Adolescent Sleep:

Learning, Memory, and Mental Wellbeing

Jessica Amy March

A thesis for the degree of Doctor of Philosophy (PhD)

Submitted to the Department of Psychology

Royal Holloway University of London

9th June 2023

Corrected version: Monday 16th October 2023

Acknowledgements

Thank you to my supervisors, Dr Jakke Tamminen and Professor Jessie Ricketts, who showed unwavering support throughout my PhD and supported me in developing my career. I am very grateful for all your guidance, support, and advice. I would also like to thank the other PhD students at the time that let me sleep on their sofas when I was collecting data and commuting, and who provided me a good sense of community and support. I am also incredibly grateful to all the participants, schools, and families that took part in my research. I would also like to thank our undergraduate research assistants for their support with data collection, collation, and transcription for Chapter 3: Haniya Zaka, Emily Todd, Mollie Brown, and Selan Lee. Thank you to Dr Johnny Downs and other members of the CAMHS Digital Lab who provided advice and support in managing the data used in Chapter 6. Thank you also to Dr Laurence Telesia who supported with inter-rater validation and in obtaining final analyses and figures once my access to the dataset expired for Chapter 6.

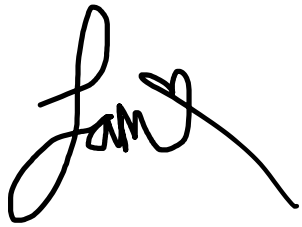
I would like to acknowledge the funding for my PhD from the Waterloo Foundation and Royal Holloway University of London. Clinical Records Interactive Search (CRIS) provided the secondary data for the study in Chapter 6. CRIS is supported by the National Institute of Health Research (NIHR) Biomedical research Centre for Mental Health (BRC) Nucleus, which is located at the South London and Maudsley NHS Foundation Trust within the Institute of Psychiatry, Psychology and Neuroscience (IoPPN) at King's College London (KCL). The BRC Nucleus is funded by the Guy's and St Thomas' (GSST) Trustees and the South London and Maudsley Trustees. Access to this data set was supported by the CAMHS Digital Lab, chaired by Dr Johnny Downs, at the IoPPN at KCL.

Thank you to my granddad who bought my first Psychology textbooks when I was an undergraduate and didn't mind me doing work and sleeping on his sofa when I visited him, I miss you so much. I also want to thank my amazing husband, Liam March, who supported

me through all the highs and lows of this journey and kept the house from falling into complete disarray in my chaos! I really appreciate that you have been by my side throughout this process, keeping me on track. Thank you to my mum who also supported me throughout this journey and dealt with all the blood, sweat, and tears over this time. I would like to thank Liam and my mum for both allowing me to practice sticking electrodes to their heads! I also appreciate all my friends and family who supported me in so many different ways, including my dad, step-mum, siblings, nan, niece and nephews. I would also like to thank the teams I worked with before and during my PhD that inspired my interests in development, sleep, and cognition: SEND team at Honywood Secondary School, St. Aubyn's Centre, Thorpe Lea Primary School, NELFT CAMHS Eating Disorder Team, and North Essex CAMHS. Thank you also to my undergraduate lecturers and teachers through GCSE and A-Level who sparked my interest in Psychology and supported me to apply to university when I did not consider this path as an option for me.

Authorship Declaration

I, Jessica Amy March, hereby declare that this thesis and the work presented in it is entirely my own. Where I have consulted the work of others or sections have been co-authored, this is always clearly stated



Please note, Chapter 3 is based on a registered report that has been published and co-authored by my supervisors (Dr Jakke Tamminen and Professor Jessie Ricketts), with suggestions made by anonymous reviewers at Stage 1 and Stage 2 of the registered report process. For the published registered report, please see: March, J. A., Ricketts, J., Tamminen, J. (2023). Is word learning capacity restored after a daytime nap? *CORTEX*, https://doi.org/10.1016/j.cortex.2022.10.013. Chapter 3 overlaps substantially with the published report but has been amended to fit the thesis format and narrative.

Abstract

This thesis explores the relationship between adolescent sleep, learning, memory, affect, and mental health. In Chapter 2 (*N* = 11), multi-night sleep recordings showed 13-16-year-olds were sleep restricted during school nights compared to free nights. At-home polysomnography found slow-wave sleep was relatively preserved despite sleep restriction. In Chapter 3 (*N* = 45), spindles during a pre-encoding nap were positively correlated with overnight improvements in lexical integration. Chapters 2 and 3 did not find slow-wave sleep to be associated with behavioural performance, despite predictions made by the synaptic homeostasis hypothesis (Tononi & Cirelli, 2003, 2006, 2014). In relation to affect and mental health, two online studies were conducted. Chapter 4 (*N* = 247) found good sleepers on the PSQI had higher neutral than positive memory, whilst the opposite was seen for poorer sleepers at the one-year follow-up, not during lockdowns. Chapter 5 found reactivity toward, or recognition of, emotional stimuli did not significantly mediate the relationship between sleep and mental health, contrary to Goldstein and Walker's (2014) framework, but was mediated by lower thought control ability supporting Harrington & Cairney's (2021) model. High positive affect mediated the relationship between better sleep and lower depression for the younger (*N* = 118) but not older adolescents (*N* = 136). Lastly, in Chapter 6 (*N* = 7,167) NHS CAMHS patients with an ADHD diagnosis were significantly more likely to be prescribed sleep medication than patients with most other mental health diagnoses (except psychosis and autism or learning disability). Longer service use was also associated with greater reporting of sleep issues and prescription of sleep medication. Together, this thesis demonstrates the importance of sleep for adolescents' learning, memory, and mental wellbeing. This is particularly important, as some adolescents may need sleep interventions to benefit their education (e.g., GCSEs, A-Levels, University) and mental health. However, Chapter 6 suggests some interventions (i.e., sleep medication) may not be accessed equally.

Table of Contents

Glossary

Term	Definition
Actigraphy	A wrist-worn device used to objectively measure sleep based primarily on movement
Adolescence	10 – 24 years old (Sawyer et al., 2019)
Affect	Feelings including emotions and moods
Amygdala	The brain area that detects and reacts to emotional stimuli
Chronotype	Biological peaks in cognition and a preference to complete activities at a particular time-of-day associated with circadian rhythms
Circadian rhythms	Internal biological clocks (or 'body-clocks')
Emotions	Transient forms of affect
Hippocampus	The brain area where (declarative) memories are initially encoded
Learning	The acquisition of new information (e.g., through reward, exposure)
Memory	The ability to recall information that has been learnt
Moods	Longer lasting forms of affect
Negative affect	Feelings such as fear, anger, sadness
Polysomnography	EEG worn during sleep (including EOG and EMG) to detect sleep stages (e.g., REM, slow-wave sleep) and sleep architecture (e.g., spindles, slow-wave oscillations)
Positive affect	Feelings such as joy, enthusiasm, and motivation
Pre-frontal cortex	The brain area associated with decision making and inhibitory control
REM sleep	Rapid eye movement sleep. Previously known as paradoxical sleep, as EEG recordings resemble wakefulness

Restricted sleep	Sleep of an insufficient duration (see sufficient sleep definition)
Sleep deprivation	Total lack of sleep (usually for >24 hour period)
Slow-wave sleep	A deep sleep stage characterised by slow-wave oscillations
Spindles	Synchronised bursts of neuronal activity seen during slow-wave and Stage 2 sleep
Sufficient sleep duration	Based on the National Sleep Foundation recommendations (Hirshkowitz et al., 2015); 8-10 hours during adolescence

List of Tables

List of Figures

Chapter 1: General introduction and literature review

1.1. Introduction

Humans spend around one-third of their lives asleep (Raven et al., 2018). However, sleep can be viewed as evolutionarily disadvantageous, increasing vulnerability to predation, for example. Yet sleep is common across animal species; thus it must serve an important function (Rasch & Born, 2013). Sleep is theorised to be important for allowing memories to be consolidated (discussed in Section 1.2.1.) and preparing the brain for new learning (Section 1.2.2.). Sleep disturbances are a common factor across a range of psychiatric disorders (Baglioni et al., 2016) suggesting an inherent link between sleep and mental wellbeing (Section 1.3.). The relationship between sleep, affective wellbeing and mental health may be explained by control over emotional thoughts, memories, and responses to affective stimuli (e.g., Goldstein & Walker, 2014; Harrington & Cairney, 2021). The relationship between sleep, learning, memory, and mental health may be particularly relevant to adolescents (Section 1.4.) and may have implication for education and Child and Adolescent Mental Health Services (CAMHS). Education and mental health are becoming increasingly more intertwined. Recent government green papers have put a greater emphasis on mental health provisions within schools (Department of Health & Department for Education, 2017). Additionally, the relationship between sleep and schools has become a topical focus, with petitions asking secondary schools to delay their start times to 10 a.m., which were debated in parliament (Petitions Committee, 2019; Zeichner, 2019). These petitions are also developmentally relevant, as adolescence is a key time of biological change, with subsequent influences on sleep (Crowley et al., 2018). However, adolescence is often understudied in the sleep, learning and memory literature, with a large majority of research exploring adulthood or childhood. According to a non-systematic review, adolescence was

identified as a critical time for mental health vulnerability, with 75% of mental-disorders being present prior to the age of 24-years-old and prodromal symptoms emerging earlier (Fusar-Poli, 2019). Whilst the review by Fusar-Poli (2019) is methodologically limited in its non-systematic design, due to the large topic area, it does identify and discuss multiple systematic reviews and large sample studies. Together, the literature suggests that adolescence may be an important time for both sleep changes and development of mental health difficulties. Therefore, this Chapter introduces key concepts relating to the association between sleep, learning and memory (Section 1.2.), affect and mental health (Section 1.3.) and the consequences of adolescent development and changes in sleep for these relationships (Section 1.4.).

1.2. Theories of sleep, learning, and memory

Learning and memory are fundamental to academic success and other daily functions. Learning refers to the acquisition of new information, whereas memory is the ability to access or recall information after it has been acquired. Sleep may support learning and memory in two different, but complementary ways, consolidation of memories for information learned prior to sleep (Section 1.2.1.) and restoring the brain for subsequent learning (Section 1.2.2.). Sleep is considered an optimal state for memory processes (Born & Wilhelm, 2012; Rasch & Born, 2013).

Sleep entails cycles of rapid-eye movement (REM) and non-REM sleep. Non-REM sleep is made up of deep (i.e., slow-wave sleep) and shallower sleep stages (Stage 1 and Stage 2). Sleep typically starts with a short period of Stage 1 sleep, moving into Stage 2, slow-wave sleep, and finally REM sleep. Theories of the relationship between sleep and declarative memory focus on non-REM sleep; mostly slow-wave sleep and sleep spindles. Sleep spindles are bursts of synchronised neuronal activity that have been associated with various aspects of

cognition such as working memory, speed, accuracy, and executive functions (Reynolds et al., 2018). Sleep spindles are seen in both Stage 2 and slow-wave sleep. Slow-wave sleep typically follows Stage 2 sleep and dominates the first half of the night (Diekelmann & Born, 2010). Slow-wave sleep is a deep sleep architectural stage, in which large peaks and troughs in the EEG recording are seen. The peaks represent synchronised neuronal firing, and the troughs are due to synchronised inhibition (Rasch & Born, 2013). The duration and intensity of slow-wave sleep (e.g., frequency and magnitude of slow-wave oscillations) is considered to reflect the amount of learning throughout the day, as a function of homeostatic mechanisms thought to occur during slow-wave sleep (e.g., Olcese et al., 2010; Tononi & Cirelli, 2003).

Two key theories explore the benefits of sleep for learning and memory, the synaptic homeostasis hypothesis (Tononi & Cirelli, 2003, 2014), and systems consolidation models (e.g., Born & Wilhelm, 2012). Both theories primarily address the role of sleep for declarative memories rather than procedural memories. Declarative memories can be recalled such as episodic (e.g., factual information, events) or semantic (e.g., abstract knowledge such as grammar, general knowledge) memories whereas procedural memories relate to performing skills (e.g., playing an instrument, riding a bicycle). The synaptic homeostasis hypothesis focuses on the synaptic potentiation during wakefulness (proliferation of synapses) and downscaling during sleep (trimming of synaptic connections) to maintain homeostasis within the brain, allowing both the consolidation of memories (Section 1.2.1.) and the restoration of encoding capacity (Section 1.2.2.). The active systems consolidation account focuses on memory consolidation via communication between brain areas and the reorganisation of information from temporary storage to long-term memory storage. These two theories may not be mutually exclusive; for example, both theories emphasise the role of slow-wave sleep in learning and memory, and the synaptic homeostasis hypothesis may

explain the neuronal processes occurring when the hippocampus and neocortical regions communicate as described in systems consolidation accounts. The idea of both cellular and system accounts to explain the consolidation of declarative memory has also been suggested by other researchers (e.g., Genzel & Wixted, 2017). Both the synaptic homeostasis hypothesis and active systems consolidation frameworks may only be applicable to hippocampal, or declarative, memories, further supporting the idea that they may discuss complementary mechanisms.

1.2.1. Sleep after learning (consolidation)

The majority of sleep research focuses on consolidating memories for information learned before sleeping. The synaptic homeostasis hypothesis, also known as the synaptic downscaling hypothesis, proposes a model that aims to explain the function of sleep for consolidation as well as restoring encoding capacity (Tononi & Cirelli, 2003, 2006, 2014). Here, the synaptic homeostasis hypothesis is discussed in the context of memory consolidation. Section 1.2.2. discusses how this model explains the restoration of encoding capacity following sleep. The synaptic homeostasis hypothesis suggests that learning occurs via synaptic potentiation (the formation of synaptic connections) during wakefulness. In this context, learning refers to all waking interactions with the environment which include learning associations from rewards and punishments, performing tasks, and learning from the environment. Further, increases in synaptic strength during wakefulness are associated with decreases in selectivity of synaptic responses, resulting in greater noise and less signal (e.g., greater difficulty recalling memories). Sleep is considered to be an optimal state for synaptic down-scaling, as the brain is disconnected from the environment, preventing new information being encoded and allowing spontaneous activation of the entire sample of neurons. During slow-wave sleep, synaptic downscaling (removal of excess synapses) is theorised to occur, as

the transitions between intracellular up (peak) and down (trough) states allow the weakening of less strong synaptic connections. This synaptic downscaling is considered the mechanism behind consolidation, in that downscaling reduces noise and increases signal traces, thus consolidating relevant memories and removing excessive potentiation. In essence, weaker synaptic connections are removed, making stronger connections (those that were frequently and strongly activated during wake) more accessible for memory retrieval (Tononi & Cirelli, 2014). Tononi and Cirelli (2014) explain integration of new and old information in that associated memories (whether new or old) can be coactivated during, whilst new memories that are less associated with old memories may be selectively downscaled. However, there is evidence that potentiation can also occur during sleep, particularly during slow-wave sleep (Raven et al., 2018), potentially suggesting different mechanisms associated with consolidation during sleep, such as the formation of new connections, or integration between new and old information.

An alternative model, that focuses on the consolidation function of memory, is the active systems consolidation model (Diekelmann & Born, 2010; Rasch & Born, 2013). This model develops on associative network models such as the complementary learning systems account (McClelland et al., 1995), which suggest a two-stage process for memory consolidation. The active systems consolidation account posits that connections can be formed between memories, via hippocampal replay, and interleaving of new and old memories during the transition of information from the hippocampus to the neocortex explaining integration of new and old memories during sleep. The hippocampus is the area in which declarative memories are encoded and temporarily stored, however, the hippocampus has limited capacity. Therefore, memories are reorganised into the neocortex for longer-term storage. Active systems consolidation models suggest that consolidation occurs when information is moved from temporary hippocampal storage to long-term neocortical storage in association

with sleep spindles, hippocampal ripples, and slow-wave oscillations (Born & Wilhelm, 2012). The active systems consolidation hypothesis suggests that memories are actively consolidated rather than passively protected from interference and deterioration of memory traces that may occur during wakefulness.

Active systems consolidation theories predict that the synchrony of slow-wave oscillations, sleep spindles, and hippocampal ripples is essential for memory consolidation. Hippocampal ripples and spindles occur at a temporally similar time, and are thought to be coordinated by slow-wave oscillations (Born & Wilhelm, 2012). Hippocampal ripples are considered to represent memory reactivation; spindles are theorised to increase connections between memory traces; and slow-wave oscillations synchronise the timing between memory reactivation, the formation of connections between memories, and the transfer of information from the hippocampus to the neocortex (Born & Wilhelm, 2012; Hahn et al., 2020). The temporal synchrony between sleep spindles and hippocampal ripples suggests that sleep spindles may represent a mechanism in communication between the hippocampus and neocortical areas, allowing memory replay during sleep, which is posited to support learning and memory in active systems consolidation accounts (Born & Wilhelm, 2012).

Recent research has shown that a daytime nap prior to encoding is associated with greater activity in the hippocampus following the nap, and restored encoding capacity (Ong et al., 2020). The increased hippocampal activity was significantly associated with spindle count, supporting the active systems consolidation account that sleep spindles function to reorganise memories from the hippocampus to the neocortex. This study is described further in the next section (Section 1.2.2.) in regard to how sleep before learning is associated with restored encoding capacity. A review by Hoedlmoser et al. (2022) explores how the temporal coupling of slow-oscillations and sleep spindles changes across development and has been associated with changes in memory consolidation from infancy to older adulthood. The role of this

slow-oscillation and spindle coupling is thought to be a mechanism of communication in the active systems consolidation model. As such, understanding the cognitive impacts of developmental changes in this coupling (e.g., cognitive decline in older adulthood; cognitive maturation from infancy through childhood and adolescence) could further inform active systems consolidation accounts. Hoedlmoser et al. (2022) suggest that the relationships slow-oscillation and sleep spindle coupling with cognition is supportive of the hippocampal-neocortical dialogue proposed by the active systems consolidation account.

Another prediction made by the active systems consolidation model is that memories are reactivated repeatedly and spontaneously during sleep, this functions to allow integration of new and old information and consolidation of memories. Sleep spindles have been associated with integration of new and old information via lexical competition tasks (e.g., Tamminen et al., 2010). The active systems consolidation model would explain the association between sleep spindles and lexical integration as the interleaving of new and old information via hippocampal-to-neocortical dialogue. As such, the active systems consolidation model suggests a key role of memory reactivation for sleep consolidation, as memory reactivation is required for the interleaving of new and old information. A review of the literature found that memory reactivation during sleep supported consolidation (Schreiner & Rasch, 2017). These memory reactivation studies involve presenting a 'cue' during learning, such as a sound (auditory reactivation) or scent (olfactory reactivation). These cues are later repeated during sleep to cue the learned memory and induce memory reactivation. The review by Schreiner and Rasch (2017) describes how both olfactory and auditory cues can stimulate memory reactivation during sleep, leading to greater consolidation of memories. A meta-analysis of 91 studies and over 2000 participants supports the role of targeted memory reactivation for sleep consolidation (Hu et al., 2021). Specifically, targeted memory reactivation is found to be effective when cues are presented during Stage 2 (g = 0.32) or slow-wave sleep (g = 0.27),

but not during REM sleep or wake. Together, the polysomnographic, neuroimaging and targeted reactivation research described above supports the suggestions that memories are actively reactivated and reorganised during sleep, and that sleep spindles may support such active consolidation by communicating between the hippocampus and the neocortex.

However, recent replication failures of research related to fundamental aspects of active systems consolidation models has raised some questions about this theory (Cordi & Rasch, 2020; Pöhlchen & Schönauer, 2020). Seminal findings by Ellenbogen et al. (2009, 2006, as cited by Cordi & Rasch, 2020) supported the suggestion that sleep actively (rather than passively) consolidated memories, as participants who slept retained a memory benefit over participants that remained awake, despite an interference task occurring prior to recall – suggesting sleep does not just passively protect memories from interference. Cordi and Rasch (2020) identify that these findings did not replicate in a within-subjects design (Pöhlchen et al., 2020). However, the replication study used a small sample size (N = 12), based on power analyses using the predicted effect sizes from the original research. Therefore, the failed replication results found by Pöhlchen et al. (2020) cannot necessarily conclude that sleep does not actively consolidate memories, but perhaps that the effect sizes in the original study were inflated. Further, the original sample was aged between 18-22-years-old whereas the replication sample was aged 18-30-years-old. As such, there may be developmental differences in that the original study encapsulates the latter stages of adolescent development, whilst the replication study extends into adulthood.

Cordi and Rasch (2020) also highlight the current mixed evidence regarding the role of slow-wave sleep for memory consolidation, emphasising the importance of more recent exploration into specific aspects of slow-wave sleep (such as coupling with spindles, or slow-oscillation power). Their review also emphasises a lack of meta-analyses investigating the robustness of the relationship between slow-wave sleep and memory consolidation. Meta-

analyses and other studies should also consider the developmental influence during adolescence that may influence the relationships between sleep architecture and memory, and the theoretical implications of these changes. For example, the developmental changes in slow-wave sleep and sleep spindles during adolescence (explored in Section 1.4). However, most research, and thus most theories focus primarily on adulthood, assuming that 18 years and older encapsulates adulthood, with brief consideration, if any, of developmental changes such as the continued maturation of the brain, and the implications for learning and memory.

Additionally, initial encoding, or potentially sleep before encoding may also influence these relationships (discussed further in Section 1.2.2.). The interaction between encoding and consolidation may explain the "Matthew effect". This refers to the 'rich getting richer' in that those with large existing vocabularies can obtain new vocabulary knowledge more easily. For example, research has demonstrated that weakly encoded information can be consolidated during sleep, but not to the same extent as more strongly encoded information (S. Walker et al., 2019). Improvements in explicit memory for weakly encoded information (five presentations of the stimuli) were seen after sleep consolidation, similarly to standard (ten presentations) and strongly encoded (20 presentations) information.

1.2.2. Sleep before learning (restored encoding capacity)

Given that encoding may influence subsequent consolidation, it is worth considering factors that impact learning capacity. In addition to the consolidation benefits of sleep (discussed above in Section 1.2.1.), the synaptic homeostasis hypothesis further predicts that desaturation of the brain during sleep functions to restore encoding capacity. In essence, learning throughout the day results in the over-saturation of encoding capacity, as higher strength synapses (synaptic potentiation) consume greater levels of energy, resulting in cellular stress and reduced encoding capacity during extended wakefulness. The synaptic homeostasis hypothesis therefore predicts that sleep *before* learning benefits sleep by: (1)

preventing deterioration in learning abilities associated with greater potentiation (and saturation) during wakefulness; and (2) restoring encoding capacity via the desaturation and downscaling of synaptic connections during sleep. The suggestion that sleeping before learning functions to both prevent deterioration during wake and restore encoding capacity has been supported in a nap study by Mander et al. (2011). However, in this study, the restoration of encoding capacity was associated with sleep spindles, rather than slow-wave sleep, which is the mechanism suggested in the synaptic homeostasis hypothesis. Sleep spindles may be a marker of the communication between the hippocampus and neocortical regions (as discussed in Section 1.2.1. above), suggesting a potential role for active systems consolidation accounts in explaining the restoration of encoding capacity following sleep. The active systems consolidation account could explain such restoration via the reorganisation memories from the hippocampus to the neocortical regions, restoring the hippocampus for encoding new information after sleep (Born & Wilhelm, 2012). Indeed, recent research has found that spindle count during a daytime nap was correlated with increased hippocampal activity following the nap, and subsequent improvements in encoding declarative memories (Ong et al., 2020). In this study initial memory was investigated prior to participants undergoing a daytime nap or equivalent period of wake, which allows control for individual differences in memory ability, and can examine the differences in encoding before and after sleep. Increased hippocampal activity was seen following the nap compared to prior to the nap, suggesting that the daytime nap functioned to restore hippocampal encoding capacity. However, similarly to Mander et al. (2011), Ong et al. (2020) did not find a significant relationship with slow oscillations despite predictions made by the synaptic homeostasis hypothesis.

Nonetheless, slow-wave sleep has been causally implicated in renewing encoding capacity in one study. To investigate the role of slow-wave sleep in restoring the brain for

subsequent learning, Antonenko et al. (2013) used transcranial slow-oscillation stimulation to manipulate sleep architecture during a daytime nap. This research supports the synaptic homeostasis hypothesis, in that increasing slow-wave sleep via transcranial slow-oscillation stimulation was found to contribute to desaturating the brain and restoring its capacity for subsequent learning. The role of slow-wave sleep is also key in other theories of sleep and cognition, such as systems consolidation accounts. Furthermore, whilst a causal role of slow-wave sleep is established in Antonenko et al.'s (2013) study, downscaling cannot be assumed to be the key mechanism. For example, it may be that the increased slow oscillations support greater communication between the hippocampus and neocortex, as predicted by active systems consolidation hypotheses. However, Antonenko et al. (2013) did not find a significant influence of sleep spindles for subsequent encoding capacity.

The synaptic homeostasis hypothesis also predicts that excessive potentiation and over-saturation can occur during extended wakefulness, resulting in cognitive deficits associated with sleep deprivation, suggesting the need for sleep to restore the brain for subsequent learning. However, downscaling can also occur during extended wakefulness particularly in the hippocampus. For example, hippocampal shrinking after chronic sleep deprivation is likely a function of unregulated synaptic downscaling, resulting in poorer cognition (see Raven et al. (2018) for a review), rather than excessive potentiation and over-saturation predicted by the synaptic downscaling hypothesis.

Sleep before learning also appears to selectively effect memory for emotional stimuli (Tempesta et al., 2016, 2018), supporting previous findings in consolidation research (Vargas et al., 2019; Walker, 2009). In Tempesta et al.'s (2016) study, participants were either sleep deprived or well-rested prior to the encoding of emotional stimuli. After encoding, participants were allowed recovery sleep prior to recall. At recall, participants who were sleep deprived prior to encoding had poorer memory for neutral and positive stimuli

compared to participants who were well-rested before encoding. However, negative stimuli were encoded equally regardless of sleep prior to encoding. Therefore, whilst sleep deprivation prior to encoding is associated with poorer memory in general, as predicted by the synaptic homeostasis hypothesis for example, this effect is not seen for negative memories, which were selectively preserved despite sleep deprivation. Sleep prior to encoding emotional stimuli is examined in Chapter 4, which investigates the impacts of self-reported sleep prior to encoding emotional word-pairs in an online study. Understanding the impact of sleep prior to encoding may help to explain mixed results regarding the preferential consolidation of emotional memories (e.g., Lipinska et a., 2019).

1.3. Theories of mental health and wellbeing

In addition to the relationship between sleep, learning and memory (discussed in Section 1.2. above), sleep may also be important for mental health and wellbeing. The combination of learning, memory, mental health, and wellbeing is becoming increasingly more relevant to educational policies. For example, recent governmental green papers have increased schools' involvement and responsibilities for young people's mental health (Department of Health & Department for Education, 2017). Therefore, it may be of increasing importance to integrate the two fields of sleep research, investigating mental health and affect alongside learning and memory processes. Some recent research has begun to link mental health and education (Downs et al., 2019; Epstein et al., 2019; Wickersham et al., 2020), yet even less sleep research has investigated this integration. The theories and research described in the following sections considers the relationship between poor sleep and mental health and affective wellbeing, including the role of cognition (e.g., memory biases, thought control) within these relationships. Sleep issues are transdiagnostic, meaning that they can occur across different mental health disorders, and even in the absence of

mental health disorders. Almost all mental health disorders are accompanied by abnormal sleep architecture, as explored in a meta-analysis of polysomnography data (Baglioni et al., 2016). More research into the relationship between sleep, affect, and mental health can further improve our theoretical understanding of these relationships, which may function to inform effective interventions.

1.3.1. REM sleep and mental health

In addition to the role of sleep for restoring encoding capacity (Section 1.2.2.), sleep is considered to prepare the brain for next day emotional regulation. Much of the literature exploring the relationship between sleep, learning, and memory focuses on non-REM sleep. However, REM sleep is considered to be of paramount importance in the function of sleep for regulating affect, wellbeing, and mental health. The second half of the night consists largely of REM sleep (Diekelmann & Born, 2010), which is characterised by rapid eye-movements. The electroencephalogram (EEG) recordings during REM sleep resemble wakefulness, hence the historical name "paradoxical sleep". Whilst no mental health disorders share the exact same patterns of sleep disruption, REM sleep is frequently abnormal in participants with affective (e.g., depression, post-traumatic stress disorder) and anxiety disorders (Baglioni et al., 2016). Goldstein and Walker (2014) suggest that REM sleep is required to recalibrate processes involved in the top-down inhibition of the amygdala via the prefrontal cortex.

The amygdala is an area in the brain that responds to emotional stimuli and is associated with emotional arousal. The prefrontal cortex is thought to be involved in higher-order cognitive functions and controlling the activity of other brain regions via top-down inhibitory control. Goldstein and Walker (2014) suggest the prefrontal cortex maintains inhibitory control over the amygdala via the noradrenergic system. Noradrenaline is a

neurotransmitter and a hormone that is associated with emotional arousal in the brain and body, and levels of noradrenaline are regulated by the noradrenergic system. Goldstein and Walker (2014) propose that REM sleep functions to recalibrate the noradrenergic system, influencing the amount of noradrenaline in the brain and body.

Goldstein and Walker's (2014) framework suggests that if (REM) sleep duration is insufficient, or if a person is completely deprived of (REM) sleep, their noradrenergic system remains uncalibrated. In this instance, there is excessive noradrenaline present in the system. The uncalibrated noradrenergic system reduces prefrontal cortex control over the amygdala, allowing the amygdala to become overly reactive and less specific in its responses. This leads to the hypervigilance associated with anxiety and post-traumatic stress disorder due to hyperarousal caused by excessive noradrenaline. Thus, physiological symptoms of anxiety and stress (e.g., sweating, shaking), and processing of neutral stimuli as threatening occurs. The lack of amygdala specificity is proposed to lead to over-reactivity to all stimuli, including neutral and positive stimuli, which Goldstein and Walker (2014) use to explain the consumption of highly calorific foods and increased reactivity to monetary rewards in the presence of insufficient sleep. Amygdala over-reactivity to rewarding stimuli could also explain the relationship between poor sleep and increased risk-taking (Lau et al., 2019) and substance misuse (Logan et al., 2018). However, Goldstein and Walker's (2014) framework would not explain decreased sensitivity to risk in these situations.

Such a theory of amygdala over reactivity is also less adequate at explaining reductions in positive affect associated with short sleep duration. Goldstein and Walker (2014) acknowledge that some research shows that sleep deprived participants are *less* reactive to emotional stimuli, perceiving it as less salient, as well as becoming less expressive of emotions themselves. Berger et al. (2012) found that children aged 30-to-36-months old displayed fewer positive reactions to solving puzzles and towards positive stimuli, but higher

negative reactions to neutral and negative stimuli including when they failed to solve a puzzle. This suggests that lack of sleep may result in reduced positive affect, but higher negative affect. Goldstein and Walker (2014) suggest this supports their framework in that affective reactivity is dysregulated following lack of sleep. However, this suggests that negative and positive affect are influenced differently by sleep, which would not be explained by a broad over-reactivity of the amygdala to all stimuli. Reductions in positive affect may result in feelings of numbness and anhedonia, which are symptoms of depression. However, people with major depression disorder have shorter REM sleep latency and longer REM sleep duration. Goldstein and Walker (2014) proposed that excessive REM sleep similarly reduces prefrontal cortex inhibition of the amygdala, still resulting in lack of specificity and over-reactivity of the amygdala. More REM sleep is associated with lower levels of noradrenaline. Due to these lower levels of noradrenaline in the system, people with depression experience blunted affect, rather than hyperarousal associated with excessive noradrenaline in anxiety disorder. The noradrenergic system disruption in major depression could explain the interpretation of ambiguous, neutral, or non-salient stimuli as negative, therefore cumulating in the overall negativity bias witnessed in participants with depression. The different implications of positive and negative affect for mental health are discussed further in Section 1.3.3.

