. (1978). REM sleep and memory consolidation. *Biological Psychology*, *6*(4), 293–300. https://doi.org/10.1016/0301-0511(78)90031-5

Timofeev, I., Grenier, F., Bazhenov, M., Sejnowski, T., & Steriade, M. (2000). Origin of slow cortical oscillations in deafferented cortical slabs. *Cerebral Cortex*, *10*(12), 1185–1199. https://doi.org/10.1093/cercor/10.12.1185

Tononi, G., & Cirelli, C. (2003). Sleep and synaptic homeostasis: a hypothesis. *Brain Research Bulletin*, *62*(2), 143–150. https://doi.org/10.1016/j.brainresbull.2003.09.004

Tononi, G., & Cirelli, C. (2006). Sleep function and synaptic homeostasis. *Sleep Medicine Reviews*, *10*(1), 49–62. https://doi.org/10.1016/j.smrv.2005.05.002

Tucker, M. A., & Fishbein, W. (2008). Enhancement of declarative memory performance following a daytime nap is contingent on strength of initial task acquisition. *SLEEP*, *31*(2), 197–203. https://doi.org/10.1093/sleep/31.2.197

Ulloor, J., & Datta, S. (2005). Spatio-temporal activation of cyclic AMP response element-binding protein, activity-regulated cytoskeletal-associated protein and brain-derived nerve growth factor: a mechanism for pontine-wave generator activation-dependent two-way active-avoidance memory processing in the rat. *Journal of Neurochemistry*, *95*(2), 418–428. https://doi.org/10.1111/j.1471-4159.2005.03378.x

van den Berg, R., Shin, H., Chou, W.-C., George, R., & Ma, W. J. (2012). Variability in encoding precision accounts for visual short-term memory limitations. *Proceedings of the National Academy of Sciences*, *109*(22), 8780–8785. https://doi.org/10.1073/pnas.1117465109

Van Der Werf, Y. D., Van Der Helm, E., Schoonheim, M. M., Ridderikhoff, A., & Van Someren, E. J. W. (2009). Learning by observation requires an early sleep window. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(45), 18926–18930. https://doi.org/10.1073/pnas.0901320106

Van Dort, C. J., Zachs, D. P., Kenny, J. D., Zheng, S., Goldblum, R. R., Gelwan, N. A., Ramos, D. M., Nolan, M. A., Wang, K., Weng, F.-J., Lin, Y., Wilson, M. A., & Brown, E. N. (2015). Optogenetic activation of cholinergic neurons in the PPT or LDT induces REM sleep. *Proceedings of the National Academy of Sciences*, *112*(2), 584–589. https://doi.org/10.1073/pnas.1423136112

Vanni-Mercier, G., & Debilly, G. (1998). A key role for the caudoventral pontine tegmentum in the simultaneous generation of eye saccades in bursts and associated ponto-geniculo-occipital waves during paradoxical sleep in the cat. *Neuroscience*, *86*(2), 571–585. https://doi.org/10.1016/S0306-4522(98)00045-1

Vazquez, J., Hall, S. C., Witkowska, H. E., & Greco, M. A. (2008). Rapid alterations in cortical protein profiles underlie spontaneous sleep and wake bouts. *Journal of Cellular Biochemistry*, *105*(6), 1472–1484. https://doi.org/10.1002/jcb.21970

Verschoor, G. J., & Holdstock, T. L. (1984). REM bursts and REM sleep following visual and auditory learning. *South African Journal of Psychology*, *14*(3), 69–74. https://doi.org/10.1177/008124638401400301

Vertes, R. P., & Eastman, K. E. (2000). The case against memory consolidation in REM sleep. *Behavioral and Brain Sciences*, *23*(6), 867–876. https://doi.org/10.1017/S0140525X00004003

Vertes, R. P., & Siegel, J. M. (2005). Time for the sleep community to take a critical look at the purported role of sleep in memory processing. *SLEEP*, *28*(10), 1228–1229.

https://doi.org/10.1093/sleep/28.10.1228

von der Kammer, H., Mayhaus, M., Albrecht, C., Enderich, J., Wegner, M., & Nitsch, R. M. (1998). Muscarinic acetylcholine receptors activate expression of the EGR gene family of transcription factors. *Journal of Biological Chemistry*, *273*(23), 14538–14544. https://doi.org/10.1074/jbc.273.23.14538

Wagner, U., Gais, S., & Born, J. (2001). Emotional memory formation is enhanced across sleep intervals with high amounts of rapid eye movement sleep. *Learning & Memory*, *8*, 112–119. https://doi.org/10.1101/lm.36801

Wagner, U., Gais, S., Haider, H., Verleger, R., & Born, J. (2004). Sleep inspires insight. *Nature*, *427*(6972), 352–355. https://doi.org/10.1038/nature02223

Walker, M. P., Brakefield, T., Hobson, J. A., & Stickgold, R. (2003). Dissociable stages of human memory consolidation and reconsolidation. *Nature*, *425*(6958), 616–620. https://doi.org/10.1038/nature01930

Walker, M. P., Brakefield, T., Morgan, A., Hobson, J. A., & Stickgold, R. (2002). Practice with sleep makes perfect: Sleep-dependent motor skill learning. *Neuron*, *35*(1), 205–211. https://doi.org/10.1016/S0896-6273(02)00746-8