Additionally, REM sleep may be associated with emotional memory biases (discussed in Section 1.3.2. below). In a polysomnographic and neuroimaging study, the consolidation of negatively salient emotional memories was associated with both REM sleep and slow-wave sleep (Cairney et al., 2015). 19-28-year-old participants encoded negative, neutral, and positive images prior to an opportunity for overnight sleep consolidation. They then encoded a further set of emotional images before undergoing fMRI scanning, to examine differences in memory for consolidated and non-consolidated memories. Cairney et al. (2015) found that

slow-wave sleep was positively associated with consolidation of negative memories, whilst longer REM sleep was associated with lower consolidation of positive memories. They found that REM sleep was associated with increases in connectivity between the hippocampus and neocortex during the recall test for consolidated negative memories. Further, slow-wave sleep was associated with reduced recall-related activity in the hippocampus. They suggest that REM sleep may function to select which memories are consolidated during slow-wave sleep. REM sleep may also have functioned to support hippocampal-to-neocortical dialogue, as suggested in complementary learning systems accounts that inform the active systems consolidation models (discussed earlier in Section 1.2.).

1.3.2. Sleep, thoughts, memories, and mental health

In a meta-analysis, laboratory effects were found to influence the relationship between sleep and emotional memories, for example, no overall support for the preferential consolidation of emotional memories was found (Lipinska et al., 2019). However, when controlling for initial learning, sample characteristics, or memory task (e.g., free recall, old/new recognition), emotional memories were found to be consolidated more than neutral memories. The importance of initial learning for consolidation may suggest that sleep prior to encoding is important for emotional memories. Whilst Cairney et al. (2015), described in Section 1.3.1. above, measured memory before and after sleep, this study did not consider the impact of sleep *prior* to encoding on memory (discussed in Section 1.2.2.). They only used this data for examining differences in consolidated and non-consolidated memory. The role of sleep *prior* to encoding is investigated in Chapter 3. Chapter 4 also considers the impact of sleep prior to encoding for emotional memory biases, whilst considering the role of positive and negative affect within these relationships.

Selective consolidation of emotional memories may be explained by the Affective Tagging and Consolidation (ATaC; Harrington et al., 2017) model which suggests that

selectively consolidated emotional memories during sleep may be a vulnerability factor for depression. Specifically, the Affective Tagging and Consolidation model suggests that emotional reactivity during encoding can influence later consolidation. As such, the amygdala responses for people with depression may be lower for positive and higher for negative stimuli, resulting in the negative information being selectively 'tagged' for later consolidation during REM sleep. They propose that excessive REM sleep observed in depression may be a genetic vulnerability to depression associated with circadian rhythms, which allows extra opportunity for the consolidation of emotional memories. However, these memories are only skewed toward negative biases in the presence of affective tagging and selective consolidation of negative relative to positive information. Therefore, this model suggests that usually, positive, and negative emotional memories are considered more salient and consolidated to a higher degree than neutral memories. However, in the case of major depression, negative memories are specifically tagged during encoding, resulting in consolidation biases for negative memories. The Affective Tagging and Consolidation model suggests that the role of sleep for positive affect and positive memories is important for understanding depression but is typically understudied in comparison to negative and neutral memories. Chapter 4 explores how sleep and positive and negative affect prior to encoding interact to predict immediate emotional memory. Rather than focusing on inhibition, as in Goldstein and Walker's (2014) model, the Affective Tagging and Consolidation model suggests selection is important in creating cognitive biases. Further, excessive REM sleep in major depression disorder as a cause of the negativity bias associated with depression has not been fully considered in Goldstein and Walker's (2014) model suggesting overall amygdala reactivity toward all stimuli which would not explain reduced positivity.

Nonetheless, the role of sleep in dysregulating the connectivity between the prefrontal cortex and amygdala, proposed by Goldstein and Walker (2014), has also been considered in

other theories. Recently, ideas of reduced prefrontal cortex control in the presence of poor sleep has evolved to include the impact on the hippocampus and a cyclical relationship with intrusive thoughts and memories (Harrington & Cairney, 2021). As discussed in Section 1.2.1. the hippocampus is a fast-learning, short-term memory storage system. Memories are then reorganised from the hippocampus into the slower-learning, longer-term storage of the neocortex. Harrington and Cairney (2021) posit that the retrieval of emotional memories is usually controlled by the dorsolateral prefrontal cortex. In their model, poor sleep reduces the dorsolateral prefrontal cortex's ability to maintain inhibitory control over both the amygdala and the hippocampus. As discussed in the section above, the uninhibited amygdala can result in emotional dysregulation. Harrington and Cairney's (2021) model further proposes that the uninhibited hippocampus gives rise to intrusive memories. These unwanted memories and thought intrusions, paired with emotional dysregulation, cause further disruptions to sleep. For example, staying up at night in emotional distress and ruminating over intrusive thoughts and memories. This causes a vicious cycle of poor sleep, intrusive thoughts, and emotional dysregulation. Harrington and Cairney (2021) do not discuss the potential role of REM sleep within their framework, however, they share the idea of the importance of inhibitory control with Goldstein and Walker (2014). Similarly, to Goldstein and Walker (2014) this model does not provide an adequate explanation for lower positive affect in depression, but instead claims to be a model explaining the role of sleep for a broad range of mental health disorders. The role of emotion dysregulation is further explored in Section 1.3.3. below, and the role of thought control and inhibition in the relationship between poor sleep and mental health is explored in Chapter 5.

1.3.3. Positive and negative affect and mental health

Affect is an umbrella term that encompasses "feelings" broadly, such as mood and emotion. Mood is a longer-lasting form of affect, and emotion is more transient and

changeable (hence "emotion regulation"). Affect is important for everyday wellbeing, and mental health issues involving emotion regulation or disordered mood. According to tripartite theory (Clark & Watson, 1991), high negative affect is associated with increases in depression and anxiety, which can explain common comorbidities and difficulties in differentiating between the two disorders. However, tripartite theory posits that anxiety is uniquely characterised by physiological fear responses such as sweating and increased heart rate. In contrast, depression is distinctively associated with low positive affect, as in symptoms such as anhedonia (i.e., lower motivation, lack of enjoyment or pleasure). In a meta-analytic review, short sleep duration during adolescence (defined as 10-19-years-old in this instance) was associated with decreases in positive affect, increases in negative affect, anxiety, and depressed mood, and reductions in emotion regulation (Short et al., 2020). Most of the studies included in the meta-analysis used self-report sleep measures (53 out of 74).

Objective sleep measures have been used to measure the influence of experimental sleep restriction on affect. Using a within-subjects design, 14-17-year-olds were randomly assigned, in a counterbalanced crossover design, to a sleep restriction (six-and-a-half hours in bed) or well-rested (ten hours in bed) condition for five nights (Baum et al., 2014). Participants wore actigraphy wristwatches to gain an objective measure of sleep and ensure that they engaged with the sleep manipulation at home. When participants were sleep restricted, they reported greater negative affect and lower ability to regulate their emotions compared to baseline and when they were well-rested. This was seen for both self-report and parental reports. Baum et al. (2014) also included hyperactivity, which is not considered to be sensitive to sleep, to assess potential demand characteristics, and found that participants and parents did not report significantly different levels of hyperactivity across the manipulation, suggesting that demand characteristics do not explain the results. In another objectively measured, experimental sleep restriction paradigm, with adolescents aged 15-19-years-old

negative affect was not significantly different between sleep restricted and well-rested participants (Lo et al., 2016). However, positive affect was significantly lower for participants in the sleep restriction group (five hours in bed) compared to the well-rested group (nine hours in bed). The lack of significant relationship between sleep and negative affect may partially be explained by participants reporting that negative affect items on the Positive and Negative Affect Scale were not relevant to them. Additionally, effect sizes for the relationship between short sleep duration and positive affect have been found to be significantly larger than for negative affect during adolescence (Short et al., 2020). However, positive affect is understudied, and the majority of research, including Baum et al. (2014), primarily examines negative affect.

Poor emotion regulation has also been identified as a mediator between poor sleep quality and depression symptoms (O'Leary et al., 2017). The relationship was seen cross-sectionally, and prospectively, which the authors suggest supports the notion that habitually poor sleep worsens abilities to regulate emotions over time, which ultimately results in increased depression symptoms. These conclusions are supported by previously discussed research that found manipulating sleep resulted in changes in emotion regulation (Baum et al., 2014). Sleep has also been found to predict suicidal ideation on a nightly basis for participants with existing history of suicidal ideation (Littlewood et al., 2019). Suicidal ideation can be considered a form of thought intrusion that may be explained within Harrington and Cairney's (2021) framework, discussed in Section 1.3.2. above. However, that is not to say that a single night of poor sleep is sufficient for eliciting suicidal ideation, as 84% of Littlewood et al.'s (2019) sample had sleep issues persisting longer than one year, demonstrating potential cumulative effects of poor sleep on mental health and wellbeing.

Changes in affect associated with sleep may have subsequent influences for emotional memory biases (Section 1.3.2.). If sleep were simply to prevent interference and deterioration

of memory traces, then all memories would be consolidated equally. Instead, the consolidation of emotional memories rather than neutral memories may be selective, depending on affective tagging at the point of encoding (e.g., Affective Tagging and Consolidation model, described in Section 1.3.2. above). However, the role of sleep *before* encoding emotional stimuli may explain which memories are selectively encoded or 'tagged' during encoding, resulting in memory biases (examined in Chapter 4). Therefore, it is important to understand how sleep influences affect, and the subsequent role this may have for affective processes underlying the relationship between sleep and mental health. As such, the models explored in Section 1.3.1. and 1.3.2. may benefit from considering the tripartite model of affect and mental health. Anxiety and post-traumatic stress disorder may be explained by excessive noradrenaline (e.g., Goldstein & Walker, 2014), resulting in the physiological symptoms of stress such as sweating and shaking. High negative affect may also be explained by emotional disruptions associated with lack of inhibitory control over intrusive memories and thoughts (e.g., Harrington & Cairney, 2021). However, these frameworks, positing uninhibited amygdala reactivity, do not provide an adequate explanation for symptoms of high negative affect in the presence of diminished positive affect (e.g., anhedonia), which is fundamental in distinguishing between depression and anxiety in the tripartite model. Instead, the Affective Tagging and Consolidation model may be better positioned to explain the differences in positive and negative affect. Positive affect is widely understudied, which may explain why fewer theories consider the role of sleep for positive affect (A. D. Ong et al., 2017). This may also be particularly important for adolescents, given the relationship between sleep duration and positive affect has been found to have larger effect sizes than the relationship between sleep and negative affect, depression, and anxiety for adolescents (Short et al., 2020).

1.4. Developmental changes in sleep

Adolescence is the developmental period between childhood and adulthood. This thesis adopts a broad definition of adolescence as spanning the years of 10-24-years-old, based on biological and social development during this time (Arain et al., 2013; Sawyer et al., 2018). From childhood to adolescence there is a circadian shift, in which adolescents circadian rhythms shift toward an eveningness chronotype, rather than the morningness preference typically seen during childhood (Merikanto et al., 2018). Chronotype refers to a time-of-day preference and a biological peak in cognitive abilities associated with circadian rhythms. Circadian rhythms refer to internal biological clocks, or 'body-clocks'. Shifts toward later circadian rhythms are seen across mammalian species and in different cultures across the globe, often in line with pubertal development, suggesting biological underpinnings (Hagenauer et al., 2009). This biological shift in circadian rhythms may be evolutionarily advantageous. Socially, adolescence is often characterised by an increased responsibility for oneself (Sawyer et al., 2018), this may include greater focus on peer groups and distancing from parents, particularly in individualistic cultures. Therefore, delayed circadian rhythms may provide adolescents with greater autonomy in the evening, to create space from parents to interact with peers independently. Sleep architecture and the processes modulating sleep are considered to be relatively universal across mammalian species (Borbély, 1982), including developmental changes occurring during puberty such as the circadian shifts described above. Such sleep changes during adolescence may have consequences for learning and memory (discussed later in Section 1.4.2.) as well as mental health and wellbeing (Section 1.4.3.). To understand the impact of developmental changes on sleep, it is important to first understand the processes behind sleep.

In the two-process model of sleep regulation, sleep is considered to be regulated by two processes, process C and process S (Borbély, 1982). Process C is considered the

circadian process, due to its reliance on chronological time and lack of reliance on previous sleep and wake behaviours, hence the disruption to our biological rhythms when jetlagged (Borbély et al., 2016). Process C is thought to regulate daily rhythms, such as feelings of sleepiness, throughout periods of wakefulness. For example, when people are sleep deprived their desire for sleep fluctuates rhythmically rather than consistently increasing. This rhythmic fluctuation in sleepiness is considered to be regulated by process C. Process S is considered to be independent of circadian time and instead is dependent on prior wakefulness and sleep behaviours. Process S can be conceptualised as representing sleep pressure, in that it is theorised to be reduced during sleep but high during wakefulness, increasing with extended periods of wakefulness. These two processes (processes C & S) are hypothesised to work in combination to influence sleep architecture. Slow-wave sleep duration is increased with greater levels of waking activity or learning and therefore is more likely to be regulated by process S, which represent sleep pressure. REM sleep is less reliant on influence from waking activities or previous sleep and instead appears to be controlled by circadian rhythms, in that REM sleep is more time-dependent than influenced by waking activity, which are regulated by process C. Both processes S and C are developmentally relevant during adolescence, due to changes in circadian rhythms and slow-wave sleep, which appear to be associated with pubertal development. For example, circadian shifts toward evening chronotypes could be explained by potential delays in process C in the two-process model of sleep regulation and may have consequences for adolescent sleep duration.

1.4.1. Adolescent sleep restriction and the perfect storm model

People with evening chronotypes (later circadian rhythms) may be more susceptible to having insufficient sleep durations. The social jetlag hypothesis (Wittmann et al., 2006) considers the relationship between chronotype and sleep restriction within the general

population. Individuals differ in their morningness-eveningness preference, potentially on a normal distribution, or a continuum (Natale & Cicogna, 2002). These chronotype preferences have been shown to have genetic underpinnings, in that developmental trajectories of delayed circadian rhythms during adolescence were associated with genetic predispositions towards morning or evening chronotypes (Merikanto et al., 2018). Participants with evening-type preferences are likely to suffer from sleep restriction, in relation to clashes between biological preferences for delayed sleep times and societal pressures for early rise times, this conflict is coined social jetlag (Wittmann et al., 2006). Social jetlag has consequences for mental health and wellbeing, as well as academic performance. These deficits have previously been associated with evening-chronotypes, with interventions focused on shifting circadian preferences (e.g., Zerbini et al., 2018; Facer-Childs et al., 2019). However, other research has investigated whether the timing of social commitments influences these factors, or whether chronotypes are more important. When attending a morning shift at school, participants with an evening chronotype have significantly poorer grades than morning-types; however, in the afternoon school shift, students perform equally regardless of school timing (Estevan et al., 2018). Performance in afternoon examinations has been found to be equivalent across chronotype preferences, however, students with morningness chronotypes outperform students with eveningness chronotypes in morning examinations (Van Der Vinne et al., 2015). Together, this suggests that the interaction between chronotype and social commitments (e.g., school timings) are underlying poorer academic performance for eveningness chronotypes, supporting the social jetlag hypothesis.

With new research, theoretical frameworks of adolescent sleep restriction have been expanded and further developed. The “Perfect Storm” model (Crowley et al., 2018) considers the interaction between bioregulatory and psychosocial pressures in delaying bedtimes, and societal pressures (mostly in relation to school start times) in creating earlier rise times, thus

resulting in adolescent sleep restriction. One bioregulatory influence on delayed sleep timings is the circadian shift towards evening-preferences during puberty. These circadian influences may also influence evening activity levels, potentially due to the wake maintenance zone. The wake maintenance zone is a period of roughly three hours before the onset of melatonin, and thus sleepiness, in which people are alert and have a reduced desire for sleep (Shekleton et al., 2013). Therefore, in the evening hours, adolescents may be more inclined to engage in activities that may add psychosocial pressure to their sleep onset. For example, evening screen time (a psychosocial pressure) may further delay bioregulatory processes (e.g., melatonin onset) together delaying sleep onset even further (Crowley et al., 2018). Evening screen-time may also be associated with increased bedtime autonomy and lower parental monitoring from childhood to adolescence (Tashjian et al., 2019), another psychosocial factor considered within the perfect storm model.

The perfect storm model also considers that delayed sleep onset latency and later bedtimes alone are not the cause for adolescent sleep restriction. Given the opportunity, adolescents will delay their wake-times in order to receive a sufficient sleep duration. For example, in a review of worldwide adolescent sleep patterns, longer sleep duration and later wake-times were observed on weekends compared to weeknights (school nights) (Gradisar et al., 2011). Instead, early school start times, paired with later bedtimes, results in insufficient sleep. As such, delaying school start times for adolescents has become a topic of debate (Kelley et al., 2017; Marx et al., 2017) and has recently garnered public attention with petitions being debated in parliament (Petitions Committee, 2019; Zeichner, 2019). The impact of school on adolescent sleep restriction is explored in Chapter 2, a study investigating adolescent sleep architecture during the school term and holiday, and the impact of potential sleep restriction on consolidation (see Section 1.2.1. for more information on sleep and memory consolidation).

Across different countries, adolescents are found to have sleep of an insufficient duration during the school week, and obtain compensatory sleep on the weekend by delaying their wake-times on average by two-and-a-half-hours (Gradisar et al., 2011). Whilst there were cultural differences, for example, on weekdays students in China woke up much earlier than in other countries, and students in Iceland much later, school day wake times were consistently earlier than on the weekends. This suggests that adolescence is generally a period of vulnerability to sleep restriction, and that this may be associated with school start times. Adolescents who were more likely to have optimal bedtimes were typically younger than 15-years-old, with older adolescents having later bedtimes, thus preventing sufficient sleep duration. From childhood, social jetlag gradually increases up until 18-years-old, and can reach a discrepancy of 3 hours and 18 minutes between the amount of sleep needed and the amount required (Randler et al., 2019). Social jetlag peaks at roughly 15-16-years-old, a key age also identified by Gradisar et al. (2011), and reduces, but is still present by the age of 25-years-old (1 hour and 47 minutes discrepancy). At 16-years-old adolescents in the United Kingdom sit compulsory examinations (GCSE) whereas many other countries have their main compulsory exams at 18-years-old. Therefore, the peak of adolescent social jetlag may be occurring at a similar time to GCSE examinations, making the investigation of adolescent sleep in the UK of paramount importance. However, this relationship also has importance for further and higher education (e.g., A-levels, university) given sleep restriction is present throughout adolescence. This demonstrates that adolescence and young adulthood are periods of vulnerability to sleep issues and insufficient sleep duration. This developmental period is also a time of vulnerability to mental health issues (e.g., Fusar-Poli, 2019) in addition to educational and career milestones. Sleep restriction and adolescent development also has implications for sleep architecture, which may have subsequent consequences for learning, memory (discussed in the next section, Section 1.4.2), and mental wellbeing (Section 1.4.3.).

Therefore, considering developmental influences on sleep may be of importance throughout education, including higher education, and into the workplace.

1.4.2. Consequences for learning and memory

During adolescent pubertal development, there is roughly a forty percent reduction in slow-wave sleep compared to childhood or pre-pubescence (Jenni & Carskadon, 2004). These developmental reductions in slow-wave sleep may have implications for learning and memory theories such as the synaptic homeostasis hypothesis. The reduction in slow-wave sleep during adolescence may be pragmatic, as childhood is characterised by vast acquisition of new information (e.g., learning to walk, talk, read, and write) whereas adolescence is a time more focused on maturation of the brain. Therefore, the mechanisms supporting sleep and learning may differ throughout development. During childhood, vocabulary knowledge may be sparser, and thus slow-wave sleep becomes a key mechanism in supporting the acquisition of new vocabulary knowledge (James et al., 2017). In adulthood, a large vocabulary base may support further knowledge acquisition, in that more connections can be made between words, supporting consolidation (James et al., 2017). However, less is known about how changes in slow-wave sleep during adolescence may affect sleep consolidation. James et al. (2017) identifies the need for further experimental investigation into consolidation of vocabulary across developmental periods. Both slow-wave activity and cortical grey matter decrease during adolescent maturation, and are positively correlated with one another (Buchmann et al., 2011). Slow-wave frequency held the strongest relationship with grey matter decreases, and these relationships were strongest in brain regions associated with adolescent maturation such as the prefrontal cortex. This may offer support for the synaptic homeostasis hypothesis in that maturational decreases in grey matter may occur

during slow-wave sleep, suggesting that synaptic downscaling is occurring during slow-wave sleep.

As discussed in Section 1.4. above, the circadian shifts seen during adolescence may represent developmental changes in process C of the two-process model of sleep regulation. However, process S may also be influenced by adolescent development, as reflected in developmental changes in slow-wave sleep. The time course of slow-wave sleep is hypothesised to be a function of homeostatic need for synaptic down-scaling according to the synaptic homeostasis hypothesis (Tononi & Cirelli, 2003, 2014). That is, if more learning occurs prior to sleeping, there will be a greater need for synaptic downscaling via slow-wave sleep. This may also relate to the two-process model (Borbély, 1982; Borbély et al., 2016), as process "S" is representative of sleep pressure, and slow-wave sleep appears to increase in accordance with increased need for sleep (Rasch & Born, 2013) reflecting mechanisms associated with sleep pressure.

Sleep restriction can also influence sleep architecture; however, slow-wave sleep seems relatively preserved. When adolescents were sleep restricted (5-hour sleep opportunity) for seven consecutive nights, followed by three nights of recovery sleep (9-hour sleep), all sleep architectural stages were restricted with the exception of slow-wave sleep, which was relatively preserved (Ong et al., 2016). Slow-wave sleep duration was similar at baseline and during sleep restriction, suggesting that slow-wave sleep is relatively protected against sleep restriction. Ong et al. (2016) highlight that this is similar to adult sleep restriction, in which slow-wave sleep is also preserved (e.g., Belenky et al., 2003). However, adult (24-55-years-old) sleep architecture is found to return to baseline levels within three nights of recovery sleep (e.g., Belenky et al., 2003). Three nights of recovery sleep was found to be insufficient to recalibrate and re-establish baseline sleep architecture for sleep-restricted adolescents (Ong et al., 2016), which might suggest that the influences of sleep restriction in

adolescence are longer-lasting than in adults. Despite the relative preservation of slow-wave sleep during sleep restriction, there still appears to be cognitive deficits associated with this sleep restriction, suggesting that other sleep stages may be important for cognition. There may also be interactions between the function of sleep for memory consolidation and restoration of encoding capacity.

There is a particular sparsity of literature investigating adolescent sleep before learning, which is an emerging field in the general sleep literature currently dominated by consolidation research. The studies described in Section 1.2.2. typically investigate sleep prior to learning in young adult participants. However, one study has investigated the impact of adolescent sleep restriction prior to encoding. Participants aged 15-18-years-old, were randomly assigned to either sleep restriction (five hours in bed) or healthy sleep duration (nine hours in bed) for five nights prior to encoding pictorial stimuli (Cousins et al., 2018). Memory was then assessed after three nights of recovery sleep. Participants that were sleep-restricted prior to encoding had significantly lower memory than participants who were allowed a healthy sleep duration. Due to the recovery nights prior to the memory task, these results controlled for potential vigilance deficits (measured by psychomotor vigilance tasks) or sleepiness that may have influenced the test. These results demonstrate that poor sleep may reduce encoding capacity and hinder later memory.

The relative preservation of slow-wave sleep during adolescent sleep restriction has previously been used to explain why adolescents were not found to have poorer performance on certain cognitive tasks following lack of sleep. Despite multiple nights of sleep restriction, adolescents had intact attention, processing speed, executive function, long-term memory, and working memory abilities (Voderholzer et al., 2011). Thus, it was concluded that adolescents are exceptionally resilient to the cognitive effects of lack of sleep due to preserved slow-wave sleep. Yet these conclusions have been questioned by a recent expanse

in the adolescent sleep literature. Lo and Chee (2020) proposed that the tasks (measuring attention, processing speed, long-term memory, working memory, and executive functioning) used by Voderholzer et al. (2011) may not have had high enough sensitivity to observe effects of sleep. In more sensitive tasks, declines in performance associated with adolescent sleep restriction are observed. Tasks such as the psychomotor vigilance task, which measures attentional lapses via delayed reaction times, and is known to be sensitive to sleep, shows detriments of sleep restriction in child, adolescent, and adult research (Lo & Chee, 2020).

Poor sleep and daytime sleepiness are also associated with poor academic attainment for adolescents (Dewald et al., 2010) outside of laboratory cognitive tasks. This suggests that adolescents are not resilient to the cognitive deficits of poor sleep, and that poor sleep can have real life consequences for adolescents. Additionally, sleep restriction before or after learning may also affect the relationship with cognition. For example, the cognitive effects of adolescent sleep restriction after learning shows fewer effects than if sleep is restricted before learning (Lo & Chee, 2020). Therefore, the preservation of slow-wave sleep during sleep restriction has implications for theories of sleep and cognition (e.g., synaptic homeostasis hypothesis). Despite the apparent preservation of slow-wave sleep during sleep restriction, restricted sleep still results in cognitive deficits, suggesting that slow-wave sleep alone is not the only sleep architecture responsible for the relationship between sleep and cognition.

Other research has highlighted the importance of sleep spindles for cognition. Spindles and slow-wave oscillations may interact according to systems consolidation models of sleep and cognition. Slow-wave oscillations are theorised to act as a mechanism synchronising sleep spindles and hippocampal ripples to aid communication between the hippocampal and neocortical regions during sleep (as discussed in Section 1.2.1.). In addition to the maturation of slow-wave sleep, sleep spindles also mature during adolescence. Some researchers categorise spindles as fast (>13 Hz) or slow spindles (<13 Hz, e.g., Reynolds et

al., 2018), or group them by their location in the frontal, central, and parietal hemispheres (e.g., Hahn et al., 2018). These categorisations are supported by different developmental trajectories of frontal and centro-parietal spindles, and fast and slow spindles during adolescence (Hahn et al., 2018; Hoedlmoser, 2020). These developmental changes in both sleep spindles and slow-wave oscillations may have genetic and environmental underpinnings, similar to chronotype. In a study investigating adolescent twin-pairs, the vast majority (up to 90%) of variance in slow-wave and spindle activity had genetic foundations (Rusterholz et al., 2018). Which may represent some genetic determinism in adolescent sleep architecture with subsequent implications for mental health, learning and memory. This also further supports the importance of individual differences (Reynolds et al., 2018) and potential implications for tailored sleep interventions (e.g., Lo & Chee, 2020). However, some aspects of slow and fast spindle activity (such as amplitude and power) were mostly attributed to shared environmental influences, suggesting that these may be better targets for environmental interventions.

Maturation of sleep is also seen in increased coupling between sleep spindles and slow-wave oscillations (Hahn et al., 2020). The combination of sleep spindles and slow-wave oscillations may have importance for cognition (e.g., hippocampal-to-neocortical dialogue, discussed in Section 1.2.). In adult research, spindles and slow-wave activity together were found to predict the integration of new and existing information (Tamminen et al., 2013). The coupling of slow-wave oscillations and spindles during adolescence has also been identified as being important for memory. In a longitudinal polysomnography study, the synchrony between slow-wave oscillations and sleep spindles was found to improve from childhood (*M* = 9.5 years old, *SD* = 0.8 years) to adolescence (16 years old + 0.9), and this improved coupling was associated with increased performance on declarative memory tasks (Hahn et al., 2020). In a meta-analysis of adolescent polysomnography literature, sleep spindles were

found to be important for cognitive performance (Reynolds et al., 2018). Reynolds et al. (2018) discuss that during adulthood, greater sleep spindle activity has been found to be associated with improved cognitive performance. In contrast, a relationship between greater spindle activity and lower cognitive performance has been observed during childhood. The meta-analysis of adolescent sleep spindles and cognition found that most cognitive domains showed higher performance in association with greater sleep spindle activity, demonstrating a more adult-like profile (Reynolds et al., 2018). Different aspects of the sleep spindle characteristics (such as power, frequency, and location) were found to have a small influence on effect sizes, which the authors suggest represents the importance of individual neural profiles and brain maturation processes.

Reynolds et al. (2018) highlighted a lack of longitudinal studies investigating sleep spindles during adolescent development, which would address the query regarding developmental, maturational, and individual differences influencing the role of spindles for cognition. Longitudinal studies employ a within-subjects design which allow the controlling of individual differences to identify separate effects of maturation. Some longitudinal studies are now available, and similarly support the importance of sleep spindle maturation for cognition during adolescence. In a longitudinal study, Hahn et al. (2018) found that during childhood or early adolescence (8-11-years-old) participants had more slow than fast spindles throughout the brain. At follow-up seven years later (aged 14-18-years-old), fast spindle density had become greater in central and parietal regions. Analysis of individual differences revealed that increased central fast-spindle density during adolescence compared to childhood predicted greater overnight consolidation of memories across the seven-year period. Fewer developmental changes were seen in spindles in the frontal regions, reflecting the continued maturation of frontal brain areas up until 24 years of age. However, slow spindle maturation from childhood to adolescence was associated with general intelligence,

which the authors suggest is representative of the maturation of frontal brain regions and cognitive processes associated with adolescent development. Together, the adolescent maturation of slow-wave sleep may have implications for the synaptic homeostasis hypothesis, whilst the maturation of sleep spindles may also have implications for active systems consolidation accounts. Coupled with adolescents' vulnerability to sleep restriction (e.g., the perfect storm model), the consequences of adolescent sleep restriction require further investigation to inform our current theoretical understanding of these relationships.

1.4.3. Consequences for affect, mental health, and wellbeing

The relationship between sleep and mental health may also vary developmentally, particularly as adolescence is a period of vulnerability to the onset of mental health disorders (e.g., Fusar-Poli, 2019). Adolescent changes in sleep spindles have also been associated with social anxiety and emotion regulation (Wilhelm et al., 2017). Participants with social anxiety had less slow and fast sleep spindle activity than participants without social anxiety. Lower fast spindle activity in the social anxiety group was associated with increased responsiveness to emotionally salient stimuli. Socially anxious participants rated both positive and negative stimuli as more arousing than non-socially anxious participants. In particular, lower fast spindle activity was correlated with increased ratings of arousal for negative stimuli. Furthermore, whilst Wilhelm et al.'s (2017) study did not find significant correlations between slow-wave activity and social anxiety, they suggest future research examine this link. This may be due to the small sample of 14 participants with social anxiety and 14 without and a wide age range (9 – 17 years old, age matched groups) which may struggle to separate the effects of individual differences (e.g., in the maturation of spindles and slow-wave sleep). The investigation of slow-wave activity within this relationship makes sense, given the relationship between slow-wave sleep and the maturation of the adolescent brain

(Buchmann et al., 2011), and the association between extended prefrontal cortex maturation and adolescent social anxiety (Pfeifer & Blakemore, 2012, as cited by Wilhelm et al., 2017). Together, this further emphasises the importance of investigating both slow-wave and spindle activity during adolescent development for both learning and memory as well as emotional and mental health processes.

As discussed in Section 1.3.1., REM sleep also has implications for mental health and affective wellbeing. In addition REM sleep issues, disruption to sleep continuity, that is, greater wake after sleep onset or sleep fragmentation, and reduced depth of sleep have been observed in adults with affective disorders such as depression (Baglioni et al., 2016). However, in young people with depression (classed as under 18-years-old in this meta-analysis), only continuity of sleep issues was observed, without the presence of disrupted sleep depth of REM sleep pressure. Consequently, different processes in the relationship between sleep and mental health may be occurring during adolescence, which cannot be fully explained by Goldstein and Walker's (2014) model of REM sleep recalibration. Thus, developmentally relevant models of adolescent (and child) sleep and mental health and emotions are required. For example, sleep continuity and the maintenance of healthy sleep may be of greater importance in younger ages than sleep architecture per se.

However, the meta-analysis by Baglioni et al. (2016) did observe slightly shorter REM sleep latency for young people with depression compared to controls without depression. Therefore, suggesting that excessive REM sleep as in Goldstein and Walker's (2014) theory may still be involved in the development of depression, or may be a biological vulnerability to depression as posited by the Affective Tagging and Consolidation model (Harrington et al., 2017). Baglioni et al. (2016) speculatively suggest that interventions targeting sleep, particularly sleep continuity, at a younger age may be beneficial in preventing further detriments to sleep in later years and improving clinical outcomes. Baglioni et al.

(2016) identify the lack of childhood and adolescent research as a significant gap in the literature. This gap is of especial importance considering that early interventions may need to occur during adolescence when sleep and mental health issues begin to emerge.

Developmental influences are not well considered but may be relevant in Goldstein and Walker's (2014) framework and Harrington and Cairney's (2021) model due to prefrontal cortex maturation occurring throughout adolescent development. Young and mid adolescents were compared to adult participants regarding changes in affect following sleep deprivation, and all developmental groups were found to have similar relationships, in that sleep deprivation was associated with reduced positive affect (Talbot et al., 2010). However, there were some developmental influences, in that the younger adolescents rated their biggest fears as significantly more threatening after being sleep deprived compared to being well-rested, which was not seen for adults and older adolescents. The authors suggested this finding may be associated with prefrontal cortex maturation. In turn, this could be associated with less prefrontal cortex control over the amygdala, causing the amygdala to become overly reactive to threatening stimuli, as proposed in the Goldstein and Walker (2014) framework.

The meta-analysis conducted by Short et al. (2020) suggests that the relationship between poor sleep and lower positive affect has a stronger effect size than for negative affect during adolescence. This is interesting, as adolescents are found to be more risk-taking than adults, and this is thought to be associated with prefrontal cortex maturation and reward-seeking (e.g., Logan et al., 2018). Adolescents are also particularly susceptible to sleep restriction (e.g., the perfect storm model, discussed in Section 1.4.1 above). Logan et al. (2018) suggested that sleep restriction and prefrontal cortex maturation together makes adolescents more vulnerable to substance misuse, for example increased caffeine intake and addictions, due to increased reward sensitivity, risk-taking and lower inhibition of impulsivity. It may also be that the increased reward-sensitivity acts as a compensatory

mechanism for low positive affect in the presence of poor sleep. Therefore, understanding adolescent development may provide some insight into the different relationships between poor sleep and positive versus negative affect, and inhibitory processes (explored in Chapters 4 and 5), which may have further implications for our understanding of the relationship between sleep and mental health (e.g., depression).

1.5. Thesis overview

As discussed above, sleep and its architecture are important for learning and memory, as well as affective processes and mental wellbeing. The fields of sleep and learning and memory, and sleep and mental wellbeing are often segregated. However, both fields are important for understanding everyday functioning, and may overlap for example in emotional memory biases associated with wellbeing (e.g., Chapter 4). The current thesis focuses on the relationship between adolescent sleep and both learning and memory, as well as affect and mental health. Research within this field is of relevance to policymakers and practitioners, as it may inform theoretical frameworks, interventions for education and mental health, and policies to promote better sleep, particularly during adolescence.

Together, the literature suggests that sleep functions to regulate processes involved in learning, memory, mood, and mental health. Without sufficient sleep, homeostasis is not maintained, and detrimental effects may be seen for learning, memory, affect and mental wellbeing. During adolescence, sleep changes are occurring that may clash with social commitments (e.g., school), resulting in insufficient sleep. Further, adolescence is a critical time for educational attainment, making the relationship between development, sleep, learning and memory of particular importance during adolescence. Understanding the interactions between sleep and adolescent development, alongside the consequences for learning, memory, and mental wellbeing may inform educational policies and mental health

interventions. This thesis explores the importance of sleep for learning and memory, both regarding naturalistic sleep restriction associated with school terms (Chapter 2), and how napping may serve to restore the brain for subsequent learning (Chapter 3). The influence of sleep may also extend to affect and mental health, which is further investigated in this thesis, with potential implications for prioritising sleep within mental health interventions. The relationship between sleep and affect is explored in regard to emotional memory biases (Chapter 4) and symptoms of depression, anxiety, and stress (Chapter 5). The relationship between sleep and mental health was further examined in a large secondary NHS dataset, examining predictors of sleep issues and sleep medication prescription in the Child and Adolescent Mental Health Services (Chapter 6).

The first study in this thesis is a pilot study, investigating the consolidation of memories, during the school term and holidays, in relation to adolescent sleep restriction and architecture. This builds on previous work demonstrating the effects of experimentally controlled sleep restriction on adolescent sleep architecture (Ong et al., 2016), and the effects of multiple-nights of sleep restriction on cognitive performance (Cousins et al., 2018). However, rather than experimental sleep restriction, this study utilised a naturalistic design, investigating adolescent (13-16-year-olds) sleep during the school term and school holiday in a within-subjects design. Polysomnography (sleep EEG) was used within participants' own homes, rather than in a laboratory or hospital setting, which allowed a closer measure of natural sleep (Bruyneel et al., 2011). The study described in Chapter 2 also used the "cathedruke" non-word learning paradigm with adolescent participants for the first time to our knowledge. This non-word learning paradigm is able to maintain experimental control over the learned stimuli and investigate memory processes associated with both episodic memory consolidation and integration. Whilst the ideal scenario would be to conduct a randomised controlled trial, manipulating school timings, this is not currently feasible within

the United Kingdom (Illingworth et al., 2018). However, this study still has practical relevance for understanding the effect of school start times in the UK on adolescent sleep restriction, which is a popularly debated topic in academia (e.g., Kelley et al., 2017) and in the general population (Petitions Committee, 2019).

The "cathedruke" non-word learning paradigm is also used in a novel way in the second study, in that this is the first time the "cathedruke" paradigm has been used to investigate learning *after* sleep. The second study, in Chapter 3, is a registered report (March et al., 2023) investigating the influence of daytime napping in restoring encoding capacity. In a within-subjects design, older adolescents (18-24-years-old) took a polysomnographically monitored daytime nap, compared to an equivalent period of wakefulness, before encoding non-word stimuli. Chapter 3 addresses a key gap in the literature, in that sleep before learning is largely understudied. This study tests the predictions made by current theories of how sleep may function to restore the brain for subsequent encoding (e.g., Tononi & Cirelli, 2014). The results of this study may also have implications for the institutionalisation of daytime napping as a strategy to improve learning and memory, that could be considered within university or school timetable scheduling (e.g., Cousins et al., 2019).

Similarly, to Chapter 3, the next study also investigated sleep before encoding and recall during older adolescence (18-24-years-old). Chapter 4 used the Pittsburgh Sleep Quality Index (Buysse et al., 1989) self-report sleep questionnaire in an online study. These relationships were examined during COVID-19 lockdowns, and again a year later (outside of lockdown conditions). The study described in Chapter 4 explored theoretical predictions made about sleep before learning and applied this concept to the relationship between sleep and affect. This Chapter investigated the association between sleep and positive and negative affect with the recall of emotionally valenced word-pairs. Therefore, Chapter 4 combined theories of sleep and emotional memory (Goldstein & Walker, 2014; Harrington et al., 2017)

and affect and emotional memory (Bower, 1981). This Chapter bridges the gap between our understanding of sleep, learning, memory, and affective processes that may underlie the relationship between sleep and mental health.

To further explore factors that may underlie the relationship between sleep and mental health, Chapter 5 also considered the role of positive and negative affect alongside other theories predicting mediators of this relationship. Chapter 5 builds on the literature investigating the relationship between sleep and affect during adolescence (e.g., Short et al., 2020) that currently lacks theoretical grounding (A. D. Ong et al., 2017). This study also considered the tripartite theory of affect and mental health (Clark & Watson, 1991) in relation to sleep, which was suggested as an avenue for future research in Baum et al.'s (2014) discussion. These relationships were explored for 13-16-year-old and 18-24-year-old participants in an online study. Chapter 5 included the ability to control intrusive thoughts (e.g., Harrington & Cairney, 2021) and the inhibition of responses to affective stimuli (e.g., Goldstein & Walker, 2014) in a Go/No-Go paradigm as mediators of the relationship between sleep and mental health. This study builds on some findings relating to potential emotional memory inhibition in Chapter 4, considering other inhibitory processes as potential mediators of the relationship between sleep and mental health.

The final empirical study in this thesis (Chapter 6) also explores the relationship between sleep and mental health. The data used in this study were extracted from the Clinical Records Interactive System (CRIS) held at South London and Maudsley, which includes data from the Child and Adolescent Mental health Services (CAMHS) up until the age of 18-years-old. This study investigated the characteristics of CAMHS patients that reported sleep-issues compared to those who did not. Further, this study compared patients that did (or did not) receive pharmacological treatment for such sleep issues. Whilst this data is correlational and exploratory, it is a starting point for investigating potential barriers in accessing

medication to support sleep issues, or potentially characteristics of patients that may be over-prescribed sleep medication.

Altogether, this thesis provides research examining claims made by current theories of sleep and cognition, such as the synaptic homeostasis hypothesis (Tononi & Cirelli, 2014), active systems consolidation models (Diekelmann & Born, 2010; Rasch & Born, 2013), and theories of sleep and mental health (e.g., Goldstein & Walker, 2014, Harrington & Cairney, 2021, Harrington et al., 2017). Most current theoretical knowledge is based on adult studies, with some recent focus on adolescents and children. However, the topical debates around the impacts of sleep restriction are focused on secondary schools and adolescent years (Kelley et al., 2017; Petitions Committee, 2019; Zeichner, 2019). Potential developmental influences on the relationship between sleep and learning, memory, and mental health, such as prefrontal cortex development continue into the mid-twenties (Arain et al., 2013; Sawyer et al., 2018). Therefore, these debates may also be important for sixth-form colleges and universities. However, research during adolescence is required in order to answer questions regarding the role of sleep and the consequences for learning, memory (Chapters 2, 3, and 4), mental health, and affective processes (Chapters 4, 5, and 6).

Chapter 2: Adolescent sleep, learning and memory during the school term and holiday: a pilot study

2.1. Introduction

2.1.1. Adolescent sleep restriction

As discussed in Section 1.4., sleep changes during adolescence can clash with early school start times, resulting in insufficient sleep on school nights and compensatory sleep on free nights (Crowley et al., 2018; Díaz-Morales & Escribano, 2014; Gradisar et al., 2011). Pubertal development has been associated with circadian rhythms shifting towards eveningness preferences during adolescence (Díaz-Morales & Escribano, 2014). Such circadian shifts often result in a chronotype preference for performing activities later in the day, in accordance with biological peaks in cognitive performance (Zerbini & Merrow, 2017). Developmental changes in circadian rhythms are evident in other mammalian species (Hagenauer et al., 2009), suggesting they are biologically driven. The perfect storm model proposes that shifts in bioregulatory processes, such as circadian rhythms alongside psychosocial changes (e.g., increased bedtime autonomy and screen usage) result in later bedtimes (Crowley et al., 2018). Later bedtimes alone are not sufficient to restrict adolescent sleep, as wake-times could be delayed in order to compensate for later bedtimes.

Therefore, the perfect storm model highlights that early school start times may be a key factor in causing insufficient sleep duration for adolescents. The recommended sleep duration decreases across the lifetime. The National Sleep Foundation recommend that adolescents get 8 – 10 hours sleep per night (Hirshkowitz et al., 2015). In the current thesis, sleep restriction refers to achieving sleep of an insufficient sleep duration in regard to these recommendations. Across the globe, adolescents have restricted sleep on school nights and extended sleep duration on the weekends (Gradisar et al., 2011). Sleeping for longer on the

weekends compensates for school night sleep restriction, and so may be longer than their sleep need. When adolescents are able to sleep freely (e.g., during the school holidays), this may be more representative of their sleep need. Gradisar et al.'s (2011) review showed that 53% of the samples in the meta-analysis achieved less than the recommended eight hours sleep on school nights, whilst weekend sleep was never reported to be of insufficient sleep duration. The average sleep time on school nights varied from approximately 7.5 – 8.5 hours, whilst weekend seep was approximately 8.75 – 10 hours. Adolescents in Asia were more sleep restricted than participants from Northern America and Europe, potentially due to cultural and societal factors such as longer school days, earlier school start time (e.g., 06:30 – 07:30), and later school finishing times (e.g., 17:30 – 21:30; Liu, Zhao, Jia, & Buysse, 2008).

Adolescents (15-19-years-old) that were experimentally sleep restricted over multiple nights had significantly restricted sleep architectural stages, except for slow-wave sleep, which remained relatively preserved (Ong et al., 2016). The sleep manipulation was over multiple nights, similar to a school week. However, the manipulation was longer than the average five-day school week, with sleep opportunities limited to five hours for seven nights. This study also found that three nights of recovery sleep was insufficient to restore sleep architecture to baseline levels, suggesting that two-day weekends are not enough to remedy weekday sleep restriction. Using a similar sleep manipulation, but over five consecutive nights (more similar to a school week) deficits in memory encoding were observed when sleep was restricted (Cousins et al., 2018), demonstrating the impact of multiple nights of partial sleep restriction on adolescent (15-18-year-olds) cognition. However, these manipulations of 5-hour sleep opportunities do not represent more naturalistic adolescent sleep restriction. Therefore, we do not know the extent of sleep restriction adolescents in the UK are experiencing, or what impact this may have on learning and memory.

2.1.2. The importance of sleep for learning and memory

Insufficient sleep during adolescence has been found to have consequences for physical health (e.g., obesity, cardio-vascular issues), mental health (e.g., Chapter 5) and cognition such as executive dysfunction and inattention (Owens & Weiss, 2017). The present chapter investigates the impact of adolescent sleep restriction on memory consolidation and integration. One prediction made by the synaptic homeostasis hypothesis is that information learned prior to sleep is more easily accessed after a period of sleep, due to a reduction in noise and increase in signal following synaptic downscaling during slow-wave sleep (Tononi & Cirelli, 2003, 2006, 2014). However, the synaptic homeostasis hypothesis does not account for other behavioural benefits of sleep, such as the integration of new and old information.

Lexical competition is a behavioural marker of lexical integration. Specifically, similar words (e.g., "banana" and "banara") become associated, causing slower reaction times toward competing words, due to increased effort in distinguishing between the associated words. These slowed reaction times are referred to as 'lexical competition' and suggest that the new and old information have been integrated within the mental lexicon. Dumay and Gaskell (2007) found that lexical competition was only seen after a period of sleep, not wake, suggesting that the processes associated with lexical integration were sleep dependent. However, other research has shown that lexical competition can occur after a wake-filled delay, particularly with interleaved exposure (e.g., Lindsay & Gaskell, 2013). Therefore, sleep may be an optimal state for lexical integration, with spontaneous interleaved reactivation of memories, but similar processes may also be possible during wake. One paradigm that measures the role of sleep for both episodic memory and the integration of new and old information is the "cathedruke" non-word learning paradigm, developed by Tamminen and Gaskell (2008). This non-word learning paradigm can measure integration

and word-form memory, using non-words derived from real English words with a late uniqueness point (e.g., cathedral/cathedruke, biscuit/biscal).

A systematic review and meta-analysis of adult (18-65-years-old) research investigating the consolidation and integration of new and old memories during sleep was conducted by Chatburn et al. (2014). Three of the 27 studies included within this meta-analysis used the "cathedruke" non-word learning paradigm described above. After controlling for potential publication biases, the results indicated a moderate effect size for the role of sleep, rather than wake, in active consolidation and integration of newly learned information. This is also supported by a more recent meta-analysis of the benefits of sleep for episodic memory that found the overall effect size of sleep for episodic memory was small-medium (g = .44), but when controlling for potential selective reporting biases, this reduced but was still significant (g = .28; Berres & Erdfelder, 2021). However, the polysomnographic evidence was less consistent, and less available. Only 29% of the included studies reported any data investigating associations between polysomnography data and behavioural outcome measurements, meaning that the authors could not meta-analyse the polysomnography data. From the limited polysomnographic data, Chatburn et al. (2014) concluded that associative memory processes are not generally tied to a specific sleep stage, such as slow-wave sleep (e.g., as suggested by synaptic homeostasis hypothesis). Instead, the authors suggest that the polysomnography correlates are more related to the type of information being learned. The study by Tamminen et al. (2013) is highlighted as an example of this, as the interaction between lexical competition and sleep spindles was identified for spindles located towards the left-hemisphere, which is implicated in the processing of language. Therefore, the authors of the meta-analysis suggest that the correlation between spindles and performance may be due to a combined or individual influence of associative or language processing. Chatburn et

al. (2014) related their findings to reactivation theories, suggesting that perhaps associative processes occur during the same time as general consolidation of memories during sleep.

Berres and Erdfelder (2021) instead suggest that different strengths of the relationships between sleep and episodic memory may be related to the task types and study designs. For example, multiple training repetitions, compared to single sessions of training, were associated with greater benefit of sleep for memory. This supports the idea that memories that are more strongly encoded benefit more from sleep consolidation. Walker et al. (2019) found that the benefits of sleep for the integration of new and old lexical information were only seen for standard or more strongly encoded stimuli, not weakly encoded information. This may suggest that weakly encoded information does not receive the same interleaving consolidation processes that may support lexical integration. In regard to word learning, effect sizes were largest for free recall, then cued recall, and lastly old/new recognition. These findings are consistent with the suggestion that recall (explicit recollection) and old/new recognition (e.g., implicit familiarity and/or explicit recollection) may use different neural processes. Explicit recollection processes may be more hippocampal dependent than familiarity processes (e.g., Brown & Aggleton, 2001) and potentially more sensitive to sleep. Furthermore, when participants have natural sleep or a night-time nap, effect sizes were stronger than in a daytime nap or depriving participants of slow-wave sleep. Therefore, the present study aims to collect naturalistic sleep data by using polysomnography in participants' own homes.

Chatburn et al. (2014) also identified that different tasks have different relationships with sleep. For example, they suggest that there were not correlations between gist extraction (such as in the Deese-Roediger-McDermott (DRM) false memory paradigm) and polysomnographic markers in adults (Chatburn et al., 2014). However, one study in adolescents found that adolescent girls with more sleep spindles had a lower number of

critical lures (false memories from gist extraction) in the DRM paradigm, and longer sleep was associated with more intrusion words (Kuula et al., 2019). This suggests that developmental differences should be considered within theories of sleep and cognition. Adolescent sleep has been understudied, with a recent resurgence (e.g., Lo & Chee, 2020) that could inform theories of sleep and cognition across development.

2.1.3. Current gaps in the literature

The research discussed in Section 2.1.1. suggests that adolescents are sleep restricted on school nights. However, we do not know how restricted sleep during adolescence impacts overnight memory consolidation, particularly with regards to lexical integration (discussed in Section 2.1.2. above). Many studies of adolescent sleep restriction do not include polysomnographic recordings of adolescent sleep, for example the meta-analysis by Gradisar et al. (2011) included only self-reported sleep. On the other hand, experimentally controlled laboratory sleep restriction paradigms, which do include polysomnography (e.g., Cousins et al., 2018; Ong et al., 2016), do not assess the magnitude of naturalistic adolescent sleep restriction in the UK during school nights compared to free nights. The current study aims to fill this gap by using polysomnography in participants' own homes during the school term and holiday in order to assess the levels of sleep restriction naturalistically and the consequences for overnight sleep consolidation.

To investigate the impact of adolescent term-time sleep restriction, cross-sectional studies have used academic attainment (e.g., Enright & Refinetti, 2017; Phillips et al., 2017). However, behavioural paradigms are a more controlled method that can also explore theoretical hypotheses associated with our understanding of the relationship between sleep, learning, and memory. For example, the complementary learning systems model, which informs the active systems consolidation account, highlights in importance of integrating new (e.g., "cathedruke") and old (e.g., "cathedral") information. Sleep is thought to facilitate this

integrative process through hippocampal replay and synchronised reactivation of long-term memory storage, allowing interleaving of new and old information (e.g., Rasch & Born, 2013). One paradigm that can investigate these relationships is the "cathedruke" non-word learning paradigm (described in Section 2.1.2. above). The "cathedruke" non-word learning paradigm is well-established in the adult sleep literature and has also been used in studies with children. However, the present study is the first to our knowledge to use this paradigm with adolescents.

Using the "cathedruke" non-word learning paradigm (described in Section 2.1.2. above), Tamminen et al. (2010) was the first study to consider the role of sleep spindles for lexical integration. Young adult participants were randomly allocated to either a sleep or wake condition. In both conditions, participants were trained and then tested immediately, again after a 9.5-to-10-hour interval (delayed test), and once more a week later. For the wake group, the training and *immediate* test occurred in the morning, whilst this was in the evening (followed by a period of overnight sleep) for the sleep group. Both groups were tested at a similar time a week later. Lexical competition was not observed in the immediate test but was observed in the delayed test for both the wake and sleep group. However, the lexical competition effect was stronger in the sleep condition than the wake condition, suggesting sleep is the optimal state for memory consolidation. Sleep spindle activity was associated with lexical competition but not with increases in recollection or reaction times for old/new recognition task. This suggests the role of sleep spindles was unique to the integrative processes underlying lexical integration. The effect of spindle activity was found in both Stage 2 and slow-wave sleep spindles. Slow-wave sleep was the only sleep stage that showed correlation with performance on any task. Slow-wave sleep was correlated with faster reaction times in recognising if non-words had been learned or were unlearned (old/new recognition). This suggests that sleep spindles may play a role in lexical integration

processes, such as memory reactivation and interleaving of new and old information during communication between the hippocampus and neocortex (i.e., active systems consolidation accounts). Whereas, slow-wave sleep may be associated with consolidating memory, potentially by making the memory trace more easily available (e.g., synaptic homeostasis hypothesis, active systems consolidation model).

Overnight sleep consolidation has also been found to play an important role in childhood vocabulary acquisition using the "cathedruke" non-word learning paradigm (Henderson et al., 2012). Children (7-12-years-old) were trained on the non-word paradigm in the morning or evening and then had a test ~12 hours later, after a delay filled by wake for morning trained children and sleep for evening trained participants. Participants were tested again after 24-hours (when both groups had a sleep opportunity) and at a one-week follow-up. At the 12-hour test, only the evening trained participants, who had a sleep filled delay, demonstrated integration of the newly learned non-words into their mental lexicon. This research demonstrates that sleep supported lexical competition to a greater extent than a delay with just wake, supporting findings in adult studies. The current study addresses a gap in the literature by using the "cathedruke" paradigm to examine adolescent learning and memory.

Measuring sleep architecture with polysomnography is important for addressing theories related to neural processes underpinning the relationship between cognition and sleep, such as the synaptic homeostasis hypothesis (Tononi & Cirelli, 2003, 2014) and the active systems consolidation models (Born & Wilhelm, 2012). There is currently a sparsity of research investigating these theories in relation to adolescent sleep restriction, as a lot of research focuses primarily on subjectively measured or experimentally manipulated sleep duration. Polysomnography used at home, rather than in hospitals, is found to be more representative of natural sleep and is often preferred by patients (Bruyneel et al., 2011). Therefore, polysomnography used in participants' homes, rather than in a sleep laboratory,

may also be more naturalistic and better accepted by participants. Polysomnographic data were collected both on a school night and free night during the school holiday. The present study investigated the naturalistic sleep architecture that may be associated with adolescent sleep restriction and the potential consequences for memory consolidation and integration. Unfortunately, data collection was interrupted by COVID-19 lockdowns, meaning insufficient data were collected to complete the study, and therefore, the present study is only useable as pilot data.

2.1.4. Hypotheses

1) Hypothesis 1: Adolescents will be significantly sleep restricted during the school term compared to the holiday:
    a) Total sleep duration will be shorter during the term than the holiday (e.g., Crowley et al., 2018; Gradisar et al., 2011).
    b) During sleep restriction, all sleep architectural stages will be significantly shorter, whilst slow-wave sleep duration will be relatively preserved (Ong et al., 2016).
2) Hypothesis 2: Participants' performance on the non-word learning paradigm will be significantly poorer during the term compared to the holiday:
    a) There will be a greater difference in reaction times for control words compared to base words on the lexical decision task during the holiday, demonstrating more lexical integration, than during the term.
    b) There will be greater accuracy in the cued recall task during the holiday compared to the term, demonstrating greater ability to recall the novel words.
3) Hypothesis 3: The sleep restriction that adolescents experience will be correlated with their behavioural performance on the non-word learning task:

a) There will be smaller overnight increases in lexical competition and cued recall when sleep is restricted (e.g., during the term-time) compared to when sleep duration is longer (e.g., the holiday). In essence, there will be a significant interaction between test time (immediate versus delayed) and school phase (term versus holiday).

2.2. Methods

2.2.1. Participants

Thirteen participants aged between 13-16-years-old were recruited from three secondary schools in the South of England (Essex and Surrey). Characteristics of the sample are described in Section 2.4.1. Each participant volunteered to take part by completing and returning parental and participant consent forms. Participants were paid £15 cash and entered a prize draw to win a £25 voucher of their choice. Two participants are missing data in one school phase due to lockdown preventing further data collection.

2.2.2. Materials

2.2.2.1. Screening Questionnaire

After written consent was received, participants were given an initial questionnaire to complete prior to setting dates for data collection. The initial questionnaire asked participants to respond ‘yes/no’ as to whether they had: a sleep disorder, a psychiatric disorder, known special educational needs, were within the study age range (13-16-years-old) and whether English was their first language.

2.2.2.2. Background measures

2.2.2.2.1. Wechsler's Abbreviated Scale of Intelligence (WASI-II; Wechsler, 2011)

Previous research has found a positive correlation between general intelligence and memory gains after sleep, but no correlation was found for memory change during wake, suggesting that participants with higher intelligence may have greater offline processing abilities (Fenn & Hambrick, 2015). The current research used the Matrix Reasoning subscale of the WASI-II as a non-verbal intelligence measure. When completing the WASI-II matrix reasoning task, participants were asked to complete a pictorial pattern from a multiple-choice selection. Accuracy was scored as correct (1) or incorrect (0), and data collection ceased after three continuous scores of zero. A total score of correct responses was then calculated, with the maximum score being 30. Raw scores were then converted into a standardised score in relation to the participants chronological age on the date of the test. These data were used to characterise the sample.

2.2.2.2.2. British Picture Vocabulary Scale (BPVS)

A meta-analysis by James et al. (2017) suggests that a larger existing vocabulary may support further acquisition of new vocabulary for adults and children. Therefore, the current research measured participants' existing vocabulary with the BPVS. The BPVS is appropriate for use with children and adolescents. When completing the BPVS, participants are asked to select one of four pictures in a quadrant that represent a word spoken by the examiner. Different age groups start on different items, and all prior items are considered to be correct. Participants' responses were scored as correct (1) or incorrect (0) and data collection was discontinued after eight consecutive incorrect scores. A standardised score according to participants age can be calculated from the raw score. The raw scores for the age group

included in the current sample could range from 97 -168. These data were used to characterise the sample.

2.2.2.2.3. Automated Operation Span Task (AOSPAN; Stone & Towse, 2015)

Working memory capacity has also been associated with word-pair learning after a period of sleep, suggesting a role of working memory capacity in offline memory consolidation (Fenn & Hambrick, 2012). To measure working memory capacity, this research used the operation span task, similarly to Fenn & Hambrick (2012). The AOSPAN task was delivered automatically on a laptop via Tatool. The task asks participants to remember a series of numbers whilst making a keyboard button press to indicate whether an equation (e.g., 3 * 3 = 9) is correct or incorrect after the presentation of each number. For each trial, there were between 2 and 6 number and equation pairs for participants. At the end of each presentation phase, participants were asked to recall the numbers, in sequence, that they memorised during the trial. To successfully pass a trial, the participants must get both the equation and recollection of the number correct. These data were used to characterise the sample.

2.2.2.2.4. Socioeconomic status (SES)

Low socioeconomic status has been associated with poor sleep in children (Bagley et al., 2015; Buckhalt et al., 2007) and young adolescents (M = 12.6, SD = 0.6 years old; Marco et al., 2011). In the UK free-school meals status has been considered to be a reasonable measure of socioeconomic status (Ilie et al., 2017), and this information can be collected from schools. The intention was to use free school meals as a measure of socioeconomic status in the current study. However, this was no longer practical due to delays in recruitment and burdens on schools during the COVID-19 pandemic. Given that at-home polysomnography data were collected, postcode data were available as an alternative measure

of socioeconomic status. Postcodes can be used to calculate indices of multiple deprivation, which consider income, employment, education/skills/training, health and disability, crime, barriers to housing and services, and the living environment of the neighbourhood. This data was accessed using a website tool (Ministry of Housing, Communities & Local Government, 2019) that calculates index of multiple deprivation based on the 2019 census. The deciles range from 1-10, with higher scores indicate less deprivation (i.e., higher socioeconomic status).This information was used to characterise the sample.

2.2.2.2.5. Puberty (Physical Development Scale; Earis et al., 2000)

The physical development scale (Earis et al., 2000) is a questionnaire that participants used to self-report their pubertal development stage. Physical examination (for example, by a paediatrician) or self-report (using visual photographs or diagrams) using the Tanner Sexual Maturity Scale may be a more accurate measure of pubertal development. However, ethically, and practically working with adolescents within schools this would not be appropriate. Therefore, the physical development scale is a preferred method for use within schools (Petersen et al., 1987). From the physical development scale, we used three questions; these questions ask about height, body hair and skin changes. We also used the two additional questions for boys (voice, facial hair) and girls (breast development, menstruation) and one question about their development in relation to peers of the same sex and chronological age. Participants responded to the questions on a scale from one ("not yet started") to four ("seems completed"), except for the menarche question, for which participants respond 'yes' (4) or 'no' (1) as to whether they have begun menstruating. This produces scores from 5 – 20. These scores can be translated into categories of pubertal development (Petersen et al., 1987).

2.2.2.2.6. Chronotype

The Morningness-Eveningness Questionnaire (MEQ, Horne & Österberg, 1976) is a 19-item self-report questionnaire, in which participants answer questions such as "at approximately what time of day do you usually feel your best?" and "if you got into bed at 11 p.m., how tired would you be?". Two questions (items 1 and 2) are on a 6-point scale, 3 questions (items 10, 17 and 18) on a 5-point scale, and 14 questions (items 3-9, 11-16, and 19) on a 4-point scale with responses varying according to the question. From the scores produced by the MEQ, participants can be categorised morning, evening and intermediate chronotypes. Alternatively, the MEQ scores can be conceptualised as a continuum between extreme evening and extreme morning types, with the majority of the population falling around the middle (intermediate chronotype; Natale & Cicogna, 2002). Continuous scores are used in the current study, with lower scores indicating greater evening-type preferences.