Walker, M. P., Stickgold, R., Alsop, D., Gaab, N., & Schlaug, G. (2005). Sleep-dependent motor memory plasticity in the human brain. *Neuroscience*, *133*(4), 911–917. https://doi.org/10.1016/j.neuroscience.2005.04.007

Watts, A., Gritton, H. J., Sweigart, J., & Poe, G. R. (2012). Antidepressant suppression of non-REM sleep spindles and REM sleep impairs hippocampus-dependent learning while augmenting striatum-dependent learning. *Journal of Neuroscience*, *32*(39), 13411–13420. https://doi.org/10.1523/JNEUROSCI.0170-12.2012

Webb, W. B., & Agnew, H. W. J. (1975). The effects on subsequent sleep of an acute

restriction of sleep length. *Psychophysiology*, *12*(4), 367–370. https://doi.org/10.1111/j.1469-8986.1975.tb00002.x

Welch, B. L. (1951). On the comparison of several mean values: An alternative approach. *Biometrika*, *38*(3/4), 330–336. https://doi.org/10.2307/2332579

Wickham, H. (2007). Reshaping data with the reshape package. *Journal of Statistical Software, Articles*, *21*(12), 1–20. https://doi.org/10.18637/jss.v021.i12

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag. http://ggplot2.org

Wierzynski, C. M., Lubenov, E. V., Gu, M., & Siapas, A. G. (2009). State-dependent spike-timing relationships between hippocampal and prefrontal circuits during sleep. *Neuron*, *61*(4), 587–596. https://doi.org/10.1016/j.neuron.2009.01.011

Wilhelm, I., Diekelmann, S., & Born, J. (2008). Sleep in children improves memory performance on declarative but not procedural tasks. *Learning & Memory*, *15*(5), 373–377. https://doi.org/10.1101/lm.803708

Wilhelm, I., Prehn-Kristensen, A., & Born, J. (2012). Sleep-dependent memory consolidation–what can be learnt from children? *Neuroscience & Biobehavioral Reviews*, *36*(7), 1718–1728. https://doi.org/10.1016/j.neubiorev.2012.03.002

Williams, R. L., Agnew, H. W., & Webb, W. B. (1964). Sleep patterns in young adults: An EEG study. *Electroencephalography & Clinical Neurophysiology*. https://doi.org/10.1016/0013-4694(64)90160-9

Williams, R. L., Agnew Jr, H. W., & Webb, W. B. (1966). Sleep patterns in the young adult female: an EEG study. *Electroencephalography and Clinical Neurophysiology*, *20*(3), 264–266. https://doi.org/10.1016/0013-4694(66)90092-7

Wilson, M. A., & McNaughton, B. L. (1994). Reactivation of hippocampal ensemble

memories during sleep. *Science*, *265*(5172), 676–679. https://doi.org/10.1126/science.8036517

Witt, K., Margraf, N., Bieber, C., Born, J., & Deuschl, G. (2010). Sleep consolidates the effector-independent representation of a motor skill. *Neuroscience*, *171*(1), 227–234. https://doi.org/10.1016/j.neuroscience.2010.07.062

Yang, G., Lai, C. S. W., Cichon, J., Ma, L., Li, W., & Gan, W.-B. (2014). Sleep promotes branch-specific formation of dendritic spines after learning. *Science*, *344*(6188), 1173–1178. https://doi.org/10.1126/science.1249098

Yang, H.-W., Lin, Y.-W., Yen, C.-D., & Min, M.-Y. (2002). Change in bi-directional plasticity at CA1 synapses in hippocampal slices taken from 6-hydroxydopamine-treated rats: the role of endogenous norepinephrine. *European Journal of Neuroscience*, *16*(6), 1117–1128. https://doi.org/10.1046/j.1460-9568.2002.02165.x

Yaroush, R., Sullivan, M. J., & Ekstrand, B. R. (1971). Effect of sleep on memory: II. Differential effect of the first and second half of the night. *Journal of Experimental Psychology*, *88*(3), 361–366. https://doi.org/10.1037/h0030914

Yoder, R. M., & Pang, K. C. (2005). Involvement of GABAergic and cholinergic medial septal neurons in hippocampal theta rhythm. *Hippocampus*, *15*(3), 381–392. https://doi.org/10.1002/hipo.20062

Yordanova, J., Kolev, V., Wagner, U., Born, J., & Verleger, R. (2012). Increased alpha (8–12 Hz) activity during slow wave sleep as a marker for the transition from implicit knowledge to explicit insight. *Journal of Cognitive Neuroscience*, *24*(1), 119–132. https://doi.org/10.1162/jocn_a_00097

Yoshitake, H. (1978). Three characteristic patterns of subjective fatigue symptoms.

Ergonomics, *21*(3), 231–233. https://doi.org/10.1080/00140137808931718

Zhang, H., Fell, J., & Axmacher, N. (2018). Electrophysiological mechanisms of human memory consolidation. *Nature Communications*, *9*, 4103. https://doi.org/10.1038/s41467-018-06553-y

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, *453*(7192), 233–235. https://doi.org/10.1038/nature06860

Zimmerman, W. B. (1970). Sleep mentation and auditory awakening thresholds. *Psychophysiology*, *6*(5), 540–549. https://doi.org/10.1111/j.1469-8986.1970.tb02243.x