2.2.2.2.7. Background sleep measures

Actigraphy (described in Section 2.2.2.2.7.2.) and self-report diary (described in Section 2.2.2.2.7.1.) data were used to collect multiple night recordings for each participant around the night of polysomnography data collection. School night sleep may be restricted, and weekend sleep may compensate for this by extending sleep duration. Therefore, the analyses examined school nights and weekends separately in the term-time, weekends were not separated in the holiday. Both actigraphy and diary background sleep data were considered separately. This maximises the use of the existing data, as participants sometimes completed the diary but did not wear the actigraphy device or vice versa on different days. As each of these measurements includes multiple observations, influential outliers were identified and removed using Rosner's test. Rosner's test can detect several statistical outliers simultaneously to prevent masking closely positioned outliers (Soetewey, 2020).

2.2.2.2.7.1. Daily diary

Participants were asked to complete an online sleep diary every morning for approximately one week prior to the night of polysomnography. The daily diary was completed on Qualtrics (www.qualtrics.com) and consisted of six questions related to sleep, checklist questions about consumption of caffeine products, and four questions relating to screen usage. The daily diary took an average of approximately 3 minutes to complete (*M* = 179 seconds, *SD* = 101.57 seconds). Participants were asked questions about their sleep such as the time they went to bed and awoke, their self-reported sleep duration, quality of sleep and how alert they felt in the morning from sleep diary questions on the sleep council's website (sleepcouncil.org.uk, accessed February 2019). Participants also reported caffeine intake the previous day, with questions inspired by a modified version of the caffeine intake questionnaire (Heaton et al., 2012), including time frames of consumption to aid memory. Using this self-report, an estimated milligrams of caffeine consumption can be calculated using data from the Committee on Toxicity (2001). Screen usage in the evening and around bedtime was measured with questions from Polos et al. (2015). Some questions, for example, asked about duration spent using screens after dinner but before bed, whilst in bed, and awakenings caused by messaging.

2.2.2.2.7.2. Actigraphy

During the week of background sleep measurements, participants were asked to wear an ActiLife wrist-worn actigraphic device on their non-dominant wrist. Participants were instructed to always wear this device during the day and night, only removing the watch for showering/bathing or sports. The Cole-Kripke and Actigraph algorithms were used to automatically detect sleep periods as this is shown to be an appropriate algorithm to use with adolescent actigraphic sleep data (Quante et al., 2018).

2.2.2.3. Polysomnography

In an attempt to gather sleep architectural data in the most naturalistic setting possible, at-home polysomnography was used. Polysomnography in participants' own homes has been reported to be more acceptable by subjects (Bruyneel et al., 2011). Sleep efficiency has also been demonstrated to be superior and sleep duration was longer for at-home, in comparison to in-hospital, polysomnography (Bruyneel et al., 2011). At-home polysomnography has also been deemed appropriate for use in children and adolescents up to 16-years-old (Stores et al., 1998). Therefore, the current study collected at-home polysomnography data using an Embla N7000 device. Data were recorded on REMLogic via a laptop. Electrodes were placed on the scalp on the left and right frontal (F3, F4), central (C3, C4) and parietal (P1, P2) positions. These electrodes were referenced to the average of the left and right mastoid (behind the ear). The ground electrode was placed on the forehead. Electro-oculogram (EOG) was used to measure eye movements and electromyography (EMG) was used to measure muscle tone around the jaw. We aimed to achieve impedances below 5Ω for EEG electrodes and below 10Ω for EOG and EMG electrodes where possible. Signals were sampled at 500 Hz.

2.2.2.4.Novel word learning paradigm

66 stimulus triplets consisting of a familiar base word (e.g., 'cathedral'), a fictitious novel word (e.g., "cathedruke"), and a similar-sounding non-word foil to be used in the old-new categorisation task (e.g., "cathedruce") were selected from stimuli used in Tamminen and Gaskell (2008). These were divided into three lists of 22 triplets. The words included in each of the three lists are reported in Appendix 2.1. The lists were roughly matched on number of syllables (M = 2.45, range 2-3), number of phonemes in the words (list 1: M = 6.45, range = 5-8; list 2: M = 6.55, range = 5-9; list 3: M = 6.45, range = 5-9), and SUBTLEX-UK (van Heuven et al., 2014) frequency of the use of the base words (list 1: M = 3.37, range = 2.43-4.52; list 2: M = 3.23, range = 2.43-4.06; list 3: M = 3.47, range = 2.89-

4.10). Two of the lists were used for encoding and testing, one in the holiday and the other in the term-time school phase. The third list remained an untrained control. The lexical competition task required 88 filler words, half used in the holiday and the other half in the term-time school phase. This ensured that only one quarter of real words presented in the lexical decision task at each session were base words related to the encoded list, making it unlikely that participants would explicitly notice the overlap between base words and the phonologically overlapping trained novel words. In each list of 44 real filler words, 15 were monosyllabic, 15 were bi-syllabic and 14 were trisyllabic. These filler words were available from the stimulus set of Tamminen and Gaskell (2008). One hundred and fifty-two filler non-words were also available from Tamminen and Gaskell (2008). This means that most filler non-words were not repeated across the two school phases (i.e., in each of the two lists of 88 non-word fillers), but 23 non-word fillers were used in both school phases. Since responses to the filler items are not analysed, this partial repetition of fillers did not impact on the results of the study. Each list of non-word fillers contained monosyllabic (N = 40), bi-syllabic (N = 24), and trisyllabic (N = 24) non-words. These non-word fillers were generated by Tamminen and Gaskell (2008) and were created by changing one phoneme of a real word to form word-like legal non-words. None of the real words used for this purpose occurred in any of the tasks used in the experiment. All spoken recordings of the stimuli were made by Tamminen and Gaskell (2008) and were recorded by a female speaker of southern British English in a soundproof booth.

The novel word training and test tasks closely followed the protocols established in Tamminen et al. (2010). Participants learned the novel words by listening to repeated presentations of each novel word while monitoring for one of six target phonemes (/P/, /D/, /S/, /M/, /N/, /L/). There were five blocks of phoneme monitoring: in each block each novel word was heard six times, once while monitoring for each of the six target phonemes. This

yielded a total of 30 exposures to each novel word in this task. Each phoneme monitoring trial began with the visual presentation of the target phoneme. The spoken novel word was presented through headphones 500 milliseconds after the onset of the visual target. The target remained on the screen until the end of the trial. Participants were required to indicate with a button press whether the target phoneme was present or absent in the novel word. The trial ended when a button press was made or after 3000 milliseconds had elapsed from the offset of the auditory presentation of the word. In addition to the phoneme monitoring task, there were four blocks of verbal repetition task where participants heard each novel word once and were asked to repeat the word aloud. The trial ended when the participant pressed a response key to indicate they had completed their response. Vocal responses were recorded via head mounted microphone. These blocks were interleaved with the five phoneme monitoring blocks. Therefore, the total number of exposures to each novel word across the two training tasks was 34.

In cued recall participants heard the first two or three phonemes of a trained novel word, excised from the recording of the complete word. The length of the cue (two or three phonemes) depended on how much of the word is needed to identify it reliably from other words beginning with the same sequence of phonemes. The task was to say aloud the complete novel word during the 10 seconds that followed from the presentation of the cue. As participants tended to skip through these trials without attempting recall, we piloted a free-recall task, in which participants were given 5 minutes to verbally recall as many of the non-words as they could remember, these were audio recorded, and the researcher provided scripted prompts to encourage attempts. We also piloted a cued recall task, in which the researcher played the beginning phonemes (cue) and used scripted prompts to encourage recall. No feedback was given regarding accuracy to avoid further learning.

In the lexical competition task, participants were asked to determine as quickly and accurately as possible whether a presented spoken stimulus was a real word or a non-word (i.e., make a lexical decision). The task included 22 base words (e.g., cathedral), 22 control words (e.g., dolphin), 44 filler real words, and 88 filler non-words. The task began with a practice block of ten trials, five non-words and five real words. A lexical competition task trial began with the visual presentation of a fixation cross for 500 milliseconds. The spoken word was then presented through the headphones. The trial ended when a button press (word or non-word) was made or after 2500 milliseconds had elapsed from the offset of the spoken word. At the end of the trial feedback on response accuracy was given through the presentation of a smiley face (correct) or frowning face (incorrect). The feedback remained on the screen for 750 milliseconds before the next trial began. Feedback was provided in the lexical competition task to ensure participants were accurate in their identification of real versus non-words, rather than guessing before listening to the word up to the uniqueness point. This enables the identification of potential shifts in reaction times as a consequence of integrating the novel word into the mental lexicon. This feedback methodology was used in Tamminen e al. (2010), which was used to inform the procedure of the current study. Therefore, we aimed to follow this procedure as closely as possible, given the novel use of the paradigm with an adolescent population in the current study. The order of trials was newly randomised for each participant by the software. Reaction times were measured from the onset of the spoken word.

In the old-new categorisation task, participants heard 22 trained novel words (e.g., “cathedruke”) and 22 untrained foils (e.g., “cathedruce”). Participants were asked to make a button press to indicate if the word was old (i.e., trained) or new (i.e., a foil). A trial began with the visual presentation of a fixation cross for 500 milliseconds. The spoken word was then presented through the headphones. The trial ended when a button press (old or new) was

made or after 3000 milliseconds had elapsed from the onset of the spoken word. If no response was made within the 3000 milliseconds, a message saying "too slow" was presented on the screen until the participant pressed a button to proceed to the next trial. This was to encourage fast responding. No feedback was given regarding accuracy to avoid further learning. Reaction times were measured from the onset of the spoken word.

2.2.3. Design and procedure

After parental and participant consent was received, participants were asked to complete an initial screening questionnaire to determine eligibility for the study. If participants were eligible to take part, data collection dates were set at a time convenient for the family, school, and participant. This often meant that there were limited opportunities for data collection during examination seasons (May – July) and over the winter holidays (December – January). This study was given ethical approval following the Royal Holloway University of London Ethics Committee procedures.

A within-subjects design was used, with participants completing the non-word learning tasks on four occasions. A 2 x 2 design was implemented with an immediate (evening) and delayed (morning) test, during the school term compared to the school holiday. For term-time data collection, data were collected on Wednesday to Thursday or Thursday to Friday night to allow three prior nights of weekday sleep routine and prevent any effects of weekend compensatory sleep. During the holiday, data were collected on any day of the week, provided that a minimum of three nights of holiday sleep preceded data collection, and a minimum of three nights before returning to school. Collection of holiday and term-time data were planned to be counterbalanced to prevent any biases caused by participants adjusting to the polysomnography equipment. However, this was disrupted by COVID-19 lockdowns and therefore counterbalancing was not possible within this pilot study.

Participants were trained and tested on the "cathedruke" non-word learning paradigm in the evening and tested immediately and again the next morning. During the holiday, morning data collection timing was determined by the time that the participant naturally awoke. For the term-time, data collection times were confirmed with the school and occurred during the participants' first lesson of the day after registration. The procedure for the training and testing in the non-word learning paradigm are described above in Section 2.2.2.4. Training involved audio presentation of stimuli over headphones. The testing began with the (cued) recall task. Initially, cued recall was recorded on the laptop, but this led to lots of blank responses. For other participants, free recall was recorded on a mobile phone, but participants were difficult to engage. The most effective task was cued recall lead by the experimenter with scripted prompts to encourage participation. After the recall task, two testing tasks were completed on the laptop with keyboard button responses; recognition of new (untrained foils) or old (trained non-words) words (old/new recognition) and a lexical decision task, in which participants responded as to whether a word was real or a non-word. The researcher then applied the at-home polysomnography equipment. The researcher then left the participant's home. Participants were shown how to unplug and plug-in the patient unit to ensure they were able to get in and out of bed and correctly set-up the equipment to record when they went to sleep after the researcher had left. They were instructed to unplug the unit or remove electrodes when they were finished sleeping in the morning. The researcher then came to collect the equipment and conduct the delayed test as soon as possible. In the term-time, the immediate testing session was during the participant's first lesson after they had registered attendance at school. During the holiday, the immediate testing session was as soon as possible after the participant woke up.

All information provided was stored separately from identifiable information such as the participants' names and dates of birth. All data were assigned an ID number to link

together multiple parts of the within-subject data whilst protecting participant confidentiality and anonymity.

2.3. Analysis plan

2.3.1. Word learning data

Cued recall data were phonetically transcribed using the international phonetic alphabet. However, the recall task was changed from free recall, to cued recall on the laptop, and then to researcher led cued recall, due to issues in engaging adolescent participants to recall the non-words. Therefore, due to inconsistencies in the methods used for recall, these data are not analysed or included in this pilot study.

To investigate lexical competition effects, button press reaction times were log-transformed and meaningless responses (reaction times under 300 or over 2500 milliseconds) were removed from the analysis. Using only correct responses, mean reaction times were calculated for each type of word (base, control). The mean reaction times for control words were subtracted from the mean base word reaction times; this allows a measure of the difference in responses times to control words compared to base words (e.g., “cathedral”), which have trained competitors (e.g., “cathedruke”). Lexical competition effects were calculated separately for the immediate (evening) and delayed (morning) tests, as lexical competition effects are often not seen in immediate testing (e.g., Dumay & Gaskell, 2006; Tamminen et al., 2010). Reaction times for the old/new recognition memory task were again calculated for correct responses only and were log-transformed. A mean reaction time was then calculated for the immediate and delayed tests during the holiday and the term.

2.3.2. Polysomnography data

The sleep architectural data collected via polysomnography were scored manually on REMLogic using American Academy of Sleep Medicine (AASM) guidance. The researcher’s

scorings were then compared to algorithm calculated scores conducted on REMLogic. If agreement between the researcher and algorithm scoring was below "good" inter-rater reliability, based on Kappa ratings (>0.60) and percentage of agreement (>75%), the researcher would manually rescore the data. The manual scoring conducted by the researcher was used in the final analyses. The average inter-rater reliability of the final data was 75.38% (*SD* = 0.09%) and 0.65 Kappa (*SD* = 0.12).

2.4. Results

2.4.1. Characteristics of the sample

The age of the total sample (*N* = 13) ranged from 13.39 – 16.93 years old (*M* = 14.89, *SD* = 1.20). The sample was 69.23% male (*N* = 9) and 30.77% female (*N* = 4). Participants were recruited from three schools, 30.77% from school one (*N* = 4), 30.77% from school two (N = 4) and 38.46% from school three (*N* = 5). The index of multiple deprivation deciles (proxy for socioeconomic status) for this sample ranges from 4-10 (*M* = 9.15, *SD* = 1.63) suggesting that this sample is skewed towards higher socioeconomic status.

The average raw score on the WASI non-verbal reasoning test was 20 (*SD* = 4.28, range = 13-27), the average age-normed t-score was 52.23 (*SD* = 11.59, range = 37 – 75), with 50 being the normed average. On the vocabulary scale (BPVS), the average raw score was 154.23 (*SD* = 5.00, range = 148-163), the average age-normed t-score was 105.62 (*SD* = 7.29, 91 – 115), with 100 being the normed average. On the working memory task (AOSPAN) scores ranged from 18-45 (*M* = 30.17, *SD* = 8.20). Participants' scores on the Morningness-Eveningness Questionnaire ranged from 38 (moderate evening) to 69 (moderate morning). The average of the sample was within the intermediate (range for intermediate category is 42-58) range (*M* = 51.46, *SD* = 9.63). The range of scores on the pubertal developmental scale was 6 (mid-pubertal) to 19 (post-pubertal). The average for the sample was post-pubertal (*M* = 14.5, *SD* = 4.03). Responses on the Qualtrics daily diary were used to estimate caffeine

intake in milligrams (see Table 2.1.), and measures of self-reported evening screen-time (see Table 2.2.) during term-time weekdays, term-time weekends, and school holidays. Seven participants had within-subject polysomnographic data in both the term and holiday school phase. Eleven participants had within-subject data collected via actigraphy, and nine via diary data, in the holiday, school-night and weekend school phases.

Table 2.1. Self-reported estimated caffeine intake (in milligrams) calculated from the daily diary

	Mean	Standard deviation	Median	Range
Term	38.6	46.7	20.5	0 – 170
Weekend	90.0	101	49.2	0 – 346
Holiday	86.8	87.6	63.8	0 – 326

Note: This is within-subjects data from nine participants who completed the daily diary in each school phase. 'Term' refers to school nights, i.e., weekdays during the school term. 'Weekend' refers to term-time weekends. 'Holiday' refers to school holidays. For reference, the Committee on Toxicity (2001) estimate there is 75 mg of caffeine per 190 ml cup of instant coffee.

Table 2.2. Self-reported evening screen-time from the daily diary

	Holiday (57)		Term (48)		Weekend (20)	
	N	%	N	%	N	%
Screen time between dinner and bed						
None	1	1.75	5	10.42	3	15.00
Less than 30 minutes	10	17.54	3	6.25	0	0.00
30 minutes to 1 hour	9	15.79	12	25.00	2	10.00
1-2 hours	12	21.05	18	37.50	6	30.00
2-3 hours	10	17.54	5	10.42	5	25.00
3+ hours	15	26.32	5	10.42	4	20.00
Screen time in bed						
None	23	40.35	26	54.17	6	30.00
Less than 30 minutes	9	15.79	10	20.83	5	25.00
30 minutes to 1 hour	4	7.02	6	12.5	4	20.00
1-2 hours	13	22.81	3	6.25	1	5.00
2-3 hours	3	5.26	0	0.00	0	0.00
3+ hours	5	8.77	3	6.25	4	20.00
Number of messages sent in bed						
None	44	77.19	46	95.83	15	75.00
1-10	7	12.28	2	4.17	3	15.00
11-20	3	5.26	0	0.00	1	5.00
20-30	0	0.00	0	0.00	0	0.00
30+	3	5.26	0	0.00	1	5.00

Number of times awoken by messages in bed						
0	55	96.49	47	97.92	20	100.00
1-5	2	3.51	0	0.00	0	0.00
6-10	0	0.00	1	2.08	0	0.00
10+	0	0.00	0	0.00	0	0.00

Note: The screen-time section of the daily diary is based on Polos et al. (2015) STRICT (sleep time-related information and communication technology). Nine within-subjects participants completed the daily diary in all three school phases. N = the number of observations, % = the percentage of responses. There were 57 observations included in the holiday condition, 48 in the term (school nights), and 20 for term-time weekends; this is represented as the figure in parenthesis next to the school phase labels in the table header.

2.4.2. Adolescent sleep duration during the school holiday and term

Two of the 13 participants have term-time data only, due to COVID-19 preventing the collection of holiday data. Therefore, their data are not included in the subsequent within-subjects analyses.

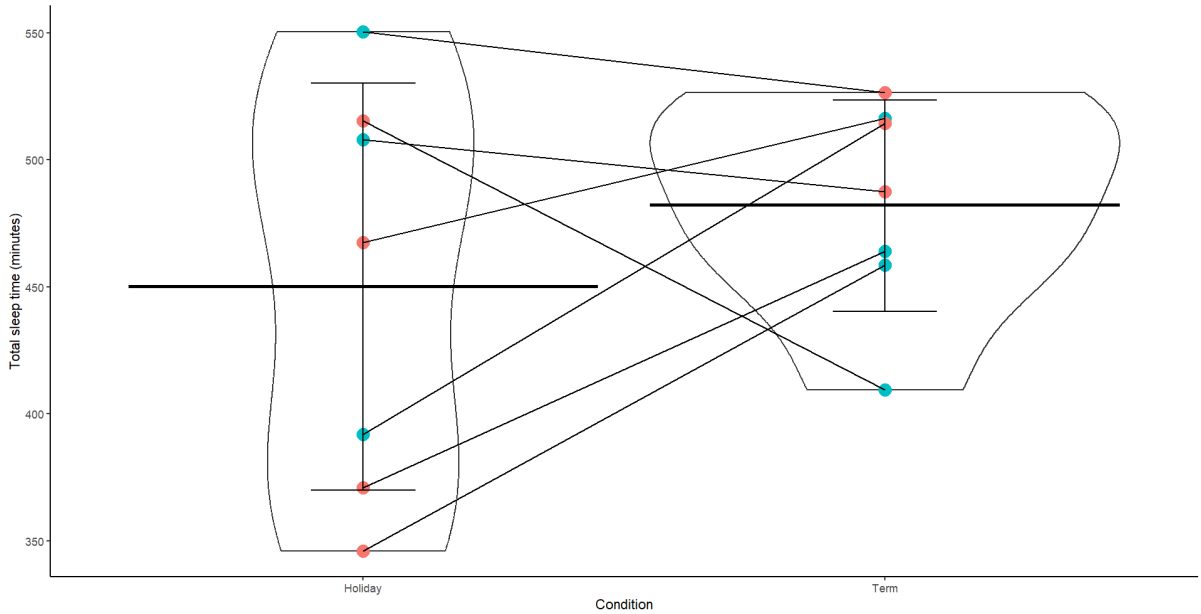
2.4.2.1.Polysomnography

From the remaining 11 participants, only seven participants had within-subject data that are included in the following polysomnography analyses. One participant's data were excluded due to low quality polysomnographic and behavioural data. One participant removed the electrodes and returned to sleep during the term-time polysomnography session, and therefore, their data are excluded from these analyses. Further two participants lost their data during the holiday due to (1) power failure resulting in data not being recorded, (2) due to the equipment and electrodes becoming disconnected.

Figure 2.1. below shows total sleep time across the term and holiday school phases. There was not a significant difference in polysomnographically measured total sleep time

during the school term compared to the holiday, $t(6) = -1.03$, $p = .350$ [95% CI = -111.22 – 46.53]. The mean difference between the term and holiday was -32.34 minutes. All other sleep variables were also quite similar across the term and holiday (see Table 2.3). Contrary to predictions, total sleep time on school nights was longer than on free nights. However, it is important to note, that due to data collection being interrupted by COVID-19, the timings of the polysomnography data collection are not fully counterbalanced. Therefore, the majority of participants (N = 4 out of 7) had their first polysomnography session during the holiday. Figure 2.1. shows that the two shortest sleep durations in the holiday are during the first night of polysomnography, whilst the shortest three sleep durations in the term school phase are on the second night of polysomnography. This highlights the importance of counterbalancing, and potentially familiarisation with the equipment prior to data collection, for future studies to get a more accurate estimate of naturalistic sleep restriction during.

Figure 2.1. Violin plot showing polysomnographically recorded total sleep time during the school term and holiday (N = 7 within-subjects participants)



Note: The mean is represented via the crossbar and standard deviation bars are present. Pink dots indicate the first polysomnography session, blue dots indicate second polysomnography session. Lines join participants first and second session across the school phases.

Table 2.3. Polysomnography monitored sleep variables during the school term and holiday (*N* = 7 within-subjects participants)

Subject	School phase	Efficiency (%)	SOL (mins)	WASO (N=)	WASO (mins)	TST (mins)	S1 (%)	S2 (%)	SWS (%)	REM (%)
1	Term	74.3	138	1	3.5	409.4	0.7	40.5	45.3	13.4
1	Holiday	80.1	106.9	2	20.9	515.5	0.5	42.2	39.8	17.6
3	Term	83.1	481.2	1	31.8	487.5	1.2	55.7	24.7	18.4
3	Holiday	90.2	31.5	1	23.6	508	1.9	64.1	15.7	18.3
5	Term	93.6	23.5	0	8.4	464	0.4	59.6	25.3	14.7
5	Holiday	86.0	47.3	2	12.5	371	1.8	61.9	20.5	15.9
6	Term	91.9	30.7	0	14.6	516.5	0.7	44.1	35.9	19.3
6	Holiday	88.0	52.9	0	10.7	467.5	0.6	42.5	37.1	19.8
7	Term	93.9	20.5	0	9.1	458.5	2.8	46.5	27.5	23.2
7	Holiday	58.4	96.1	2	150	346	0.6	50	29.8	19.7
9	Term	89.4	45.0	1	17.5	526.5	2.3	39.9	34.8	23.0
9	Holiday	85.9	80.2	1	9.9	550.5	0.5	38.1	43.7	17.7
11	Term	94.9	13.2	2	13.7	514.5	0.5	48.3	25.6	25.7
11	Holiday	66.8	15.2	0	179.6	392	0.6	47.2	31.4	20.8
Mean	Term	88.73	43.87	0.71	14.09	482.4	1.23	47.80	31.30	19.67
Mean	Holiday	79.34	61.44	1.14	58.17	450.1	0.93	49.43	31.14	18.54

Note: SOL = Sleep Onset Latency, WASO = Wake After Sleep Onset, TST = Total Sleep Time, S1 = Stage 1 Sleep, S2 = Stage 2 Sleep, SWS = Slow-Wave Sleep, REM = Rapid Eye Movement Sleep. Percentages for sleep stages indicate percentage of total sleep time spent in each stage. N = for WASO indicates the number of awakenings recorded after sleep onset. Mins = minutes. For participants with no awakenings, WASO indicates time spent in bed after the final awakening before disconnecting the equipment or removing electrodes, likely time spent trying to get back to sleep. Sleep efficiency is calculated in REMLogic using the following formula [(SOL + WASO in mins + TST) / TST] *100 this calculation includes time spent awake in bed before disconnecting the equipment.

2.4.2.2.Actigraphy

Eleven participants had within-subject data for all three school phases (term-nights, holiday, and weekend sleep). Polysomnography nights and daytime epochs (Getting into bed between 7 a.m. and 7 p.m., *N* = 15 observations) were removed to ensure only nocturnal sleep on nights undisturbed by polysomnography data collection were included. This left a total of 147 observations. Four outlying data points were identified and removed from the term-time data (6.67%), two from the holiday data (3.28%), and none from the weekend data. Participants had an average of 5.36 nights of actigraphic data (*SD* = 1.63, range = 3-9) during the school holiday, 5.09 nights (*SD* = 3.63, range = 2 - 12) during term-time weeknights, and 2.36 nights (*SD* = 1.36, range = 1 - 6) during the term-time weekend.

An average total sleep time was calculated per participant per school phase (weekend, term, or holiday night). These data were then analysed in a within-subjects ANOVA. On a nightly observation, rather than participant level, the actigraphic data (see Figure 2.2. Panel A and Table 2.4. below) show that total sleep is descriptively longest on the term-time weekend (*M* = 479.08, *SD* = 112.92, range = 261 – 714 minutes), followed by the school-holiday (*M* = 468.88, *SD* = 111.48, range = 208 - 809), with term-time sleep being the shortest (*M* = 438.59, *SD* = 82.04 range = 189 - 586). A total of 37.5% (*N* = 21 out of 56) of the term-time observations were equal to or greater than the recommended 8-hours sleep per night. During the school holiday, this increased to 44.07% (*N* = 26 out of 59), and 50% (*N* = 13 out of 26) on the weekend. Of the nights that sleep duration was identified as restricted by the actigraphy algorithms, the average number of minutes below the recommended duration was 84.2 minutes in the holiday (*SD* = 63.0 minutes), 86.5 minutes on school nights (*SD* = 70.4), and 89.9 minutes on the weekend (*SD* = 71.4).

In a within-subjects ANOVA with 11 participants and three school phases (school night, weekend (term), and holiday), school phase came close to being a significant predictor of

total sleep time, $F(2, 20) = 3.08$, $p = .068$, $\eta_p^2 = .236$. Paired-sample post-hoc t-tests showed a trend toward greater total sleep time in the holiday compared to the term, $t(10) = 2.14$, $p = .058$ (Bonferroni adjusted $p = .174$). There was also a trend toward less sleep on school nights than weekends during the term, $t(10) = -1.96$, $p = .078$ (Bonferroni adjusted $p = .235$). Term-time weekend and holiday sleep were very similar in duration, $t(10) = -0.87$, $p = .405$ (Bonferroni adjusted $p = 1.00$).

Table 2.4. Table showing actigraphy average sleep times ($N = 11$)

	Fall asleep time	Waketime	Total Sleep Time (minutes)	Number of observations included
Term	22:48:25	06:56:56	438.59	56
Weekend	23:49:47	08:31:47	479.08	26
Holiday	23:50:03	08:31:33	468.88	59

Note: Data were removed for the nights that polysomnography data were collected. Daytime epochs and statistically outlying data were also removed from the analyses. The remaining data are from 11 within-subjects participants. Fall asleep time was calculated from bedtimes plus sleep latency. Waketimes were calculated from fall asleep times, plus total sleep, and wake after onset.

2.4.2.3.Daily diary

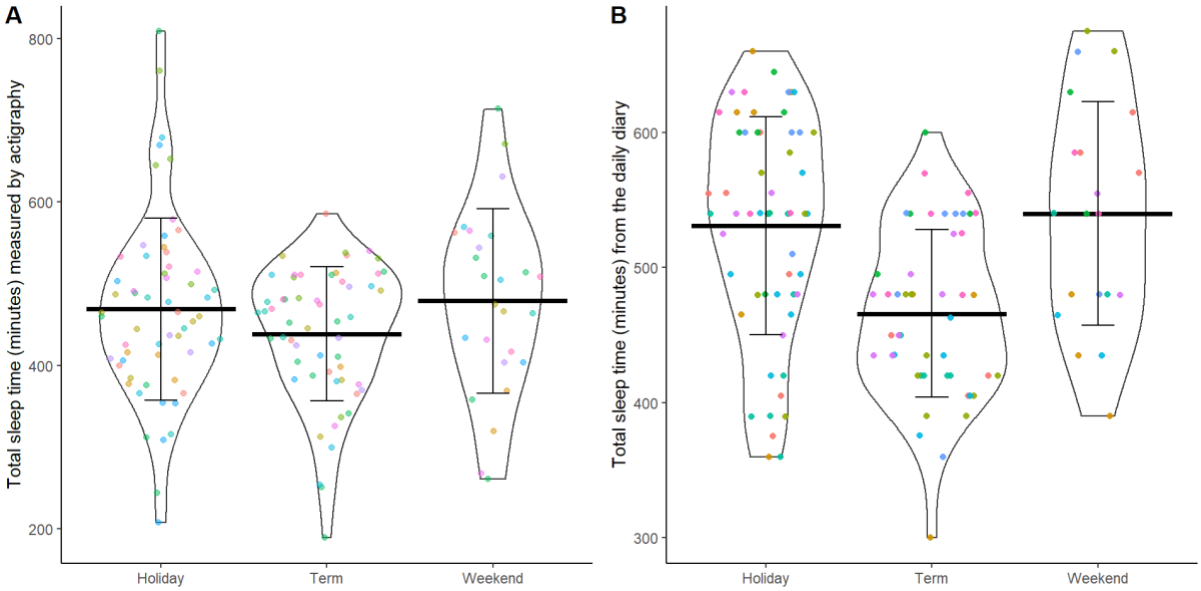
An average of total sleep time was calculated per participant per school phase (weekend, weekday, or holiday night) from the responses on the daily diary. One statistically outlying data point was identified and removed from the term-time school night condition (1.39%), none were identified in the holiday or term-time weekend condition. Nights of polysomnography data collection remained in the diary data, to provide background information (e.g., caffeine, screen-time). Nine participants had included diary data for each of

the three school phases. Total sleep time per school phase can be seen in Figure 2.2. Panel B. Table 2.5. shows average sleep timings.

For these participants, school phase was a significant predictor of total sleep time, $F(2, 16) = 8.22$, $p = .004$, $\eta_p^2 = .507$, with average total sleep time being significantly lower on the term-time weekdays ($M = 463$, $SD = 49.5$) compared to the weekend ($M = 543$, $SD = 73.0$), $t(8) = -3.23$, $p = .012$ (Bonferroni adjusted p = .036), and holiday ($M = 537$, $SD = 43.1$), $t(8) = -3.12$, $p = .014$ (Bonferroni adjusted p = .043). There was not a significant difference between total sleep time during the holiday compared to the term-time weekend, $t(8) = -0.23$, $p = .828$.

The majority of observations during the term-time weekend (80.77%, $N = 21$) reported sleeping equal to or greater than the recommended eight hours per night (480 minutes; Hirshkowitz et al., 2015), similarly to the holiday (78.95%, $N = 45$). During school nights (term-time weekdays) this reduced to 56.34% ($N = 40$). Of the nights that were restricted (under 8 hours total sleep time), the average number of minutes below the recommended eight hours sleep was 72.5 minutes in the holiday ($SD = 37.2$), 65 minutes on school nights ($SD = 34.6$), and 48.8 minutes on the weekend ($SD = 30.9$).

Figure 2.2. Violin plot showing total sleep time per school phase as reported by actigraphic (Panel A) and daily diary (Panel B) data



Note: Crossbar indicates mean, error bars are standard deviation. Coloured data points indicate participants (N = 11 actigraphy participants, N = 9 daily diary participants).

Table 2.5. Self-reported average sleep timing from the daily diary (N = 9 participants)

	Bedtime	Waketime	TST (mins)	Number of observations included
Term	22:23:45	06:56:53	466	71
Weekend	23:15:45	09:02:15	540	26
Holiday	23:21:14	08:41:51	531	57

Note: Data presented in the table are within-subjects from nine participants. Averages were calculated at the observation (i.e., nightly), rather than aggregated at participant level.

2.4.3. Behavioural performance on the non-word learning paradigm

Nine participants were included in the behavioural analyses from the "cathedruke" non-word learning paradigm. Two participants only had term-time data due to COVID-19 preventing collection of the complete datasets, and two participants were missing trials or had

low quality data (e.g., due to environmental interruptions during the task). Each training session (using a 22-item list of non-words) took an average of 36.84 minutes (*SD* = 9.58, range = 24.45 – 59.65 minutes). Each testing session (made up of recall, lexical competition and old/new recognition tasks) took approximately 21.81 minutes (*SD* = 2.67, range = 18.58 – 32.53).

2.4.3.1. Lexical competition

Larger differences in reaction times between base words and control words (higher lexical competition scores) are indicative of slower responses in identifying whether the base word is a real word or non-word. Slowed reaction times suggest that the novel word (e.g., "cathedruke") has been integrated into the mental lexicon, and therefore, is acting as a competitor for the base word. Eleven within-subjects participants were included in these analyses. Reaction times below 300 milliseconds and above 2500 milliseconds were removed. Only accurate responses were included in the analyses. Reaction times were log-transformed. The log-transformed reaction time for base words were subtracted from the log-transformed reaction times for control words in each school phase (holiday and term-time weeknight [i.e., term]) and at each test time (immediate and delayed) in a 2 x 2 design (see Figure 2.3. Panel A below). Descriptively, lexical competition (in regard to raw reaction time differences) was greater in the holiday for both the immediate (*M* = 61.4 milliseconds, *SD* = 126 milliseconds) and delayed test times (*M* = 61.4 milliseconds, *SD* = 74.5 milliseconds), compared to the term. Lexical competition was lowest during the term-time immediate test (*M* = 31.0 milliseconds, *SD* = 57.9 milliseconds). However, after sleep, lexical competition increased slightly in the term after a period of overnight sleep consolidation (*M* = 46.5 milliseconds, *SD* = 89.9 milliseconds). Perhaps better sleep before the holiday training

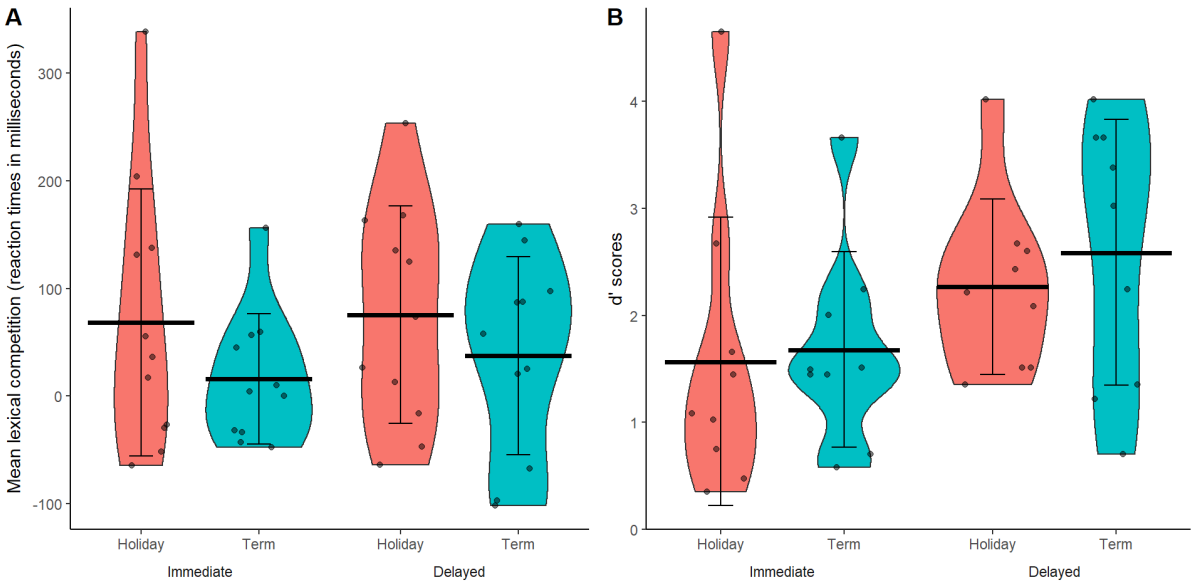
contributed to better performance, whilst overnight sleep consolidation during the term allowed for improvements in performance from the immediate to delayed test.

Statistically, there was no significant main effect of school phase (term versus holiday), $F(1, 8) = 0.21$, $p = .209$, $\eta_p^2 = .153$. There was also no significant main effect of test time (immediate versus delayed), $F(1, 8) = 0.77$, $p = .400$, $\eta_p^2 = .072$. The interaction between condition and test time was also not significant, $F(1, 8) = 0.14$, $p = .714$, $\eta_p^2 = .014$.

2.4.3.2. Old/New Recognition

D-prime (d') scores were calculated using the NORMINV function in excel. The mean was set to zero and the standard deviation was set to one, in order to standardise the data (z-transformed). The z-transformed false alarms (i.e., responding "trained" to a foil) were subtracted from z-transformed hits (i.e., responding "trained" to a trained word). Nine participants were included in the analyses. From the original 13 participants, two were removed as they only had data in one school phase. A further two participants were missing at least one session worth of data (e.g., due to E-Prime errors). There was no significant main effect of condition (term versus holiday), $F(1, 8) = .55$, $p = .479$, $\eta_p^2 = .011$. There was not a significant interaction between condition and test time, $F(1, 8) = 0.26$ $p = .627$, $\eta_p^2 = .003$. However, there was a significant main effect of test time, $F(1, 8) = 10.12$, $p = .013$, $\eta_p^2 = .130$, with participants performing significantly better in the delayed than the immediate test (see Figure 2.3. Panel B below).

Figure 2.3. Behavioural data across school phases (holiday versus term) and test time (immediate versus delayed) for lexical competition (Panel A) and old/new recognition (Panel B) tasks



Note: Crossbar indicates the group mean, error bars are standard deviation. Scatter points indicate individual participants. *N* = 11 in Panel A, *N* = 9 in Panel B.

2.4.3.3. Recall

No analyses were conducted on recall data due to the use of various methods (laptop-led cued recall, free recall, researcher-led cued recall), which would not be comparable across participants. Future research should identify the most practical solution for their sample (e.g., adolescents) and data collection method (e.g., in person or remotely).

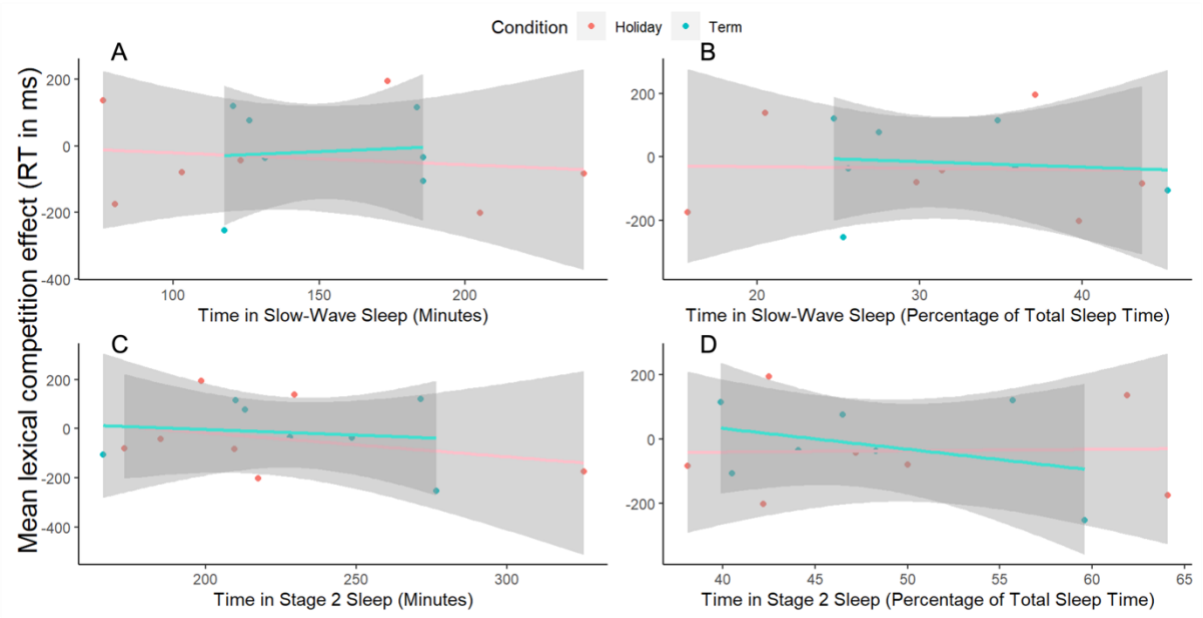
2.4.4. Correlations between sleep architecture and behavioural performance

A total of seven participants had within-subjects polysomnography data (see Section 2.4.2.1. above regarding exclusion of data). Therefore, the analyses investigating correlations between sleep architecture and behavioural performance (described below in Section 2.4.4.1. and 2.4.4.2.) are conducted with these seven participants.

2.4.4.1. Lexical Competition

The difference between lexical competition in the delayed and immediate test was used as an indicator of consolidation effects in lexical competition. Positive values indicate greater lexical competition in the delayed test (i.e., after sleep consolidation) compared to the immediate test. There was not a significant correlation between time spent in slow-wave sleep (in minutes, Figure 2.4. Panel A) and lexical competition, $r(12) = -.07$, $p = .589$, 95% *CI*[-0.51 – 1.00], or proportion of total sleep time spent in slow-wave sleep (Figure 2.4. Panel B), $r(12) = -.06$, $p = .576$, 95% *CI*[-0.50 – 1.00]. There was also not a significant relationship between number of minutes spent in Stage 2 sleep (Figure 2.4. Panel C), $r(12)= -0.24$, p = .793, 95% CI[-.63 – 1.00] or proportion of total sleep time spent in Stage 2 sleep (Figure 2.4. Panel D), $r(12) = -.14$, $p = .685$, 95% CI[-0.56 – 1.00], and lexical competition.

Figure 2.4. Correlations between lexical competition and minutes spent in slow-wave sleep (Panel A) or Stage 2 sleep (Panel C) and percentage of total sleep time in slow-wave sleep (Panel B) and Stage 2 sleep (Panel D)

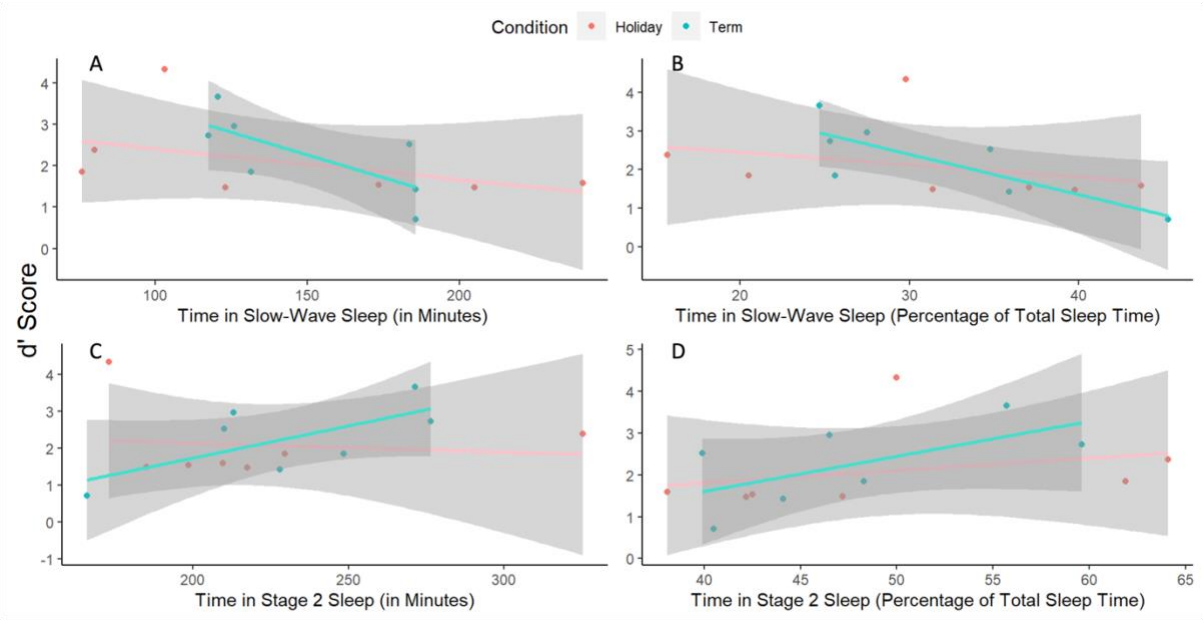


Note: RT = Reaction Times, ms = milliseconds

2.4.4.2. Old/New Recognition

There was no significant correlation between time spent in slow-wave sleep (in minutes, Figure 2.5. Panel A) and d' scores, $r(12) = -.51$, $p = .969$, 95% *CI*[-0.79 – 1.00], or proportion of time spent in slow-wave sleep (Figure 2.5. Panel B) $r(12) = -.51$, $p = .970$, 95% *CI*[-0.79 – 1.00]. Descriptively, there was a larger negative trend between slow-wave sleep and d' accuracy in the term ($r = -.72$ [slow-wave sleep minutes], -.80 [total sleep time]) than in the holiday ($r = -.46$, -.31 [total sleep time]). There was also no significant relationship between minutes in Stage 2 sleep and d' scores (Figure 2.5. Panel C), $r(12) = .23$, $p = .219$, 95% *CI*[-.26 – 1.00], nor proportion of total sleep time spent in Stage 2 (Figure 2.5. Panel D), $r(12) = .41$, $p = .073$, 95% *CI*[-0.06 – 1.00]. Descriptively, there was a larger positive trend in the term ($r = .67$ [minutes], .62 [total sleep time]) than in the holiday ($r = -.12$ [minutes], .28 [total sleep time]).

Figure 2.5. Correlations between performance on the old/new recognition task (d') and number of minutes spent in slow-wave sleep (Panel A) and Stage 2 sleep (Panel C), or percentage of total sleep time spent in slow-wave sleep (Panel B) and Stage 2 sleep (Panel D)



2.5. Discussion

The current study set out to investigate whether adolescents (aged 13-16-year-olds) in the UK are significantly sleep restricted on school nights. We also aimed to explore the effects of adolescent sleep restriction on sleep architecture, under naturalistic conditions, as previous research has used laboratory sleep restriction paradigms to examine this relationship (e.g., Ong et al., 2016). The research questions related to adolescent sleep restriction are discussed in Section 2.5.1. below. The second research question investigated whether sleep restriction was correlated with behavioural performance on a non-word learning paradigm (see Section 2.5.2. for discussion). This pilot study showed that using polysomnography in participants' homes on a school night and on a free night (i.e., the school holiday) appears to be a feasible methodology. However, some suggestions are made in Section 2.5.3. for ways to improve recruitment and methodology in future research. It is important to note throughout

that the sample size is small, hence the pilot nature of the study, and therefore the statistical analyses are underpowered and limited. Therefore, it is important to consider the conclusions from this pilot study tentatively.

2.5.1. Sleep restriction

Contrary to our hypotheses, the polysomnography data did not show a significant difference in total sleep time during the school term compared to the school holiday. Instead, mean total sleep time was approximately 30 minutes *shorter* during the holiday than the term (see Figure 2.1. and Table 2.3.). However, wake after sleep onset was much higher during the holiday (average 58 minutes) than the term (14 minutes). Sleep latency was also longer and sleep efficiency lower during the holiday (61 minutes, 79%) than the term (44 minutes, 89%). Together, this may suggest greater levels of sleep disruption and poorer quality sleep during the holiday polysomnography session than the term, which may be explained by methodological limitations such as counterbalancing issues due to COVID-19 meaning that most participants (4 out of 7) had their first polysomnography session in the holiday (see Section 2.5.3. and Figure 2.1.). Alternatively, adolescents may be experiencing lower sleep pressure during the holiday, as the actigraphy and diary data suggest they are receiving more sleep on free nights than during school nights. Therefore, during the holiday, participants may have been less tired and less likely to fall asleep immediately (e.g., increasing sleep latency).

The daily diary showed significantly longer sleep on the weekends and holiday than on school nights. The majority of observations in the holiday and on term-time weekends were reported to be of sufficient duration, whilst just over half of observations achieved of exceeded the recommended 8-hours sleep (recommended by the National Sleep Foundation; Hirshkowitz et al., 2015). This may be explained by the subjectivity of self-reports leading to

over-reporting free-night sleep and under-reporting school night sleep duration. However, the objective actigraphic data shows a similar trend towards longer total sleep time in the holiday and weekend than on school nights, which may be significant in larger samples. The number of observations reaching or exceeding the 8-hour sleep duration was much lower for the actigraphy than the diary data. Actigraphic data also revealed descriptive evidence of weekend compensatory sleep, given that sleep duration was longer on the term-time weekend than on the school or holiday nights. This highlights that sleep during the term-time weekends may be compensatory for school night sleep restriction, whilst free-sleep during the holiday may demonstrate sleep need. Multi-night recordings of sleep may be beneficial for identifying the extent of adolescent sleep restriction (suggested methodologies are discussed further, in Section 2.5.3.).

It is also worth noting that the actigraphic data showed a smaller difference (30 – 42 minutes) between free nights and school nights than the self-reported diary (75 minutes), which may be explained by actigraphy devices recording wake after sleep onset that may not be identified by participants in subjective reports. Additionally, actigraphy algorithms have been found to overestimate wake after sleep onset for adolescents aged 10-14-years-old (Quante et al., 2018). There were also a higher proportion of nights below the recommended eight-hours sleep on weeknights compared to the holiday and the weekend, across both the diary and actigraphy data. These findings support previous research suggesting that weekend sleep is longer due to compensating for sleep restriction occurring during school nights (Gradisar et al., 2011). It is important to continue exploring the extent of adolescent sleep restriction in future research, due to the numerous potential consequences (e.g., for academic grades, mental health; Sun et al., 2019). The findings of the current study also highlight the importance of collecting data across multiple nights of sleep for each participant, to examine effects across the week, including on weekends. This could have implications for

understanding individual differences in the relationships between sleep restriction and cognition (e.g., Lo & Chee, 2020) as well as examining intraindividual night-to-night sleep variability or regularity (e.g., Becker et al., 2017).

The longer sleep duration identified by multi-night recordings on free nights may be best explained by later wake-times relative to school nights. The actigraphic data showed average bedtimes of 11:50 p.m. and wake times of 8:30 a.m. during the holiday and weekends. On school nights, bedtimes were approximately one hour earlier, while wake-times were greater than one-and-a-half hours earlier, than during free nights. This suggests that school start times may have pushed adolescents preferred circadian rhythms earlier than their preferences (seen on free nights). However, the later bedtimes during free nights could be associated with psychosocial factors. For example, participants reported a higher level of caffeine consumption (87 – 90 milligrams) on free nights than school nights (39 mg). Over half of the diary entries (54%) during the term reported no time spent on screens in bed, this was lower during the holiday (40%) and the weekend (30%) suggesting greater levels of screen-time at in bed during free nights. Together, the increased caffeine intake and screen usage during the holiday and term may have contributed to delayed bedtimes and consequently delayed wake-times. Research has shown that increasing adolescent age is associated with increased autonomy over bedtimes and lower parental restriction over sleep (Gangwisch et al., 2010; Tashjian et al., 2019). The findings in the current study may also be related to increased autonomy on free nights than school nights, however, this is speculative, as we did not collect measures of parental monitoring. Therefore, the role of parental monitoring versus autonomy during school nights compared to free nights would need to be investigated in future research. There was also greater variation in total sleep time on free nights compared to school nights in the polysomnography and actigraphy data (Figure 2.1., Figure 2.2. Panel A.). Reduced variability during school nights may indicate more restriction

over bedtimes (e.g., parental monitoring) and waketimes (e.g., school start-times) on school nights. Multi-night recordings could be used to examine intraindividual night-to-night sleep variability, which may also be associated with increasing age, pubertal development, and social changes such as greater autonomy (Becker et al., 2017). Individual participants' night-to-night variability in sleep duration has been associated with poorer academic attainment (Phillips et al., 2017) and therefore, may show interesting relationships with experimentally measured cognitive performance.

In experimental studies manipulating sleep duration, sleep restricted adolescents showed restriction in the duration of all sleep stages except slow-wave sleep (e.g., Ong et al., 2016). Despite the current study not being able to identify term-time sleep restriction in the polysomnography data, likely due to methodological issues (see Section 2.5.3), a similar preservation of slow-wave sleep is observed. In the present study, slow-wave sleep made up approximately one third of the sleep cycle, yet there was only 0.16% difference in proportion of time spent in slow-wave sleep across term and holiday school phases (see Table 2.3.). For five of the seven participants in the current study, the shorter the total sleep time, the longer the proportion of slow-wave sleep. Together, this suggests that slow-wave sleep was preserved or prioritised when sleep duration was shorter, supporting experimental sleep manipulation studies (e.g., Ong et al., 2016). The largest differences in proportion of total sleep time were seen for Stage 2 and REM sleep (1.63%, 1.13% difference respectively). This suggests that future research using a naturalistic design (as in the present study) may be able to replicate results seen in the experimental literature, regarding changes in the sleep architectural stages during sleep restriction. As such, this implies that cognitive deficits associated with sleep restriction may not be explained by lack of slow-wave sleep, which contradicts predictions made by the synaptic homeostasis hypothesis (Tononi & Cirelli, 2003, 2014). Furthermore, these relationships between slow-wave sleep and cognition may vary

developmentally, given the changes in slow-wave sleep observed in the transition from childhood to adolescence (see discussion in Chapter 1, Section 1.4.).

2.5.2. Sleep and cognition

This pilot study is the first time, to our knowledge, that the well-established "cathedruke" non-word learning paradigm has been used with adolescents. Previous research with adults and children suggests that non-word learning is sensitive to sleep (e.g., Chatburn et al., 2014; Henderson et al., 2012). In the present study, the behavioural data alone indicated some trend level differences that may be significant with larger samples, outside of a pilot study. In the immediate test (before sleep), lexical competition was greater in the holiday compared to the term. After sleep, lexical competition remained stable in the holiday, however, lexical competition increased slightly in the term. The lack of sleep consolidation effects in the holiday may be due to greater disruptions to sleep during the holiday (as described in Section 2.5.1.). However, it is also worth noting that there appears to be a wider spread of lexical competition data (see Figure 2.3. Panel A) in the holiday compared to the term school phase, which is similar to the greater variability in objectively measured sleep duration during holiday compared to school night sleep (Figure 2.1.). Nonetheless, the term-time data demonstrated overnight increases in performance from immediate to delayed test, similar to findings in adults and children, suggesting that the "cathedruke" non-word learning paradigm is suitable for use in the adolescent population. Greater overall performance on the lexical decision task in the holiday compared to the term school phase may be associated with cumulative effects of multiple nights of shorter sleep duration in the school term than on free nights, as evidenced by the diary and actigraphic data. The role of sleep *prior* to encoding is explored further in Chapter 3, which also uses the "cathedruke" non-word learning paradigm with 18-24-year-olds.

The within-subjects design used in the current study benefits from controlling for individual differences in reaction times, word learning ability, and sleep. However, there may be possible practice effects from repeating the test tasks, and further from the feedback received in the lexical competition task. Some research has suggested that lexical competition emerges only after a delay, whether that is filled by wake or sleep. In Lindsay and Gaskell's (2013) research, they examined the effects of spaced exposure and retrieval practices, for example, having a single intense training session followed by spaced lexical competition testing sessions throughout the day versus spacing exposure and retrieval of non-words throughout the day and only testing for lexical competition once at the end of the day. However, the repeated measures designs used by Lindsay and Gaskell (2013) and the current study may, at least partially, be explained by practice effects either from repeated exposure or testing. However, the results of the present study appear to show no change from the immediate to the delayed test in the holiday. Perhaps this task may have reached a ceiling in the holiday for the adolescents in the present study. Nonetheless, it may be useful for future research to employ a between-subjects design, in which participants complete only one test session, either immediately after a single training session or after a delay with no additional exposure or resting sessions. Such a design may be better able to separate the impact of a delay versus potential practice effects for the emergence of lexical competition.

The old/new recognition task did not show any differences between the holiday and term school phases. However, performance was significantly better in the delayed than the immediate test, which may also be explained by practice effects. It is also important to note that the lack of counterbalancing, due to COVID-19 disruptions, in this pilot study also extends to the allocation of lists learned in the "cathedruke" non-word learning paradigm. Therefore, these relationships will likely have been impacted by differences in the properties of the lists (e.g., memorability or familiarity of base words in adolescent populations), despite

the lists being carefully matched on other criteria (e.g., number of phonemes). Further pilot testing of the "cathedruke" non-word learning paradigm would be beneficial to determine the reliability and validity of this paradigm in adolescent samples.

When considering the correlations between sleep architecture and lexical competition, a trend towards a negative correlation between slow-wave sleep duration and lexical competition was observed. Descriptively, there was also a negative correlation between slow-wave sleep duration and performance on the old/new recognition task. This is contrary to the synaptic homeostasis hypothesis, which would predict greater improvements in behavioural performance in line with increases in slow-wave sleep duration (Tononi & Cirelli, 2003, 2014). When paired with the preservation of slow-wave sleep, discussed in Section 2.5.1. above, this implies that the proposal made by the synaptic homeostasis hypothesis may require further scrutiny. It is important to note that in the current study, the sample size is small, and the counterbalancing is inadequate, due to disruptions caused by COVID-19. Therefore, it is important to interpret these data and conclusions tentatively. However, other research has also demonstrated that the relationship between slow-wave sleep and cognition may not be as robust as previously believed (Cordi & Rasch, 2020). Together, this highlights a greater need for open science practice and scrutiny of the current leading theories of sleep and cognition.

There was a trend towards a positive correlation between proportion of total sleep time spent in Stage 2 sleep and participant's average performance on the recognition task. Perhaps Stage 2 sleep may be involved in accurate old-new categorisation. Previous research has found a correlation between slow-wave sleep duration and faster reaction times in the old-new categorisation task following a period of sleep (Tamminen et al., 2010). Other research has not found a significant effect of sleep deprivation compared to sleep on recognition memory (Tamminen et al., 2020). Inconsistencies in the old-new recognition literature may

be explained by the use of explicit recollection or implicit familiarity processes, which may be employed to complete this task, and which may rely on different brain structures (Brown & Aggleton, 2001). As such these different processes may be related to different neural mechanisms during sleep, resulting in mixed findings. Explicit recall tasks, such as cued or free recall would better be able to answer questions regarding implicit versus explicit processes, as suggested by previous authors (Tamminen et al., 2020). However, a recent meta-analysis of sleep-deprivation literature found that differences in tasks (e.g., recognition or recall) did not significantly moderate the meta-analytic effect size between sleep and memory, despite greater inconsistencies in findings for the recognition than the recall literature (Newbury et al., 2021).

There were also descriptive trends toward negative correlations between Stage 2 sleep duration and lexical competition in both the holiday and term-time data. Sleep spindles, particularly in Stage 2 sleep, are the proposed mechanism behind the interleaving of new and old information that allows lexical competition effects to emerge in the active systems consolidation models. However, given the small sample size, the current study did not analyse sleep spindles. Exploratory analyses showed a trend towards a positive correlation between REM sleep duration and lexical integration. Typically, sleep spindles are observed in Stage 2 and slow-wave sleep, but not REM sleep (Diekelmann & Born, 2010). Therefore, alternative processes may be occurring during REM sleep that may also allow the integration of new and old information. Additionally, REM sleep is less common during daytime naps, whilst slow-wave sleep and Stage 2 sleep are much more frequent during napping (e.g., Studte et al., 2017). As such, daytime naps may be insufficient for compensating for any cognitive effects associated with REM sleep. Therefore, a daytime nap as an intervention to compensate for adolescent sleep restriction, as suggested by Cousins, van Rijn, et al. (2019), may have some benefits, but may not benefit all tasks equally. Further research into the

correlations between adolescent sleep restriction, sleep architecture, and cognition could inform our understanding of effective interventions.

2.5.3. Methodology and future research

The current pilot study could be used to investigate questions related to the impact of current school start times on adolescent sleep. The design of the current research provides naturalistic evidence regarding the extent of sleep restriction for UK pupils, and the consequences of such sleep restriction for learning and memory. This naturalistic design has the ability to validate experimental laboratory sleep studies that manipulate sleep duration (e.g., Cousins et al., 2018; Ong et al., 2016). Other studies have used actigraphy to examine sleep restriction (Baum et al., 2014; Phillips et al., 2017). However, actigraphy cannot inform about sleep architecture. Self-reported sleep and surveys have also demonstrated an impact of school on sleep restriction (De Souza et al., 2012; Gradisar et al., 2011), however, these are subjective reports that may allow greater biases. The current design of this pilot study allows a balance between experimental rigour and naturalistic observation. Such a design may have implications for policy and practice in higher powered and effectively counterbalanced future research. The outcome of a petition debated in parliament was that more research was required to allow schools to confidently make the decision to delay school start times (Petitions Committee, 2019; Zeichner, 2019). However, there is currently a reluctance from schools in the UK to participate in randomised controlled trials shifting school start times (Illingworth et al., 2018), despite efficacy of other such trials in countries such as the USA (Marx et al., 2017). Therefore, future research using a similar study design to this pilot would be able to provide high quality evidence to inform about the impact of adolescent sleep restriction on cognition, which could have implications for school start time policies. If disseminated effectively, such research could increase the possibility of future randomised controlled trials delaying school start-times.

Future research could ensure effective translation to policy and practice through co-production of research questions, methods, and dissemination in inter-disciplinary teams including school-staff, parents, policy makers, scientist-practitioners (e.g., educational psychologists), and researchers. Such inter-disciplinary teams should include different people with a range of complementary skills to facilitate success (Solari et al., 2020). The lessons learned in the current pilot study could inform the success of future research using a similar design. Potential methodological considerations include further piloting of the non-word learning paradigm in adolescent samples; combinations of objective and subjective sleep measures over multiples nights; COVID-19 disruptions and counterbalancing; recruitment barriers; and further consideration of background measures. All of these considerations are discussed in more detail below.

As mentioned in section 2.5.2., this pilot study was the first to our knowledge that used the "cathedruke" non-word learning paradigm with adolescent participants. Response rates were poor when the cued recall task was delivered via the laptop and headphones, with participants controlling when they move on to the next item. Free recall with a time limit was also more time-consuming, and participants often reported that it could feel uncomfortable or awkward. Researcher led cued recall with scripted encouragement appeared to be the best method for this adolescent sample, although this may be less practical in other settings, in which a researcher may not be present, such as online studies. However, future research would also be needed to validate the "cathedruke" non-word paradigm for online research. As the methods varied within this pilot study, recall data were not analysed. Participants completed all tasks within their home or school, which may explain some issues in obtaining recall data as adolescents may have felt embarrassed saying non-words within these uncontrolled environments. Additionally, participants sometimes pressed other buttons on the keyboard than the ones specified in the instructions; this resulted in E-Prime freezing and

requiring restarting on occasions. It is worth noting that these incidents of technical issues were typically during the term-time data collection, which may have biased the results. Nonetheless, this piloting indicated the importance of the testing and training scripts being adapted to account for non-anticipated key presses, or to use a millisecond accurate button box or similar equipment with limited keys.

Regarding measurement of sleep, there was reasonable adherence to wearing the actigraphic watches for approximately one week (M = 5.36 nights during the holiday, M = 5.09 for term, M = 2.36 for weekend). However, the brand of actigraphic watches used in the current study was a more prominent design, which anecdotally made participants aware that they were identifiable by peers as participating in the research. Alternative devices that are more discrete may increase adherence and reduce barriers to recruitment. The actigraphic and diary results both demonstrate sleep restriction on school nights and greater sleep duration on free nights, but to differing extents (as discussed in Section 2.5.1.). This emphasises the importance of considering both objective and subjective measures of sleep together (e.g., Dewald et al., 2010), and the potential benefit of multi-night recordings.

The current study used single night recordings of polysomnographic data during the holiday and term. Therefore, testing was only possible at the end of the school week, after three consecutive nights of school night sleep patterns. Additionally, data could only be collected with one participant per night, due to requiring researcher presence to set-up the equipment in the participants' homes. Together, this allowed a maximum of two participants per week during term-time. One feasible way to address this issue would be using multi-night sleep monitoring technologies such as actigraphy watches. However, actigraphy does not inform about sleep architectural correlates of sleep restriction. Alternatively, Dreem headbands, a dry-electrode, commercial sleep architecture monitoring technology could be used in future research. Dreem headbands can be worn for multiple nights and do not require

researcher presence within participants' homes. They are also compact and could be delivered and collected from participants' homes or schools with ease. This would reduce the burden on family's time and space and could allow data to be collected with multiple participants at the same time by a single researcher. Dreem headbands may also be preferrable as they are only worn at night, unlike actigraphy which is worn day and night, reducing the risk of participants becoming identifiable as taking part in the research. Multi-night polysomnographic recordings could answer additional research questions regarding sleep restriction over the course of an entire school week, including weekend compensatory sleep and intraindividual night-to-night sleep variability.

Multi-night sleep recordings could also reduce disruptions caused by insufficient counterbalancing. The current study was disrupted by COVID-19 lockdowns. In the current study, four out of seven participants had their first night of polysomnography data collected in the holiday. This may have contributed to shorter sleep duration and greater disruption in sleep during the holiday than the term (discussed in Section 2.5.1.). Familiarisation with the equipment and multi-night recordings could reduce this large effect of counterbalancing on the results. This would be resolved using Dreem headbands, which would further reduce face-to-face contact with a researcher – a significant barrier during the COVID-19 pandemic due to social distancing and lockdown restrictions. An additional measure to reduce face-to-face contact could be conducting training and testing sessions online. However, these tasks were time consuming *(M* = 160.92 minutes per school phase); therefore, issues associated with online testing (e.g., internet speed, participant attention) may be more likely to occur. The researcher may be able to offer video call support during testing to support adherence and engagement. Alternatively, the tasks could be shortened, with appropriate piloting. These alternative methods were considered to complete the present study. However, due to the loss

of education time during the COVID-19 lockdowns, it was not appropriate to restart the research due to this time burden on schools and participants.

Counterbalancing of word lists was also a limitation in the present study due to disruptions caused by COVID-19 and technological failures resulting in loss of data. This pilot study acknowledges the importance of effective counterbalancing for the "cathedruke" non-word learning paradigm. This paradigm was selected due to its previously established relationships with sleep. The types of learning and memory processes assessed by this paradigm (e.g., integration of new and information) may be useful for interested parties (e.g., students, educational professionals) and psychological theories such as the interleaving process in the complementary learning system account (McClelland et al., 1995). However, implicit learning (James et al., 2021) and consolidation processes (Henderson et al., 2015) associated with reading stories, and studying factual information (Cousins, Wong, & Chee, 2019; Cousins, Wong, Raghunath, et al., 2019) may be more naturalistic and similar to processes occurring in classroom settings.

More naturalistic tasks may also facilitate recruitment, which was challenging in the present study. Initially we anticipated that adolescents would be interested in volunteering for the study. However, only 2-5% of students signed up to participate after each assembly delivered to a secondary school year group. The barriers to recruitment were not solely due to student interest levels, but also due to requiring permission from families and schools. Students in assemblies and smaller workshops seemed interested in the research, yet they did not return consent forms. Online consent forms may reduce administrative burden and increase uptake. Families were also hesitant to have researchers in their homes, however, it was not possible to meet with families who did not participate in the research to further explore this barrier. Recruitment strategies targeting parent/guardian(s) may be beneficial, as the families of the participants were often enthusiastic and interested in sleep, psychology, or

science more generally. A video information sheet, rather than a long, written document, may also be useful to explain the study to potential participants. Such videos could also be an opportunity to familiarise participants and families with the equipment, tasks and the researcher. The initial design of the study aimed to complete testing as close to participants waking up as possible, meaning participants missed their first lesson during the term-time delayed test. An alternative study design may be that participants complete the delayed test after school and that this time is matched in the holiday test. This requires school permission, which reduces the opportunity to collect data via word-of-mouth, social media, or other avenues outside of schools.

Recruitment challenges could also be mitigated by conducting collaborative research. For example, using focus groups comprised of young people, families, and teachers, could identify barriers and facilitators to participation and improve the study design. Collaborative research could ensure the information and recruitment materials were understandable and appealing and could identify suitable study designs for schools, families, and young people. When discussing recruitment with a small group of secondary school students, they suggested a guaranteed payment of £15 cash upon completion of the study would support recruitment. However, even with an additional prize draw for £25 vouchers, recruitment was still challenging. Future research would benefit from involving a range of students in the study design and discussing suitable compensation for the considerable time commitment and intrusive nature of collecting data within participants' homes.

Future research may also wish to evaluate the mechanisms behind adolescent sleep restriction, which may be informed by collaborative research and theoretical models (e.g., Becker et al., 2015; Crowley et al., 2018). Potential mechanisms associated with adolescent sleep restriction that were considered in the current study include caffeine consumption, screen-time, chronotype, and pubertal development. Additionally, we included working

memory ability, vocabulary knowledge, and non-verbal reasoning which we planned to explore as covariates for learning and memory effects. The self-report caffeine intake questionnaire allowed calculation of estimated milligrams of caffeine consumed. This could be improved in future with clearer reporting of portion sizes for caffeinated products, to allow greater accuracy in estimation. Additionally, future research could measure screen-time objectively through applications on mobile phones. However, this would not account for screen-time on other devices (e.g., television, computer, gaming). Therefore, diary measures of screen-time were appropriate for the current study, as these were taken as background measures only. As discussed in Section 2.5.1. caffeine intake and screen-usage increased on free nights compared to school nights. Parental monitoring may have influenced these relationships, a potential mechanism further effecting sleep duration (e.g., Gunn et al., 2019; Hsu et al., 2012; Tashjian et al., 2019), which may be of interest for future research.

2.5.4. Conclusions

The current study identifies an appropriate methodology to provide high quality naturalistic evidence regarding adolescent sleep restriction. Using this methodology in higher powered and effectively counterbalanced research in the future may have implications for policy and educational practice such as the school start time debate. The preliminary results of this pilot study show that such methods will likely support experimental laboratory research regarding the sleep architectural correlates of insufficient sleep duration. In particular, the relative preservation of slow-wave sleep was observed in this pilot study, supporting Ong et al.'s (2016) experimental research. Together, with the preliminary behavioural data (Section 2.5.2.), the predictions made by the synaptic homeostasis hypothesis were not supported and will require further scrutiny with robust methodology. Section 2.5.3. makes suggestions for improving the study design for future research, including the use of multi-night sleep recordings, and further investigation of the

appropriateness of the non-word learning paradigm in adolescent samples. Adolescent sleep is increasingly being found to be relevant for learning and memory, with implications for schools and educators. However, large scale changes to school scheduling (e.g., delayed school start times) may still encounter barriers, but may be supported by more research using a similar design to the present study. Instead, smaller scale interventions, such as allowing nap opportunities during the school day to recover from sleep restriction and benefit learning and memory may be more easily implemented (e.g., Cousins, van Rijn, Ong, et al., 2019; Cousins, Wong, Raghunath, et al., 2019). The next chapter (Chapter 3) explores the role of a daytime nap preceding encoding. Collaborative research could help to identify the most feasible interventions for schools, families, and young people, and the potential barriers and facilitators to such interventions and could mitigate difficulties implementing interventions.

Chapter 3: Is word learning capacity restored after a daytime nap? A registered report

3.1. Introduction

3.1.1. Sleep consolidation

As described in Section 1.2.1., sleeping after learning is thought to be beneficial for the consolidation of newly encoded declarative memories. According to active systems consolidation models, based on complementary learning systems accounts of memory (McClelland et al., 1995; Diekelmann & Born, 2010) new memories are encoded in parallel in the fast-learning hippocampus and in the slow-learning neocortex during wake. Then, during sleep, the hippocampus repeatedly replays new memories and communicates with the neocortex. Through this hippocampal-neocortical dialogue, the neocortex gradually extracts regularities from overlapping representations of new memories, as well as integrating those memories with existing knowledge stored in long-term memory. The active systems consolidation theory therefore predicts that sleep after learning should not only benefit episodic memory (e.g., for the details of the experience encoded during wake) but should also benefit more complex memory processes such as memory integration and generalisation (Witkowski et al., 2020). Such sleep consolidation is explored in Chapter 2, in relation to adolescent sleep and school start times.

Much of the empirical evidence, particularly for the role of sleep in more complex memory processes, comes from language learning studies. For example, Dumay and Gaskell (2007) trained participants on 24 novel spoken words (e.g., “shadowkt”) before sleep. They found that one night of sleep after learning increased free recall of the words, whilst an equivalent period of wake after learning yielded no change from baseline. Importantly, the novel words were all highly phonologically overlapping with familiar existing base words

(e.g., “shadow”). This allowed the authors to quantify to what degree the newly learned words interfered with recognition of the familiar base words. That is, to what degree the novel words had been integrated with existing words in the mental lexicon. This lexical competition effect was only observed after a night of sleep. While this study supported the hypothesis that sleep supports consolidation of episodic memory as well as the integration of new memories with existing memories, subsequent studies using the same word learning paradigm provided support for the neural mechanisms postulated by complementary learning mechanism accounts.

In an fMRI study, Davis et al. (2009) observed a strong hippocampal response to the first presentation of spoken novel words, compared to novel words trained the day before. At the neocortical level, only those novel words that had been trained the day before scanning (i.e., with a chance to consolidate overnight) showed a similar pattern of activation as familiar words, while novel words learned on the same day (i.e., no chance to consolidate overnight) showed an activation pattern similar to untrained novel words. Regarding the neural mechanisms that operate during sleep that might underpin these changes, Tamminen et al. (2010) used the same word learning task and reported that nocturnal slow-wave sleep duration predicted overnight change in accurate recognition memory speed for the newly learned words, and an association between sleep spindle activity and the integration of the novel words, as measured through overnight change in the lexical competition effect. The latter finding was notable as sleep spindles (brief bursts of activity in the 11-15 Hz range occurring in Stage 2 sleep and slow-wave sleep) are known to occur in close temporal connection with hippocampal ripples (Siapas & Wilson, 1998; Staresina et al., 2015) and to prime cortical plasticity (Rosanova & Ulrich, 2005), suggesting they play a key role in hippocampal-neocortical dialogue.

3.1.2. Sleep prior to encoding

While the impact of sleep on memory consolidation has been extensively researched allowing theoretical development, much less is known about the impact of sleep in preparing the brain for learning (e.g., discussed in Section 1.2.2.). The synaptic homeostasis hypothesis (Tononi & Cirelli, 2003, 2006, 2014) proposes that during wake, encoding capacity becomes saturated, due to excessive synaptic potentiation, and encoding capacity is restored following synaptic downscaling during slow-wave sleep. A recent study was published investigating the fMRI and polysomnographic correlates of associative word pair learning in relation to a daytime nap both before and after encoding (Ong et al., 2020). Young adult participants (mean age = 23.3 years) were trained on word-pairs in the fMRI scanner and then completed a recall test prior to a 90-minute nap or period of rested wakefulness (watching a documentary). The participants were then trained and tested again on the word-pairs. After the nap/no-nap manipulation period (N = 20 per between-subjects group), the number of word-pairs recalled was significantly higher for the nap group compared to the no-nap group, despite no difference in baseline (pre-nap) performance. A within-subjects design may have provided greater control over individual differences and higher statistical power, however, the use of baseline measures suggest that changes observed were likely related to the sleep manipulation rather than just individual differences. Nonetheless, sleep spindles (12-15 Hz) were significantly positively correlated with memory performance following the nap. Hippocampal activity, particularly on the left side, was significantly increased during the encoding session following the nap. This increase in hippocampal activity was also positively correlated with spindle count during the daytime nap, suggesting that sleep spindles may be involved in restoring the hippocampus for subsequent learning.

Another study with encoding sessions both before and after a daytime nap, did not find any sleep architectural correlates for behavioural improvements following a daytime nap. In a

study with young adults and older adolescents (19-27-years-old), participants who had a daytime nap had significantly better memory for factual information, in comparison to participants that spent an equivalent period in rested wakefulness (Cousins, Wong, Raghunath, et al., 2019). Interestingly, a second wake group spend the 60-minute period "cramming", that is, continually studying the stimuli. The cramming and nap group performed equally in the immediate test, and both performed significantly better than the rested wakefulness group. However, a week later, only the napping group maintained an advantage over the rested wakefulness group. The authors did not find a significant correlation between behavioural performance and sleep architecture during the nap, including slow-wave sleep and spindle density. The lack of significant sleep architectural correlates may be a function of the study design, as participants studied the stimuli both before and after the daytime nap. As there has been sparse research on the restorative benefits of sleep prior to encoding, the neural mechanisms underlying this process are not well documented. If the restoration and consolidation benefits of sleep use different neural mechanisms, then this would be masked in a study design that considers both consolidation and restoration together. Nonetheless, the homeostatic hypothesis suggests a role for slow-wave sleep in relation to both consolidation and restoration processes. Based on the synaptic homeostasis hypothesis, slow-wave sleep should be significantly associated with the behavioural benefits of a daytime nap regardless of whether restorative, consolidative or combined processes are examined.

Although the synaptic homeostasis hypothesis makes clear predictions about the restorative impact of sleep on learning, few studies have tested these predictions, and none of them have explored the impact of sleep beyond declarative episodic memory, or in the domain of language learning where the impact of sleep on consolidation has been so richly documented. In one of the first studies looking at sleep before learning, Mander et al. (2011) asked participants to encode face-name pairs, followed by a daytime nap or wake, and a

second session encoding of a new set of stimuli. In the no-nap group encoding ability deteriorated from the first to the second encoding session, while in the nap group encoding ability remained stable, suggesting that the nap restored encoding ability. Analysis of subjective and objective measures of fatigue and alertness showed that the groups did not differ on these measures and the observed effect was due to sleep rather than fatigue. Furthermore, both number of fast sleep spindles and time spent in Stage 2 sleep during the nap predicted encoding ability in the second session. The authors (see also Saletin & Walker, 2012) interpreted these findings by extending the principles of complementary learning systems accounts to the memory function of sleep before encoding. They argued that sparse hippocampal representational encoding leads to reducing encoding capacity over the course of the day. Hippocampal-neocortical dialogue, of which sleep spindles are a marker, is needed to shift from hippocampal to cortical-dependence of previously learned stimuli, thus restoring hippocampal episodic learning capacity. This view differs from that of the synaptic homeostasis hypothesis. For example, the Mander et al. framework does not assign a critical role to slow-wave activity. Instead, the data reported by Mander et al. support the view put forward by Saletin and Walker (2012) that there may be a difference between spindle-driven and slow-wave driven memory consolidation mechanisms. The precise nature of the difference is unclear, but the implication is that it ought to be possible to observe a role for sleep spindles in the restoration of encoding ability in the absence of a role for slow-wave activity, or vice versa.

Other studies have however supported the prediction made by the synaptic homeostasis hypothesis that slow-wave activity (which includes slow-wave sleep duration and other slow-oscillation measures) is critical in restoring encoding ability. Antonenko et al. (2012) used transcranial slow oscillation stimulation to induce slow-wave activity during a daytime nap. Following the nap participants encoded pictures, word-pairs, and word lists. All three tasks

benefitted from transcranial slow-oscillation stimulation, compared to a sham control group, suggesting that slow-wave activity before encoding does improve encoding ability. Similarly, Ong et al. (2018) enhanced slow oscillations acoustically during a daytime nap. While acoustic stimulation did not result in improved encoding ability of pictures compared to a sham control, slow oscillation enhancement did correlate with encoding success, providing some support to the prediction that slow-wave activity restores encoding ability. Van der Werf et al. (2009) used mild acoustic sleep-perturbation to reduce slow-wave activity (but not total sleep time) during a night of sleep in a small group of older adults. Participants encoded images the next day, with a lower memory score following perturbed slow-wave activity compared to an unperturbed night of sleep. These stimulation studies provide suggestions that slow-wave activity may be a causal mechanism for enhancing memory encoding ability, as they provide experimental manipulation of theoretically indicated mechanisms (i.e., slow-wave oscillations).

3.1.3. The present study

As discussed above, both the synaptic homeostasis hypothesis and the active systems consolidation model predict that sleep restores memory encoding capacity. However, the current evidence base for the role of sleep prior to encoding is limited (Section 3.1.2.). No studies have tested this prediction in the context of language learning, a domain of declarative learning that is well documented to benefit from sleep-associated memory consolidation after encoding. In the present study we examined the function of sleep for restoring encoding capacity by training participants on a set of new spoken words. Prior to the learning session, participants either took a daytime nap, or remained awake for an equivalent period of time under controlled conditions. Episodic memory for the newly learned words was tested shortly after learning, and again the next day, using a cued recall task, which is known to benefit

from sleep after learning (e.g., Tamminen et al., 2010; Henderson et al., 2012). Note that since the same participants were tested in both the nap and no-nap condition of our study, we can rule out potential confounds due to group differences. Also, because the same participants were tested in both conditions at the same time of the day, any circadian confounds are eliminated.

The impact of pre-encoding sleep on memory integration is particularly interesting for two reasons. Firstly, no study so far has examined the impact of pre-encoding sleep beyond declarative episodic memory. Secondly, lexical integration only emerges after a consolidation opportunity and is not typically observed prior to this (e.g., Gaskell & Dumay, 2003; Dumay & Gaskell, 2007; Davis et al., 2009; Tamminen et al., 2010; Henderson et al., 2012; Tamminen et al., 2017). Lexical integration therefore appears to be a process that occurs only after encoding, and consequently may not be influenced by presence or absence of pre-encoding sleep. However, a recent study suggests that pre-encoding sleep may have knock-on effects on later consolidation processes. Walker et al. (2019) used a visual version of the lexical competition paradigm and manipulated encoding strength by varying the number of training trials during learning of their novel words. They found that when encoding strength was lower than typically used in this paradigm, the overnight consolidation effect was reduced, that is, a weaker overnight change in the magnitude of the lexical competition effect was observed. Intriguingly, overnight change in episodic memory (free recall and recognition memory) was not influenced by the encoding strength manipulation. It appears that disrupting encoding may have knock-on effects on later consolidation that are unique to the integration of new memories with existing memories. As such, we will examine the association pre-encoding sleep and overnight changes (consolidation) in lexical integration.

Additionally, slow-wave sleep may be associated with the restoration of encoding capacity. The synaptic homeostasis hypothesis postulates a key role for slow-wave activity in

the downscaling of synaptic strength. According to this view, slow-wave activity acts both as a marker of increased synaptic strength and contributes to the decrease of synaptic strength. This dual role gives rise to a slow-wave activity control loop during sleep (Tononi & Cirelli, 2014): increasing daytime learning results in increasing synaptic strength, which in turn results in increasing slow-wave activity during sleep. High slow-wave activity during sleep results in increased synaptic depression, which progressively leads to slowing of renormalisation until an equilibrium point is reached where synaptic strength is sufficiently low to further weaken synaptic connections. The clear prediction from this view is that slow-wave activity can be used as a measure of the degree to which the brain restores encoding activity during sleep. This prediction is supported by the brain stimulation studies cited earlier. We therefore predicted an association between slow-wave activity during the nap preceding novel word learning and episodic encoding ability. We employed two measures of slow-wave activity, which have been shown by previous research to be involved with word learning: slow-wave sleep duration (Tamminen et al., 2010) and slow oscillation activity (Tamminen et al., 2013).

While the synaptic homeostasis hypothesis emphasises the importance of slow-wave activity in the downscaling of synaptic strength, it also allows for a role for other neural features of sleep that are associated with the onset of slow oscillations (Tononi & Cirelli, 2014). One of these features is sleep spindles. However, the study reported by Mander et al. (2011) showed an association between Stage 2 rather than Stage 3 frontal fast spindles and restoration of encoding ability. The authors argued that these spindles are a marker of hippocampal activation and support a shift from hippocampal to cortical dependence of newly encoded memories. The suggested hippocampal-to-neocortical dialogue via sleep spindles is considered important in the active systems consolidation models (e.g., Diekelmann & Born, 2010) and could explain the restoration of encoding capacity by

restoring the hippocampus (e.g., Ong et al., 2020, Section 1.2.2.). Although Hypotheses 3 and 4 seek to find support for competing theories of the impact of pre-encoding sleep on memory, in our view these theories are not mutually exclusive. It is possible that both mechanisms outlined in the two hypotheses are in operation simultaneously. While no studies reported so far have observed effects of both sleep spindles and slow-wave activity on restoration of encoding ability at the same time, our word learning paradigm is uniquely suited to uncovering such effects as both slow-wave activity and Stage 2 sleep spindles (Tamminen et al., 2010) have been implicated in consolidation processes involved in this type of learning.

In the current study we sought to establish the impact of sleep on subsequent word learning ability. We used a registered report format (March et al., 2023), as there have been recent replication failures in the field of sleep and memory consolidation, calling into question the robustness of the relationship between slow-wave sleep and memory consolidation (Cordi & Rasch, 2020). Therefore, a greater use of open science practices in the field of sleep and cognition will help to update and establish new theoretical models founded on robust scientific research. In the present study, older adolescent (18-24-year-old) participants learned novel words in two separate sessions immediately after a daytime nap and after an equivalent period of wake. We tested both their episodic memory of the words, and the extent to which the words have been integrated with existing vocabulary. Polysomnography (PSG) was used to test the impact of slow-wave sleep and Stage 2 sleep spindles in restoring word learning ability. As the age of our participants is typical of undergraduate university students, this study may have implications for scheduling academic timetables to facilitate daytime napping.

3.1.4. Hypotheses

1) Hypothesis 1: Napping compared to wake before encoding will result in better episodic memory of newly learned words. We predicted that participants would score significantly higher when learning was preceded by sleep, compared to when it was preceded by wake. We hypothesised to observe this effect both immediately after learning and a day after learning.
2) Hypothesis 2: Napping compared to wake before encoding will result in larger consolidation effects in the integration of newly learned words with existing known words, based on Walker et al.'s (2019) study.
3) Hypothesis 3: Slow-wave activity will be associated with better encoding ability:
    a) We predicted that slow-wave sleep duration would be positively associated with better cued recall.
    b) Power in slow oscillation band would be positively associated with better cued recall.
    c) Slow-wave sleep duration would be positively associated with larger overnight consolidation effects in emerging lexical competition.
    d) Power in slow oscillation frequency band would be positively associated with larger overnight consolidation effects in emerging lexical competition.
4) Hypothesis 4: Frontal fast sleep spindle activity in Stage 2 sleep will be associated with better encoding ability, based on Mander et al. (2011).
    a) We predicted that Stage 2 fast frontal spindle activity should correlate with measures of episodic memory for novel words (cued recall).
    b) Stage 2 frontal spindle activity should also correlate with overnight increases in lexical competition.

3.2. Methods

3.2.1. Power and participants

3.2.1.1. Behavioural data

Power analyses would ideally be based on directly comparable past research. In this case the only directly comparable published study was that of Mander, Santhanam, Saletin and Walker (2011) in that it compared a nap condition with a no-nap condition (although in a less powerful between-subjects design). However, it is not advisable to base power analyses on a single study (Kiyonaga & Scimeca, 2019). For this reason, we searched the literature for all studies that sought to isolate the impact of sleep on word learning in similar paradigms to ours, where both episodic memory and integration of new words with existing words were considered, and which used ANOVAs as the statistical test (Dumay & Gaskell, 2007; Tamminen et al., 2010; Henderson et al., 2012; Wang et al., 2016; Tamminen et al., 2017). This yielded 13 effect sizes ranging from $f = 0.23$ to $f = 0.75$, with a mean of $f = 0.46$, 95% confidence interval ranging from $f = 0.36$ to $f = 0.54$. Given that effect sizes in the published literature tend to be exaggerated (Klein et al., 2018), we chose the smallest reported effect size to base our power analysis on. We used G*Power (Erdfelder, Faul, Buchner, & Lang, 2009; Faul, Erdfelder, Lang & Buchner, 2007) to determine the required sample size. To detect an effect of $f = 0.23$ we would need a sample size of 44 participants to test Hypotheses 1 and 2 (within-subjects ANOVA, 90% power, alpha at 0.02).

3.2.1.2. Polysomnography data

To evaluate correlations between sleep architecture measures and encoding ability, thus testing Hypotheses 3 and 4, we were guided by Mander et al. (2011) as well as a literature search to identify all reports of correlations between sleep architecture and word learning (Tamminen et al., 2010; Tamminen et al., 2017). This yielded four correlations, ranging from

$r = 0.47$ to $r = 0.59$, with a mean of $r = 0.53$, 95% confidence interval ranging from $r = 0.48$ to $r = 0.58$. We again chose the smallest reported effect size to base our power analysis on. To detect a correlation of $r = 0.47$ we would need a sample size of 45 participants to test Hypotheses 3 and 4 (one-tailed, 90% power, alpha at 0.02).

Based on the above calculations, we recruited 45 participants. Participants were students at Royal Holloway, University of London, aged 18-24-years-old and were recruited via posters, e-mails, the university research website, social media, a mailing list from the university sleep lab and word of mouth. Participation was voluntary and participants were paid for their time. To be eligible to participate, participants had to be native English speakers with normal hearing and normal or corrected-to-normal vision, have no currently diagnosed neurological, psychiatric, or sleep disorders, must not be currently taking medication effecting their sleep, must not be currently engaged in shift work involving working at night, have not travelled across time zones within two weeks of taking part in any of the sessions in the study, and must have no known special educational needs (e.g., dyslexia, ADHD, autism spectrum disorder).

Participants were excluded if they failed to complete all sessions of the study, or if data were lost for the reasons described in the analysis plan. Based on our experience of running nap studies, we expected about 10-15% of recruited participants would be excluded. Excluded participants were replaced until the required sample size was achieved, at which point recruitment ceased. This study was given ethical approval following the Royal Holloway University of London Ethics Committee procedures. Using a within-subjects design, participants took part in both the wake and nap conditions. The order of the conditions was randomly assigned to each participant and counterbalanced across participants.

3.2.2. Materials

Materials used in this study are available on the Open Science Framework at https://osf.io/98es5/. Legal copyright restrictions prevent public archiving of WASI-II, which can be obtained from the copyright holders in the cited references.

3.2.2.1. Screening questionnaire

Participants were required not to consume alcohol or caffeine in the 24 hours preceding the encoding session, and not to nap on the day of each encoding session. To facilitate falling asleep in the nap session and to control for time spent awake before the encoding sessions, participants were asked to wake no later than 6 a.m. the morning of both encoding sessions and not to nap before the experiment. Adherence was monitored through the screening questionnaire and additionally through an actigraph attached to the non-dominant wrist the evening before an encoding session. Each participant was reminded of these requirements through a phone call or a text message the day before each encoding session. Participants who failed to adhere were dismissed and replaced.

3.2.2.2. Polysomnography

Polysomnography measures were recorded during the nap using an Embla N7000 system. Electrodes were placed on the scalp on the left and right frontal (F3, F4), central (C3, C4) and occipital (O1, O2) positions as recommended by the American Academy of Sleep Medicine (AASM) guidelines. These were referenced to the average of the left and right mastoid. The ground electrode was placed on the forehead. Electro-oculogram (EOG) and electromyography (EMG) was used to measure eye movements and muscle tone respectively, with positioning of the electrodes following AASM guidelines. Impedances were kept below

5Ω for EEG electrodes and below 10Ω for EOG and EMG electrodes. Signals were sampled at 500 Hz.

3.2.2.3. SASS-Y (Dietch, Sethi, Slavish, & Taylor, 2019)

A retrospective sleep diary was used to measure participants' habitual sleep patterns in the week prior to data collection. The Self-Assessment of Sleep Survey - Split (SASS-Y) can also be used to determine the difference between sleep patterns on the weekdays versus the weekends. These data were only used for exploratory analyses.

3.2.2.4. Habitual napping

Participants were asked how often they nap during a typical week, if they reported napping once or more per week then participants were classified as habitual nappers (e.g., Cousin, Wong et al., 2019). Participants who reported napping less than once per week were considered as non-habitual nappers. These data were only used for exploratory analyses.

3.2.2.5. Stanford Sleepiness Scale (SSS; Hoddes et al., 1972)

The SSS was used to evaluate levels of sleepiness. Before encoding and again before each test participants were asked to report their current level of sleepiness on a scale from 1 ("feeling active, vital, alert, or wide awake") to 7 ("No longer fighting sleep, sleep onset soon, having dream-like thoughts"). These data were used to check whether participants were equally tired when conducting the test sessions in the nap and no-nap conditions.

3.2.2.6. Wechsler Abbreviated Scale of Intelligence II (WASI-II: Wechsler, 2011)

The WASI-II vocabulary subscale was used to measure participants' existing vocabulary knowledge. Participants were shown a list of words from the WASI-II stimulus

book under the vocabulary subsection (page 40-51 of the WASI-II stimulus book). The researcher read each word aloud to the participant and pointed to the respective word. Participants were asked to define each word. All participants started at item four as they were over six years of age and continued until they scored zero for three consecutive items allowing them a maximum raw score of 59. Each response was scored from zero to two points, with DK to represent when the participant said they do not know the definition or NR if the participant did not respond in approximately 30 seconds. The scoring was in accordance with the WASI-II administration manual with two points awarded for clear understanding of the word, one point for general understanding but in a vaguer manner, and zero points for incorrect or no obvious understanding of the word. This scoring allows a measurement of depth of vocabulary knowledge by assessing what the participant knows about the word. Additionally, this vocabulary measurement assesses breadth of vocabulary knowledge via the number of words a participant knows. Raw scores were converted to norm-referenced T-Scores relevant to the participants' age as indicated by the manual. These data were used to characterise the sample based on their vocabulary level.

3.2.2.7. Novel word learning paradigm

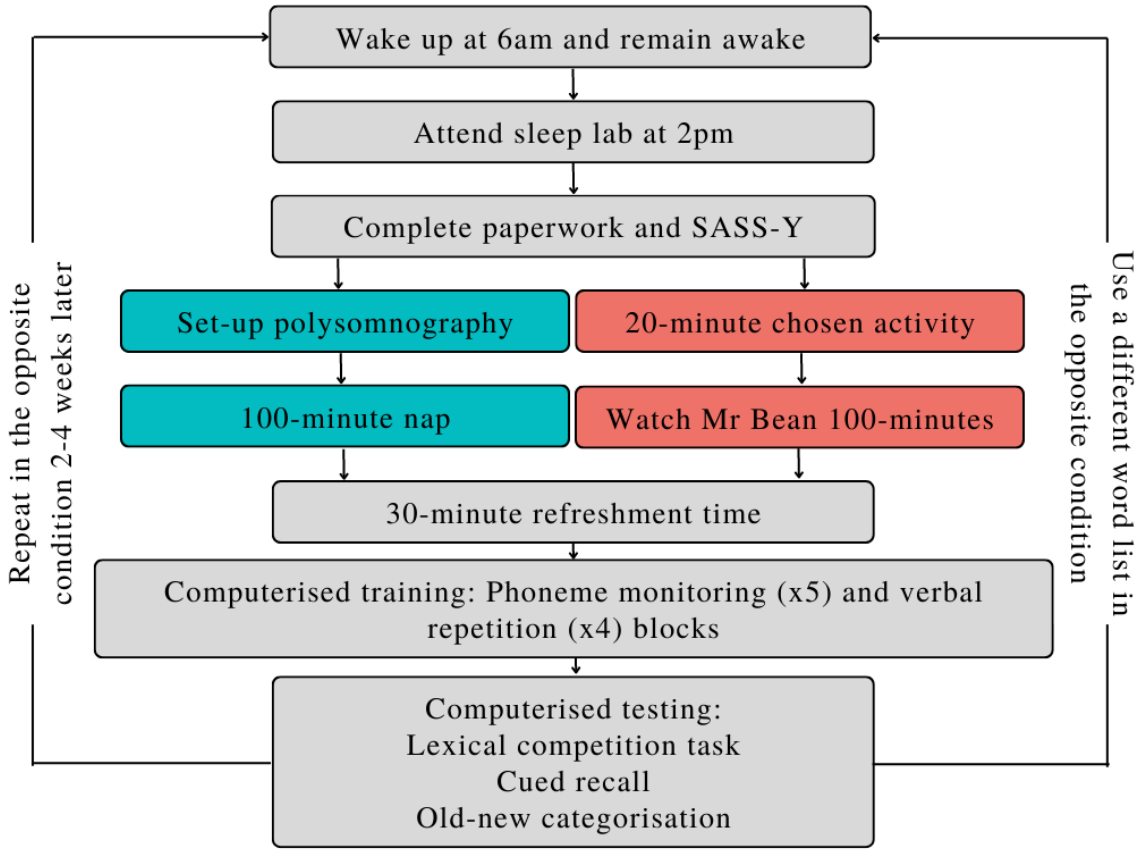
The "cathedruke" paradigm used in Chapter 2 (see Section 2.2.2.4. for method) was used again in this study. There are some minor differences in the task used in this chapter. In the Stage 1 registered report, we stated that the trials in the phoneme monitoring training task would end after 2000 milliseconds had elapsed, but this was extended to 3000 milliseconds so not to disadvantage slower responders. In the verbal repetition training task, we had planned the trial would end after 2000 milliseconds from the onset of the word, but this was changed so that participants made a button press to indicate when they were ready to move on to the next trial. The computer-based cued recall testing task was used, and trials ended

after 10 seconds. Participant's verbal responses were recorded via the headphones for offline scoring, and the order of trials were randomised by the E-Prime software for each participant. In Chapter 2, the new-old categorisation (recognition) task gave participants 3000 milliseconds to respond, whereas, in Chapter 3 participants were given 5000 milliseconds to respond so not to disadvantage slower responders. Participants in this chapter completed the test tasks in the following fixed order: lexical competition, cued recall, and old-new categorisation. Since there is inconsistent evidence regarding the impact of sleep on recognition memory in word learning, we do not have a strong basis for making predictions about the old-new categorisation task and therefore this task was only considered in exploratory analyses. The words used in the lists are reported in Appendix 2.1. Two of the lists were used for encoding and testing, one in the nap condition and the other in the no-nap condition. The third list remained an untrained control. The assignment of lists to these conditions was counterbalanced across participants.

3.2.3. Procedure

Participants underwent two two-day experimental sessions, one for the nap condition and one for the no-nap condition, separated by a minimum of two weeks and a maximum of four weeks to avoid carryover effects. Both sessions involved a 100-minute sleep or wake opportunity followed by the encoding phase and a same-day test phase. A delayed test phase took place the next day. The procedure is visualised below in Figure 3.1.

Figure 3.1. Flowchart of the study procedure. blue boxes represent the nap condition, red boxes represent the wake condition, and grey boxes are across both conditions



Both the nap and no-nap sessions began at 14:00. Participants were first asked to complete the consent forms, the initial screening questionnaire, and the SASS-Y. In the nap session participants then underwent preparation for the PSG recording. In the no-nap session participants were allowed to engage in activities of their choice within the sleep laboratory for an equivalent time (20 minutes). Following Mander et al. (2011), in the nap condition participants were then given a sleep opportunity of 100 minutes. In the no-nap condition participants watched a video with no language input (Mr Bean) for 100 minutes. At the end of the 100 minutes, the nap participants were given 30 minutes to clean up and to overcome potential sleep inertia effects. In the no-nap condition participants were allowed to engage in

activities of their choice within the sleep laboratory for an equivalent time. After this, participants filled in the SSS, and the novel word training session began.

The WASI was administered once, at the very end the study (after the final test session). In all tasks, stimuli were delivered by E-Prime. Manual responses were collected by a millisecond accurate button box. Auditory stimuli were delivered, and vocal responses collected by Beyerdynamic DT 234 Pro headsets.

3.2.4. Exclusion criteria, quality checks, and blinding procedures

Only participants who met the eligibility criteria (set out in Section 3.2.1) were recruited to take part. Once participants were recruited, they were excluded (and replaced) from the analysis if they met any of the following criteria:

1) Their average accuracy rate in the phoneme monitoring task in the training phase was at or below chance (50%). This suggests they were not paying attention to the novel word training and therefore the quality of the data in both training and subsequent tests would be compromised.
2) Their average accuracy rate in the lexical competition task was at or below chance (50%) in one or more of the test sessions. This suggests they were not attending to the task and therefore the quality of the data in this task would be compromised.
3) They failed to enter slow-wave sleep or stage 2 sleep during the nap. Theories discussed above and the existing literature (e.g., Antonenko et al., 2013) predict that slow-wave sleep and/or stage 2 sleep will be of critical importance in enhancing encoding ability. Therefore, data from participants who fail to get any slow-wave and/or stage 2 sleep will not be able to address the hypotheses of this study.
4) They failed to follow any of the instructions outlined in Section 3.2.2.1. as indicated by the screening questionnaire and/or actigraphy.

5) Their data from any of the behavioural tasks or PSG data were lost due to experimenter error or device malfunction.
6) They chose to withdraw from the study before completing all sessions and tasks.

Blinding procedures were implemented to avoid experimenter bias from impacting the results. We identified two stages of the data analysis where bias could occur. Firstly, to avoid any bias in the scoring of the PSG data, the scorer had not seen any of the participant's behavioural data at the time of scoring. Secondly, to avoid bias in scoring the accuracy of the cued recall data, the scorer remained blind to the condition (nap or no-nap) of the dataset being scored. There was no room for subjective judgement in the analysis of other aspects of the data, as the analysis strictly followed the pipeline outlined below (Section 3.3.). Potential experimenter bias during data collection was avoided by giving participants full written instructions for each behavioural task, in addition to brief computer-administered instructions at the beginning of each task.

3.3. Analysis plan

Alpha levels in all analyses were set at 0.02, unless otherwise stated. This differs from the rest of the thesis, which uses the standard alpha level of 0.05, due to the guidelines required for publication within the journal 'Cortex', in which the present study was published (March et al., 2023).

3.3.1. Hypothesis 1: Napping compared to wake before encoding will result in better episodic memory of newly learned words

3.3.1.1. Cued recall

The cued recall responses recorded by E-Prime were phonetically transcribed and scored as correct or incorrect. Only responses that matched the cued novel words in all

phonemes were scored as correct. A 2 x 2 within-participants ANOVA with nap condition (nap vs. no-nap) and test time (immediate vs. delayed) as factors was calculated on the proportion of novel words correctly recalled. A significant main effect of nap condition would support Hypothesis 1. No significant interaction was predicted, but if a significant interaction had been found, planned contrasts (one-tailed t-tests) would have been used to compare the nap and no-nap conditions at both tests times to clarify the source of the interaction. We predicted that in at least one of those comparisons the nap condition would show higher cued recall. In this case a Bonferroni correction would be applied to the alpha level, and therefore only t-tests with $p \leq .01$ would be considered statistically significant.

3.3.2. Hypothesis 2: Napping compared to wake before encoding will result in larger consolidation effects in the integration of newly learned words with existing known words

3.3.2.1. *Lexical competition*

RTs only to the base words and control words were analysed, data for filler words and non-words are not informative regarding our hypotheses. Following precedent using this same task (Tamminen et al., 2010; Tamminen & Gaskell, 2008), RTs were log-transformed to better meet the assumption of normality, and all RTs faster than 300 milliseconds and slower than 2500 milliseconds were removed to reduce the impact of outliers. Only trials where an accurate response was made were entered into the analysis. The magnitude of the lexical competition effect was calculated for each participant at both test times by deducting the mean log RT to control words from the mean log RT to base words with novel competitors. A 2 x 2 within-participants ANOVA with nap condition (nap vs. no-nap), and test time (immediate vs. delayed) was calculated on the RTs. Given that lexical competition effects are not observed immediately after training, Hypothesis 2 predicted that participants

in the nap condition would show a larger overnight increase in the lexical competition effect, manifested in a significant interaction. A planned contrast (one-tailed t-test) would have been used to confirm that the significant interaction reflected this pattern by comparing the overnight change in the magnitude of the lexical competition effect across the two conditions.

3.3.3. Hypothesis 3: Slow-wave activity will be associated with better encoding ability

Within the nap condition, sleep architecture data were assessed to investigate the role of slow-wave activity in restoring encoding capacity in accordance with the synaptic homeostasis hypothesis and existing literature (Antonenko et al., 2013; Ong et al., 2018). Two measures of slow-wave activity that have in the past been successfully implicated in novel word learning were investigated here: slow-wave sleep duration (Tamminen et al., 2010) and spectral power in the slow oscillation frequency band (Tamminen et al., 2013). We closely followed the analysis strategy and methods of these existing studies.

3.3.3.1.*Slow-wave sleep duration*

The PSG record of the nap was scored by an experienced sleep scorer following AASM guidelines (see Section 3.2.4. for blinding measures) in 30-second epochs. Total time spent in slow-wave sleep was recorded for each participant. A one-tailed Pearson correlation coefficient was calculated between slow-wave sleep duration and proportion of correct responses in the cued recall task (Hypothesis 3a). Only data from the immediate test were used here, as there was no a priori justification for expecting that a correlation would be seen only in the immediate test but not in the delayed test, or vice versa. A one-tailed Pearson correlation coefficient between slow-wave sleep duration and the overnight change in the lexical competition effect was also calculated (Hypothesis 3c). The lexical competition effect here and in subsequent analyses was calculated for each participant by deducting the mean

RT to control base words from the mean RT to base words. As two correlations involving slow-wave sleep duration were calculated here, a Bonferroni correction was applied to the alpha level, and therefore only correlations with $p \leq .01$ were considered statistically significant.

3.3.3.2.*Slow oscillations*

EEG power spectral density in slow-wave sleep was analysed in the slow oscillation frequency band (0.5 – 1 Hz). Pre-processing steps included extraction of those epochs scored as slow-wave sleep, exclusion of noisy electrodes (e.g., due to an electrode becoming dislodged during the nap), and manual removal of artefacts due to arousals or movement. The pre-processed data on each EEG electrode were then subjected to power spectral analysis following Welch's method using 4-second Hamming window length with 50% overlap. Statistical analysis of the spectral data was carried out on log-transformed absolute power ($\mu V^2/Hz$). As there was no a priori theoretical or empirical justification for restricting the analysis to specific electrode sites, we calculated the average power across all electrodes. This is further justified by data reported in Tamminen et al. (2013) who found slow-oscillation power averaged across cortical electrodes to be involved in integration of newly learned words in the mental lexicon. One-tailed Pearson correlation coefficients were calculated between slow-oscillation power and proportion of correct responses in the cued recall task (Hypothesis 3b). Only immediate test data were used again. A one-tailed Pearson correlation coefficient between slow-oscillation power and overnight change in the lexical competition effect was also calculated (Hypothesis 3d). As two correlations involving slow-oscillation power were calculated here, a Bonferroni correction was applied to the alpha level, and therefore only correlations with $p \leq .01$ were considered statistically significant.

3.3.4. Hypothesis 4: Frontal fast sleep spindle activity in Stage 2 sleep will be associated with better encoding ability

Pre-processing steps included extraction of those epochs scored as Stage 2 sleep, exclusion of noisy electrodes (e.g., due to an electrode becoming dislodged during the nap), and manual removal of artefacts due to arousals or movement. The raw EEG data were then band-pass filtered to restrict analysis to fast spindles (13.5 – 15 Hz) using a linear finite impulse response (FIR) filter. An automated detection algorithm developed by Ferrarelli et al. (2007) was used to derive the number of discrete spindle events for each electrode. The algorithm is freely available as an appendix to Warby et al. (2014). Briefly, the algorithm counts amplitude fluctuations in the filtered time series exceeding a predetermined threshold as spindles. These thresholds are calculated relative to the mean channel amplitude and set to eight times the average amplitude. This algorithm has been successfully used in previous research into sleep and word learning (e.g., Tamminen et al., 2013; Tham et al., 2015), including the same spoken word learning paradigm as in the current study (Tamminen et al., 2010). The average number of spindles detected over frontal electrodes (F3 and F4) was calculated and used in the correlational analyses. One-tailed Pearson correlation coefficients were calculated between number of spindles and proportion of correct responses in the cued recall task (Hypothesis 4a). A one-tailed Pearson correlation coefficient between the number of spindles and overnight change in the lexical competition effect was also calculated (Hypothesis 4b). As two correlations involving spindle count were calculated here, a Bonferroni correction was applied to the alpha level, and therefore only correlations with $p \leq .01$ were considered statistically significant.

3.3.5. Current sleepiness

As the main manipulation in this experiment concerns sleep, it is possible that participants may have been more or less sleepy at test in one or the other condition and this may have affected their test performance. Out of our two tests tasks cued recall could have been vulnerable to this potential confound (sleepiness may have affected global RTs in the lexical decision task but not the RT *difference* between the control words and base words with new competitors). We compared the Stanford Sleepiness Scale (SSS) scores at both test sessions to check whether participants were equally tired when completing the nap and no-nap conditions. A two-tailed paired t-test comparing data from the two conditions was calculated for the immediate test session, and another one for the delayed test session. A Bonferroni correction was applied to the alpha level, and therefore only t-tests with $p \leq .01$ were considered statistically significant. If a significant difference in SSS scores had been found at either test time, we would have repeated the ANOVA analysis under Hypothesis 1 using an Analysis of Covariance (ANCOVA). For each participant we would have calculated the sleepiness difference between the nap and no-nap conditions (based on the immediate test if that is where the difference was observed, based on the delayed test if that is where the difference was observed, or based on the mean values across both tests if the difference was observed in both). This difference would have been centred (Schneider, Avivi-Reich & Mozuraitis, 2015) and entered as an additional covariate with the same main effects and interactions as outlined under Hypothesis 1. If sleep restores episodic memory encoding ability, we predicted we would see a significant main effect of nap condition even when the sleepiness difference between nap and no-nap conditions is accounted for by the covariate.

3.4. Results

The data and analysis scripts are available at https://osf.io/98es5/ and the Stage 1 manuscript at https://osf.io/73yms/. The Stage 2 manuscript is now published (March et al., 2023). There were eight male (18%) and 37 female (82%) participants. Participants were aged between 18-24-years-old (*M* = 21.11, *SD* = 1.22 years). The minimum age-normed t-score on the WASI-II vocabulary assessment was 44 and the maximum was 77 *(M* = 54.11, *SD* = 6.06), indicating vocabulary abilities in the average range. Thirteen participants were non-habitual nappers (29%) and 32 were habitual nappers (71%). Seven further participants took part but were replaced following the criteria outlined in Section 3.2.4. Three participants withdrew before completing the study. Three participants were excluded due to equipment malfunction. One participant was excluded due to researcher error, training them on the same word list in both sessions.

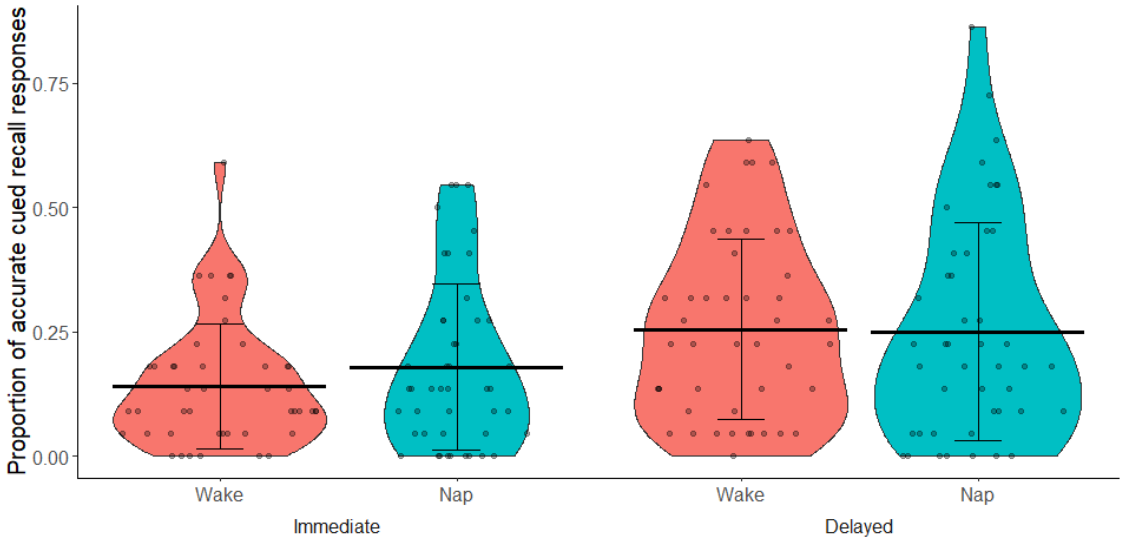
All data were collected in accordance with the registered report protocol, with two exceptions, both approved by the editor. First, one participant had their second session 34 days after the first, compared to the intended 28 days. As this minor deviation from protocol is unlikely to impact the results, we included the participant's data in the analyses. Second, two participants' data had to be excluded from the sleep spindle and spectral power analyses due to intermittent loss of signal. We were unable to replace these two participants due to a narrow data collection window caused by COVID-19 restrictions to face-to-face data collection.

3.4.1. Hypothesis 1: napping compared to wake before encoding will result in better episodic memory of newly learned words

3.4.1.1. Cued recall

Figure 3.2 shows the proportion of correctly recalled novel words in each condition and at each test time. There was no significant main effect of nap condition (nap vs. no nap) in cued recall accuracy, $F(1, 44) = 0.47$, $p = .495$, $\eta_p^2 = .011$. There was also no significant interaction between nap condition and test time (immediate vs. delayed), $F(1, 44) = 2.89$, $p = .096$, $\eta_p^2 = .062$. There was a significant main effect of test time, $F(1, 44) = 53.61$, $p < .001$, $\eta_p^2 = .549$. Proportion accuracy in cued recall was significantly higher in the delayed test ($M = 0.25$, $SD = 0.20$) compared to the immediate test ($M = 0.16$, $SD = 0.15$).

Figure 3.2. Hypothesis 1: Proportion of accurate cued recall responses across nap condition and test time



Note: The horizontal bar represents the mean, the borders around the data points are smoothed density curves, and the error bars represent the standard deviation

3.4.2. Hypothesis 2: Napping compared to wake before encoding will result in larger consolidation effects in the integration of newly learned words with existing known words

3.4.2.1. Lexical competition

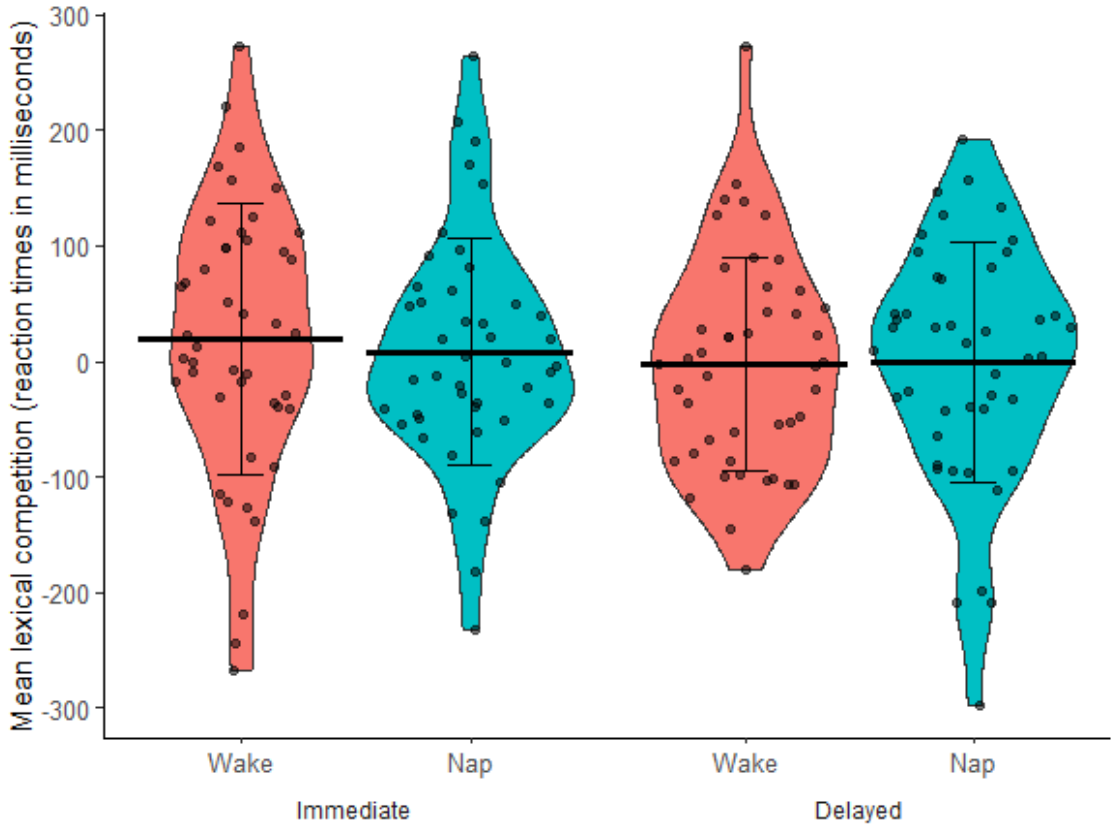
Trimming of outliers (below 300 milliseconds or above 2500 milliseconds) resulted in the removal of 1.82% (N = 144) of the trials. Removal of incorrect responses resulted in the removal of a further 7.06% (N = 549) of trials. Table 3.1 shows the RTs to base words and control words in each condition, and the error rate per condition. Figure 3 shows the magnitude of the lexical competition effect in each condition. There was no significant main effect of nap condition (nap vs. no nap) on magnitude of lexical competition, $F(1, 44)$ = 0.053, p = .819, η_p^2 = .001, and no main effect of test time (immediate vs. delayed), $F(1, 44)$ = 0.393, p = .534, η_p^2 = .009. Contrary to our hypotheses, there was also no significant interaction between condition and test time, $F(1, 44)$ = 0.502, p = .482, η_p^2 = .011.

Table 3.1. Hypothesis 2: Mean response times (in milliseconds) and error rates to base and control words in the lexical decision task

	Nap				Wake			
	Immediate		Delayed		Immediate		Delayed	
	Base	Control	Base	Control	Base	Control	Base	Control
Reaction times	1206 (180)	1199 (178)	1122 (173)	1124 (187)	1225 (171)	1206 (173)	1125 (175)	1127 (173)
Lexical competition	7.80 (98.1)		-1.21 (104)		19.3 (117)		-2.24 (92.8)	
Error (%)	9.54	7.76	7.00	4.71	8.76	6.62	7.12	5.50

Note: Lexical competition is the reaction time difference (in milliseconds) between base words and control words. Means are presented with standard deviations (SD) in parentheses. SD is calculated after aggregating trials by participant. Error is calculated after the removal of reaction time outliers.

Figure 3.3. Hypothesis 2: Magnitude of the lexical competition effect across nap condition and test time



Note: The horizontal bar represents the mean, the borders around the data points are smoothed density curves, and the error bars represent the standard deviation.

3.4.3. Hypothesis 3: Slow-wave activity will be associated with word learning capacity

3.4.3.1. Hypothesis 3a: Slow-wave sleep duration and cued recall

Table 3.2 shows the average distribution of time spent in different sleep stages during the 100-minute nap opportunity. Figure 3.4. Panel A shows the association between proportion of words correctly recalled in the cued recall task and slow-wave sleep duration in the immediate test in the nap condition. There was no significant correlation between slow-wave sleep duration (in minutes) and proportion of accurate cued recall responses, $r(43) = 0.17$, $p = .139$ [*95% CI* = -0.09 – 1.00]

Table 3.2. Average time spent in different sleep stages (standard deviation in parentheses)

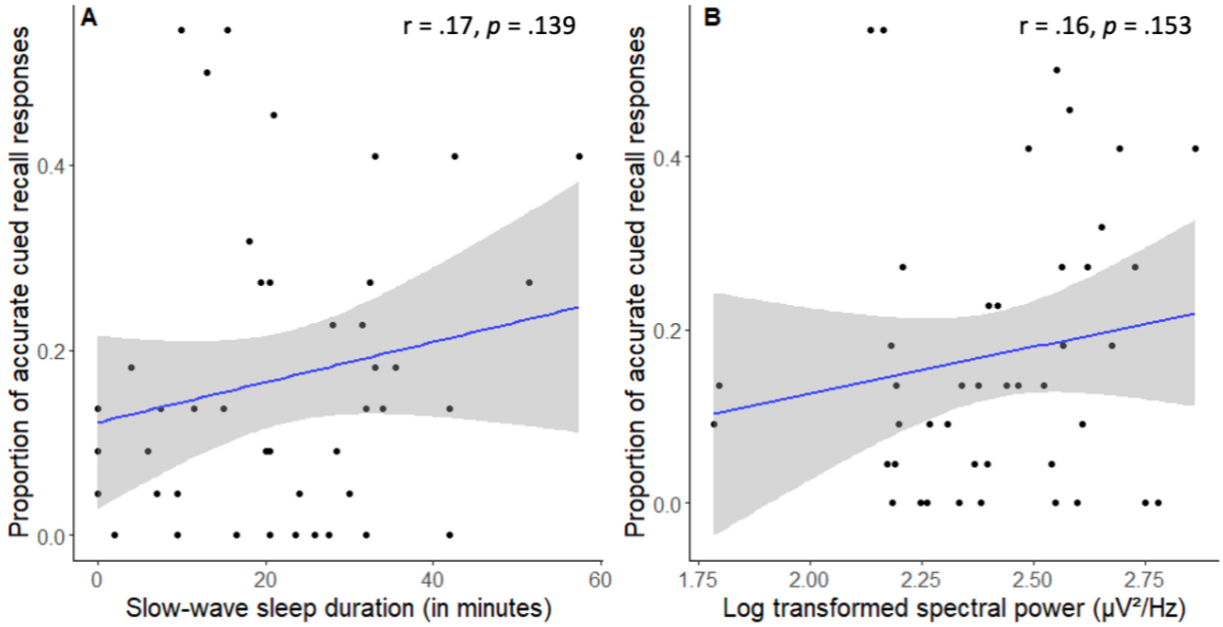
Sleep stage	Time in minutes
Stage 1	4.56 (2.52)
Stage 2	40.65 (13.51)
SWS	22.10 (13.64)
REM	5.50 (8.75)
Total sleep time	73.05 (19.43)

Note: SWS = slow-wave sleep, REM = rapid eye movement sleep

5.4.3.2. Hypothesis 3b: Slow oscillation power and cued recall

Figure 3.4. Panel B shows the association between proportion of correct answers in the cued recall task and spectral power in the slow oscillation frequency band. There was no significant correlation between these variables, $r(41) = .16$, $p = .153$ [*95% CI* = -0.10 – 1.00]

Figure 3.4. Scatterplots showing the relationships between the proportion of accurate cued recall responses in the immediate test and slow-wave sleep duration (in minutes; Panel A; Hypothesis 3A) and log transformed spectral power (Panel B; Hypothesis 3B)



Note: Grey shading around the line of best fit represents 95% confidence intervals

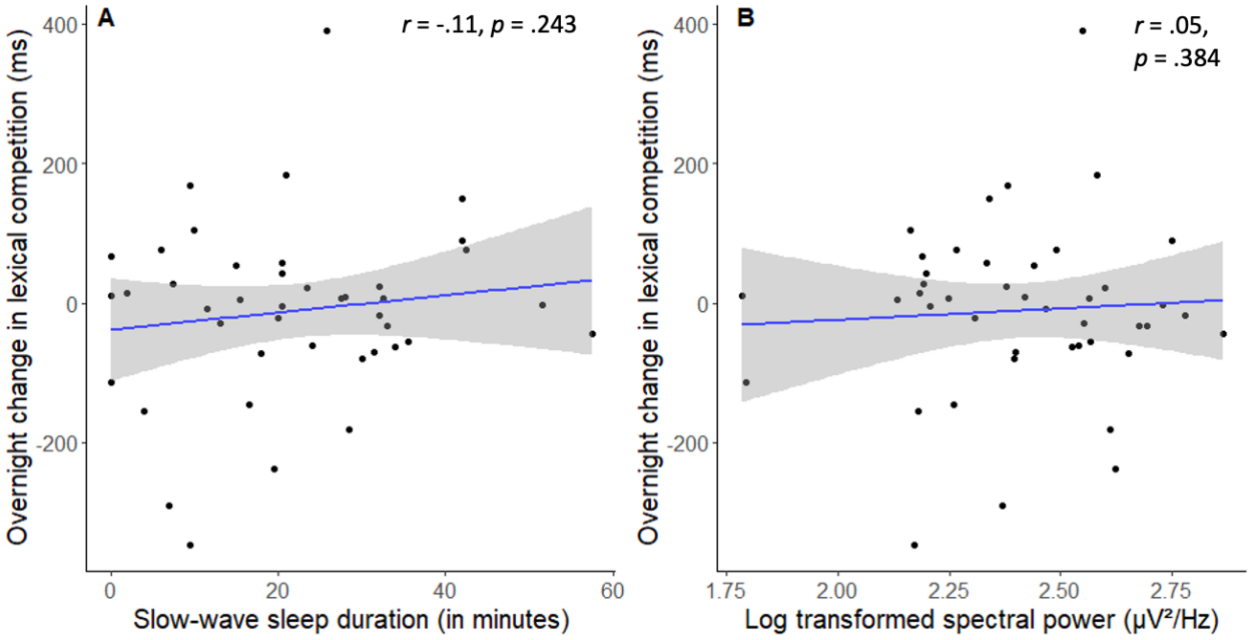
3.4.3.3. Hypothesis 3c: Slow-wave sleep duration and lexical competition

Figure 3.5. Panel A shows the association between the overnight change in the magnitude of the lexical competition effect and time spent in slow-wave sleep. There was no significant correlation between these variables, contrary to our prediction, $r(43) = -0.11$, $p = .243$, [*95% CI* = -1.00 – 0.15].

3.4.3.4. Hypothesis 3d: Slow oscillation power and lexical competition

Figure 3.5. Panel B shows the association between overnight change in the lexical competition effect and power in the slow oscillation frequency band. There was no significant correlation between these variables, $r(41) = 0.05$, $p = .384$ [*95% CI* = -0.21 – 1.00].

Figure 3.5. Scatterplot showing the relationship between overnight change in lexical competition (in milliseconds) from the immediate to the delayed test of the nap condition and slow-wave sleep duration (in minutes; Panel A Hypothesis 3C) and log transformed spectral power (Panel B, Hypothesis 3D)



Note: Grey shading around the line of best fit represents 95% confidence intervals.

3.4.4. Hypothesis 4: Frontal fast spindle activity in Stage 2 sleep will be associated with word learning capacity

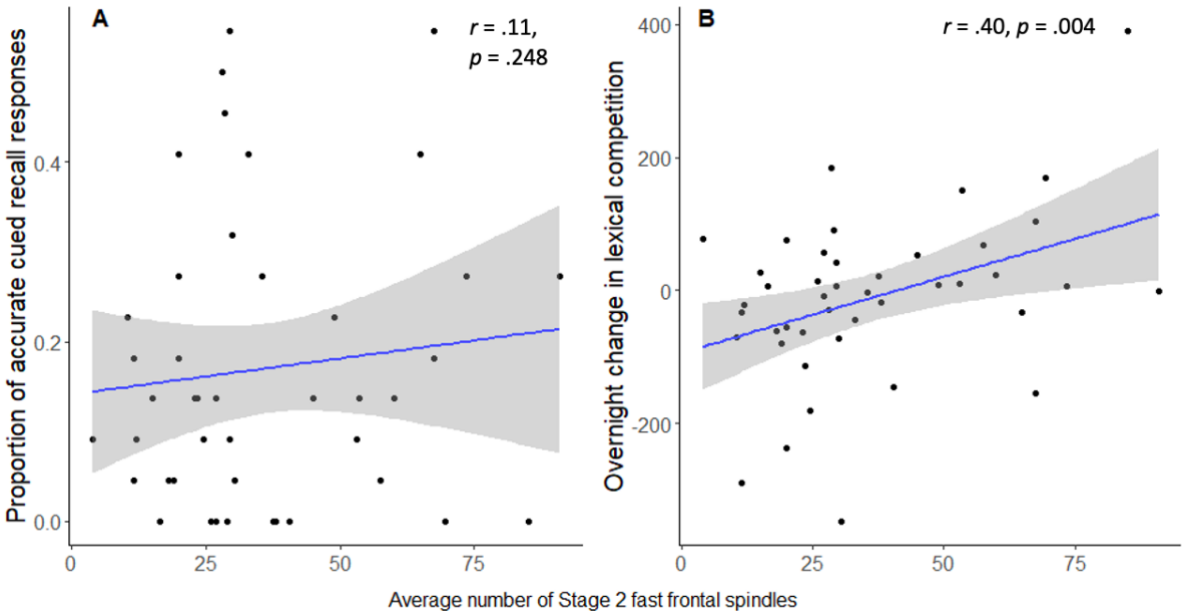
3.4.4.1. Hypothesis 4a: Sleep spindles and cued recall

Figure 3.6. Panel A shows the association between proportion of words correctly recalled in the cued recall task immediately after training and the number of frontal fast sleep spindles. There was no significant correlation between the variables, $r(41) = 0.11$, $p = .248$ [*95% CI* = -0.15 – 1.00].

3.4.4.2. Hypothesis 4b: Sleep spindles and lexical competition

Figure 3.6. Panel B shows the association between overnight change in the magnitude of the lexical competition effect and number of frontal fast sleep spindles. There was a significant positive correlation between these variables, $r(41) = 0.40$, $p = .004$ [*95% CI* = 0.17 – 1.00]. This suggests that more fast frontal spindles prior to encoding is associated with greater increases in lexical competition from the immediate test to the delayed test (after a period of overnight sleep).

Figure 3.6. Scatterplot showing the relationship between the average number of fast frontal spindles and the proportion of accurate cued recall responses in the immediate test (Panel A; Hypothesis 4A) and overnight changes in lexical competition from the immediate to the delayed test (Panel B; Hypothesis 4B) in the nap condition



Note: Grey shading around the line of best fit represents 95% confidence intervals.

3.4.5. Current sleepiness

Paired t-tests were used to examine whether there were significant differences in the Stanford Sleepiness Scale (SSS) ratings in the nap condition test sessions compared to the wake condition test sessions. There was not significant difference in the immediate test, $t(44) = -0.37$, $p = .71$, $d = -0.08$ or in the delayed test, $t(44) = -0.75$, $p = .46$, $d = -0.11$. Therefore, we do not need to use an ANCOVA in the main analyses, as stated in section 3.3.5.

3.5. Exploratory analyses

3.5.1. Cued recall using a graded measure of accuracy

In our pre-registered analyses of the cued recall data, we categorised a response as correct only if it perfectly matched the trained novel word. As shown in Figure 3.2, only a small proportion of responses met this criterion. Importantly, under this scoring both a failure to produce any response and a response that deviates from the correct word by just one phoneme are scored as incorrect, yet the participant in the latter case clearly recalled more than in the former case. In response to this problem, we (Ricketts, Dawson & Davies, 2021) and others have started calculating Levenshtein distance as a more fine-grained measure of recall accuracy, both in the domain of word learning (e.g., Ricketts et al., 2021; Frances, de Bruin & Dunabetia, 2020) and in the literature looking at the impact of sleep-associated memory consolidation (e.g., Kurdziel & Spencer, 2016). Levenshtein distance is a measure of how many changes (including insertions, deletions, or substitutions of phonemes or letters) are needed for the participant's response to match the correct response; a larger distance suggests the participant's response was further away from the target. It can be used to calculate a matching percentage that quantifies the overlap between the given response and the correct response. In calculating the matching percentage, we followed the formula used by Kurdziel and Spencer (2016): matching percentage = 100 – [(Levenshtein

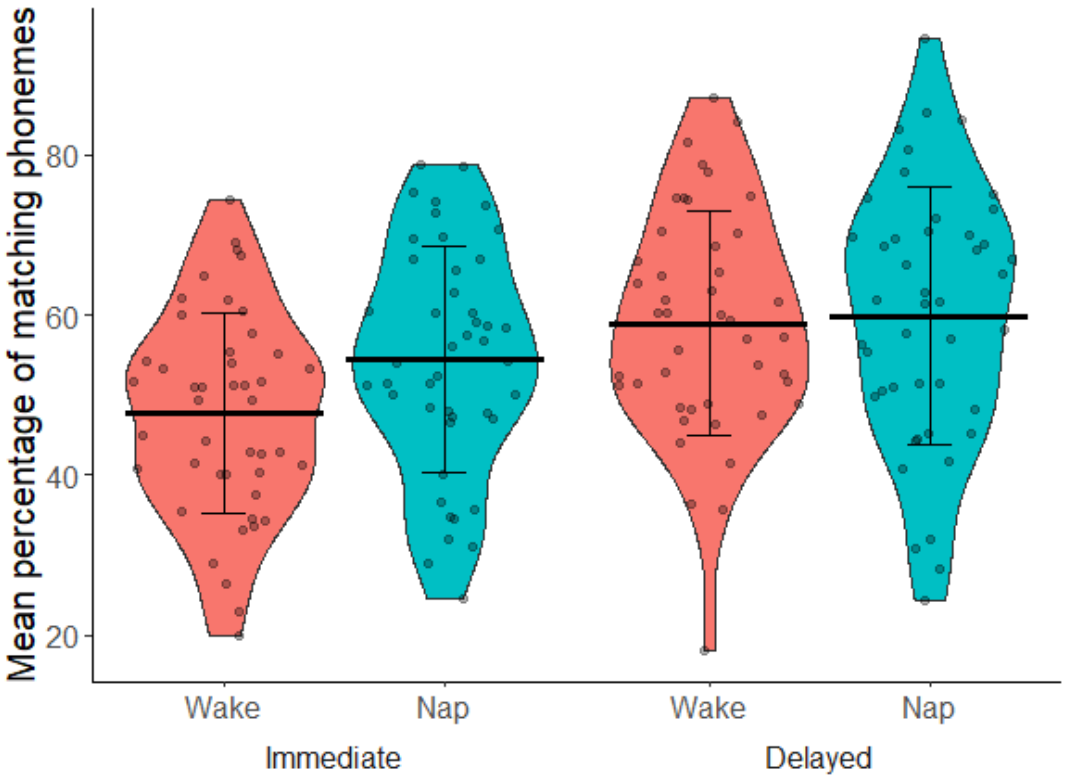
Distance)*100]/Max Length, where Max Length equals the number of phonemes in the longer of the two words (actual response or correct response) and controls for word length. A matching percentage of 100% represents a completely correct response. A matching percentage of 0% indicates a response with no overlap with the correct response, or a failure to enter a response. Importantly, Levenshtein distance captures a graded degree of accuracy that is more sensitive to individual differences than binary (correct or incorrect) accuracy (Ricketts et al., 2021); the formula from Kurdziel and Spencer (2016) also considers the impact of blank responses in the calculation of mean accuracy.

In our first exploratory analysis we repeated the main analyses on the cued recall task using matching percentage as the measure of accuracy (see Figure 3.7.). If multiple responses were given for the same item, then the item with the shortest Levenshtein distance was selected as the final answer. A mean matching percentage was calculated per participant per session. There was no significant main effect of nap condition (nap vs. no nap) on matching percentage, $F(1, 44) = 3.59$, $p = .065$, $\eta_p^2 = .075$. However, there was a significant main effect of test time (immediate vs. delayed), $F(1, 44) = 73.03$, $p < .001$, $\eta_p^2 = .624$. Critically, there was also a significant interaction between condition and test time, $F(1, 44) = 14.79$, $p < .001$, $\eta_p^2 = .252$.

We followed up the significant interaction with four two-tailed t-tests. Alpha level was Bonferroni corrected such that results were considered statistically significant if $p \leq .005$, two-tailed. There was a significant difference between nap and wake conditions in the immediate test, $t(44) = 3.162$, $p = .003$, $d = -0.50$, but not in the delayed test, $t(44) = 0.441$, $p = .66$, $d = -0.06$. The impact of time of test on the other hand was not modulated by condition, the difference between immediate and delayed test performance was significant both in the nap condition, $t(44) = 4.696$, $p < .001$, $d = -0.35$, and in the wake condition, $t(44) = 8.626$, $p < .001$, $d = -0.83$.

We also repeated the three pre-registered correlational tests on the immediate test of the nap condition for the cued recall task using the matching percentage measure. Alpha level was Bonferroni corrected such that results were considered statistically significant if $p \leq .007$, two-tailed. Matching percentage was not significantly correlated with slow-wave sleep duration, $r(43) = 0.22$, $p = .153$ [*95% CI* = -0.08 – 0.48], spectral power in the slow oscillation frequency band $r(41) = 0.11$, $p = .474$ [*95% CI* = -0.19 – 0.40], or number of fast frontal spindles, $r(41) = -0.10$, $p = .521$ [*95% CI* = -0.39 – 0.21].

Figure 3.7. Exploratory analysis of hypothesis 1: mean matching percentage of phonemes across nap condition and test time.



Note: The horizontal bar represents the mean, the borders around the data points are smoothed density curves, and the error bars represent the standard deviation.

3.5.2. Recognition memory

As outlined in the Method section (Section 3.2.2.7.), we included an old-new categorisation recognition memory test but did not include this in the pre-registered analyses as we were not confident in making firm predictions about recognition memory. However, given that some word learning studies have reported a beneficial impact of sleep following learning on old-new categorisation accuracy and response times (e.g., Tamminen et al., 2010), we conducted exploratory analyses to establish whether our nap manipulation effected accuracy and response times in this task. Accuracy was measured in d' to account for response bias. This was calculated by subtracting the z-transformed proportion of accurate "old" responses to trained words (i.e., hits) from the z-transformed proportion of inaccurate "old" responses to foils (i.e., false alarms). Table 3.3 shows the d' values in each condition, as well as hit rates, false alarm rates, and proportion of correct responses. A 2 x 2 ANOVA with nap condition (nap vs. no nap) and test time (immediate vs. delayed) showed no significant main effect of nap condition, $F(1, 44) = 0.35$, $p = .558$, $\eta_p^2 = .008$. There was also no significant interaction between condition and test time, $F(1, 44) = 1.77$, $p = .191$, $\eta_p^2 = .039$. However, there was a significant main effect of test time, $F(1, 44) = 23.78$, $p < .001$, $\eta_p^2 = .351$, with accuracy being higher in the immediate test compared to the delayed test.

Another 2 x 2 ANOVA was conducted using the reaction times for correct responses to the words, on which participants were trained. Removing the incorrect responses resulted in the removal of 18.61% of the data. Using the same cleaning procedure as Tamminen et al. (2010), reaction times slower than 3000 milliseconds and faster than 500 milliseconds were removed. This resulted in the removal of a further 1.49% of the data. The remaining reaction times were log transformed, and a mean was calculated per participant per session. There was

no significant main effect of nap condition, $F(1, 43)$[1] = 0.35, p = .558, η_p^2 = 0.008. There was also no significant interaction between test time and condition, $F(1, 43) = 0.14$, p = .715, η_p^2 = 0.003. However, there was a significant main effect of test time, $F(1, 43) = 17.19$, p <.001, η_p^2 = 0.286, with reaction times being faster in the delayed test than the immediate test (see Table 3.3. below).

Table 3.3. Means and standard deviations (in parentheses) for performance in the old-new categorisation task

	Nap		Wake	
	Immediate	Delayed	Immediate	Delayed
Accuracy (%)	86.67 (4.10)	83.83 (4.03)	85.71 (4.15)	82.42 (4.48)
Hit rate	0.84 (0.15)	0.80 (0.15)	0.81 (0.18)	0.79 (0.18)
False-alarm rate	0.11 (0.08)	0.13 (0.11)	0.10 (0.09)	0.15 (0.11)
d’	2.58 (0.94)	2.32 (0.89)	2.59 (0.87)	2.16 (0.79)
Reaction time (ms)	1302 (151)	1231 (167)	1318 (204)	1253 (228)

Note: only reaction times to accurate responses are used. ms = milliseconds

3.5.3. WASI vocabulary scores

Researchers have argued that a participant’s vocabulary size predicts word learning success, at least in children (e.g., James et al., 2017). Given that we collected a measure of vocabulary size, we sought to extend these findings to our sample that consisted of young adults. For these exploratory analyses, raw WASI vocabulary scores are used, as t-scores are

[1] After exclusion of extreme RTs and incorrect responses, one participant had no data left in one of the four test sessions, therefore this participant is not included in the ANOVA.

age adjusted, and our sample crosses an age boundary for the WASI t-scores, however, we do not anticipate any developmental differences in a sample of 18-24-year-olds. An average of performance across all test sessions for each test task was calculated per participant. However, in the lexical competition analyses, only the delayed test data were included as lexical competition effects typically only emerge after a period of consolidation. This was then correlated with the raw WASI vocabulary scores to examine whether participants with better vocabulary have a higher ability to learn the novel words

As we were conducting four two-tailed correlation tests, alpha level was Bonferroni corrected to 0.005. There was no significant correlation between vocabulary and cued recall performance either when measured as proportion of accurate responses, $r(43) = .21$, $p = .169$ [*95% CI* = -0.09 – 0.47] or when measured in matching percentage, $r(43) = .04$, $p = .781$ [*95% CI* = -0.25 – 0.33]. There was also no significant correlation between vocabulary score and old-new categorisation d' scores, $r(43) = .19$, $p = .208$ [*95% CI* = -0.11 – 0.46], or between vocabulary scores and the magnitude of lexical competition effects in the delayed test, $r(43) = 0.27$, $p = .077$ [95% *CI* = -0.03 – 0.52].

Next, we sought to establish whether participants with lower vocabulary size might benefit from the nap more than participants with higher vocabulary size. For cued recall and old-new categorisation, we first averaged the data across the immediate and delayed sessions, and then subtracted the wake condition from the nap condition, to yield the difference between nap and wake across test times. As we conducted three two-tailed correlation tests, alpha level was Bonferroni corrected to 0.007. There was no significant correlation between vocabulary size and the nap effect in cued recall matching percentage, $r(43) = 0.08$, $p = .579$, [*95% CI* = -0.21 – 0.37], or in old-new categorisation accuracy, $r(43) = -0.18$, $p = .231$, [*95% CI* = -0.45 – 0.12]. For lexical integration we calculated the magnitude of change in the lexical competition effect from the immediate to the delayed test sessions, and again

subtracted the wake condition from the nap condition. There was no significant correlation between vocabulary size and the nap effect, $r(43) = 0.27$, $p = .073$, [*95% CI* = -0.03 – 0.52].

3.5.4. Frontal slow spindles

In our pre-registered analyses, we focussed on fast sleep spindles as Mander et al. (2011) found fast spindles to be associated with restoration of episodic encoding capacity, and no such association involving slow spindles. In exploratory analyses, we sought to establish whether the association between fast spindles and the emergence of lexical competition effects that we observed is specific to fast spindles, or whether it would also be observed in slow spindles. Therefore, we repeated the spindle analyses with slow (11 - 13.5 Hz) frontal spindles in Stage 2 sleep. In all other respects the analysis was identical to the pre-registered analysis. As we were conducting three two-tailed correlation tests, alpha level was Bonferroni corrected to 0.007. Similarly, to the findings in the original analyses, we did not find a significant relationship between slow frontal spindles and cued recall accuracy when measured in proportion accuracy, $r(41) = .08$, $p = .631$, [*95% CI* = -0.23 – 0.37], or when measured in matching percentage, $r(41) = -0.14$, $p = .370$, [95% CI = -0.42 – 0.17]. There was a positive relationship between the number of slow frontal spindles in Stage 2 sleep and overnight change in lexical competition, $r(41) = 0.39$, $p = .010$, [*95% CI* = 0.10 – 0.62], that was similar in size to the same association with fast spindles, but which did not survive the Bonferroni correction here.

3.5.5. SASS-Y retrospective sleep diary

Recent studies have shown that restricted sleep and sleep/wake variability can have detrimental effects on adolescents' and young adults' ability to learn (Lo et al., 2017, Phillips et al., 2017). Such restriction and variability could be captured by the SASS-Y as it evaluates

sleep both on weekdays and weekends. Table 3.4. shows the measures collected in the SASS-Y averaged across all participants. We removed SASS-Y questionnaires with missing data, allowing only complete sets of within-subjects data (N = 39). Using the difference between weekday and weekend total sleep time as a proxy for sleep restriction, we predicted that participants with higher levels of sleep restriction would learn fewer new words. We conducted two-tailed exploratory Pearson's correlations between sleep restriction and cued recall matching percentage in the immediate test, as this appears to be the most sensitive test of sleep effects in our data. As we were conducting two two-tailed correlation tests, alpha level was Bonferroni corrected to 0.01. We found no significant correlation when using data from the nap condition, $r(37) = 0.09$, $p = .571$, [95% CI = -0.23 – 0.40] or the wake condition, $r(37) = 0.05$, $p = 0.773$, [95% CI = -0.27 – 0.36].

Table 3.4. Summary statistics (standard deviation in parentheses) on the SASS-y questionnaire

	Nap condition		Wake condition	
Measure	Week	Weekend	Week	Weekend
Bedtime	23:21:33	00:18:28	23:31:10	00:41:56
Time of attempting to fall asleep	00:09:07	01:02:27	00:20:24	01:07:42
Sleep onset latency (mins)	33.6 (26.0)	36.1 (29.3)	28.7 (21.3)	26.5 (24.0)
Number of awakenings	1.49 (1.34)	1.13 (0.98)	1.28 (1.22)	0.96 (1.05)
Wake after sleep onset (minutes)	9.24 (10.86)	9.60 (13.62)	7.38 (6.84)	7.26 (11.04)
Time of final awakening	08:31:33	09:22:19	08:42:04	09:30:47
Time of leaving bed	09:20:39	10:14:45	09:29:07	10:32:34
Sleep quality	2.82 (0.64)	2.87 (0.62)	2.77 (0.78)	2.82 (0.76)
Total sleep time (hours)	7.66 (1.46)	7.57 (1.89)	7.76 (1.40)	7.82 (1.51)

Sleep restriction in hours	-0.09 (1.38)	0.06 (1.26)

Note: Sleep quality is ranked on a five-point Likert scale from very poor (0) to very good (4), higher scores indicate better sleep quality. Total sleep time was calculated from the difference between final awakening and time of sleep onset (time of attempting to sleep plus sleep onset latency), with wake after sleep onset deducted.

3.5.6. Habitual napping

Using two-tailed exploratory independent t-tests, we examined whether there was a significant difference between habitual nappers (one or more nap per week) and non-habitual nappers (napping less than once per week on average) slow-wave sleep duration or Stage 2 fast frontal sleep spindles. There was no significant difference between habitual and non-habitual nappers in slow-wave sleep duration, $t(43) = -1.43$, $p = .16$, non-habitual nappers had an average of 26.6 minutes slow-wave sleep ($SD = 14.4$), and habitual nappers had 20.27 minutes ($SD = 13.1$). There was also no significant difference in the average number of sleep spindles, $t(41) = -1.32$, $p = .20$, non-habitual nappers had slightly more spindles on average ($M = 43.04$, $SD = 28.1$) than habitual nappers ($M = 33.52$, $SD = 18.2$).

3.5.7. Bayes Factors

We conducted exploratory post-hoc Bayesian analyses to examine the extent to which our non-significant results in the pre-registered analyses support the null hypotheses. Bayesian ANOVAs and correlations were calculated in JASP 0.16.4.0 (JASP Team, 2022) following van Doorn er al. (2021) and van den Bergh et al. (2020). In all analyses we used the default priors implemented in JASP.

3.5.7.1. Hypothesis 1: Napping compared to wake before encoding will result in better episodic memory of newly learned words

For Hypotheses 1 and 2 we conducted a Bayesian version of the ANOVA in the original pre-registered analyses. The Bayes factors (BF) of the effect analysis are reported in Table 5. The BFs associated with the nonsignificant effects are both higher than 1/3 and below 3 and therefore provide only weak evidence for the null hypothesis (van Doorn et al., 2021).

3.5.7.2. Hypothesis 2: Napping compared to wake before encoding will result in larger consolidation effects in the integration of newly learned words with existing known words

The BFs associated with the main effects are below 1/3 and therefore show moderate evidence in favour of the null hypothesis, and the BF associated with the interaction is below 1/10 and therefore provides strong evidence for the null hypothesis (Table 3.5.).

Table 3.5. Analysis of effects for the repeated measures Bayesian ANOVAs

	P(incl)	P(excl)	P(incl|data)	P(excl|data)	**BF*Inclusion***
Hypothesis 1					
Condition	0.60	0.40	0.43	0.57	**0.51**
Time	0.60	0.40	1.00	$1.44x10^{-6}$	**464249.53**
Condition * Time	0.20	0.80	0.22	0.79	**1.09**
Hypothesis 2					
Condition	0.60	0.40	0.22	0.79	**0.18**
Time	0.60	0.40	0.22	0.78	**0.18**
Condition * Time	0.20	0.80	0.01	0.99	**0.05**

3.5.7.3. Correlational analyses

Bayes Factors were calculated for the Pearson correlations reported under the pre-registered analyses (Table 3.6.). All non-significant correlations show moderate evidence in favour of the null hypothesis. The BF for the significant correlation between spindle activity and increase in the lexical competition effect is above 3 and therefore associated with moderate evidence for the alternative hypothesis.

Table 3.6. Bayes Factors (BF10) for the pre-registered correlational hypotheses

Hypothesis	Bayes Factor (BF_{10})
3a. Slow-wave sleep duration will be correlated with cued recall accuracy	0.33
3b. Slow-wave spectral power will be correlated with cued recall	0.32
3c. Slow-wave sleep duration will be correlated with lexical competition	0.24
3d. Slow-wave spectral power will be correlated with lexical competition	0.20
4a. Fast frontal Stage 2 spindle activity will be correlated with cued recall	0.24
4b. Fast frontal Stage 2 spindle activity will be correlated with lexical competition	6.31

3.6. Discussion

In the current study we examined the role of a daytime nap in restoring encoding capacity. As discussed in Chapter 1 (Section 1.2.), sleep may serve two main functions in relation to learning and memory: consolidating information learned *before* sleep and restoring encoding capacity for information learned *after* sleep. Most research and theory are based around the sleep consolidation literature (see discussion in in Chapter 1, Section 1.2.1.), which is explored in relation to adolescent sleep restriction in Chapter 2. More recently, research has considered the role of sleep *prior* to encoding (e.g., discussed in Section 1.2.2.). Therefore, the current study set out to test four sets of hypotheses based on predictions made by theories that suggest sleep restores capacity to encode new episodic memories (Tononi & Cirelli, 2003; 2014, Saletin & Walker, 2012). The current study used a within-subjects

design. In the nap condition young adult and older adolescent (18-24-year-old) participants took a short daytime nap before being trained on a set of new words. In the no nap condition participants remained awake for an equivalent duration and watched a video instead of napping. Participants were tested on their memory for the newly encoded words in an immediate test, and in a next-day delayed test. Our pre-registered analyses did not find a significant main effect of napping compared to wake on test performance across any task, although our exploratory analyses did suggest better immediate cued recall in the nap condition when using a more sensitive measure of accuracy. We did not find a significant correlation between test performance and slow-wave activity. However, in line with some earlier reports (Mander et al., 2011) we found that more Stage 2 fast frontal spindles during the daytime nap were significantly correlated with greater increase in lexical integration from the immediate to the delayed test.

Our first hypothesis stated that if the nap restores encoding ability, participants should be able to encode more words after the nap than after wake, and therefore show higher levels of recall in a cued recall task following training in the nap condition. Our pre-registered analysis did not support this hypothesis: there was no statistically significant difference between the number of words recalled correctly in the nap condition and the no nap condition. However, binary cued recall accuracy (correct or incorrect) lacks sensitivity in the current study due to the presence of floor effects. Importantly, this approach fails to account for partially correct responses and instead treats fully incorrect responses (or no response at all) and responses that only minimally deviate from the correct word as equivalent. Therefore, we (Ricketts et al., 2021) and others (e.g., Frances et al., 2020) have recently adopted more sensitive approaches to scoring recall data. In exploratory analyses, we used matching percentage (Kurdziel & Spencer, 2016), a measure based on Levenshtein distance, to calculate a graded measure of recall accuracy. In this analysis, participants' responses were significantly more

accurate in the nap condition compared to the no nap condition in the immediate test. In the delayed test, performance was equal regardless of whether a nap or wake preceded encoding. It is important to acknowledge that exploratory analyses are not hypothesis driven and must be treated with caution. Further research is needed to test the hypothesis that a nap before learning benefits cued recall when using our more sensitive measure of accuracy.

In both the pre-registered and exploratory cued recall analyses, there was significantly better performance in the delayed test than the immediate test. This may partially be explained by practice effects, especially as the old-new categorisation task may have served as additional training by providing one additional exposure to each new word in the immediate test. Memory consolidation processes during overnight sleep between the test sessions may also have contributed to the increase in recall levels. For example, Tamminen et al. (2010) showed that both cued and free recall benefitted from sleep-associated memory consolidation in the same word learning paradigm as used here, whereby recall levels increased after a night of sleep. Sleep-associated memory consolidation is thought to operate preferentially on weaker rather than stronger memories (e.g., S. Walker et al., 2019), at least in the absence of interference (Petzka et al., 2021). If consolidation processes occurring between the first and second test favoured the weaker condition (no nap), this could explain why the nap benefit in our exploratory analyses was no longer seen in the delayed test.

Our second hypothesis was that the increase in lexical competition effects from the immediate test to the delayed test, which reflects emerging integration of new and old information, would be greater in the nap condition compared to the no nap condition. We found no significant impact of the nap on the emergence of the lexical competition effects. While this appears to indicate that pre-learning sleep does not benefit later integration of new words in the mental lexicon, we note that the magnitude of lexical competition effects remained close to zero in the delayed test, where we expected to see effects significantly

above zero. It is not clear why this was the case. One possible explanation comes from an inspection of the overall reaction times in this task. Table 3.1. shows that the average reaction times in our study ranged between 1122 milliseconds and 1225 milliseconds. Tamminen and Gaskell (2008), who used the same task, the same word stimuli, and the same audio recordings of the stimuli, reported reaction times in the 970 milliseconds to 1050 milliseconds range (see Figure 2 in Tamminen & Gaskell, 2008). Quantification of the lexical competition effect requires participants to respond as close as possible to the uniqueness point of the spoken words (i.e., the point where the base word "cathedral" deviates from its new competitor "cathedruke"). If participants in our study responded substantially after this point, the magnitude of the lexical competition effect could be diluted. Given these considerations, we urge caution in interpreting the null effect observed regarding our second hypothesis.

We measured lexical competition effects in a lexical decision task. Another task commonly used for this purpose is pause detection, where participants are not asked to judge the lexicality of a word, but to detect a short pause inserted close to the uniqueness point of the spoken word (e.g., Dumay & Gaskell, 2007). The time it takes to make a decision about the presence of the pause is taken as a proxy for the level of lexical activation at the time with longer pause detection times indicating higher lexical activity (Mattys & Clark, 2002). The advantage of this task is that it is less reliant on explicit linguistic knowledge than lexical decision and may therefore be a more sensitive measure of lexical integration and its time course. In fact, studies using the lexical decision task have sometimes shown a different time course of lexical integration from studies using pause detection (e.g., Tamminen et al., 2010 vs. Dumay & Gaskell, 2007). Another reason for the absence of emerging lexical competition effects in our study could be low levels of explicit memory for the novel words. Walker et al. (2019) failed to find overnight increases in lexical competition when the number of exposures

to novel words during training was very low suggesting that sleep does not benefit integration of poorly encoded novel words.

The third and fourth hypotheses targeted the neural mechanisms of sleep-associated restoration of memory encoding capacity, and compared predictions made by two different theoretical accounts of restoration. Based on the synaptic homeostasis hypothesis (Tononi & Cirelli, 2014) we predicted, in our third set of hypotheses, an association between slow-wave activity (i.e., slow-wave sleep duration in hypotheses 3a and 3c; slow oscillation spectral power in hypotheses 3b and 3d) and both cued recall accuracy and the overnight emergence of lexical competition effects. The extension to the Complementary Learning Systems account proposed by Mander et al. (2011) and Saletin and Walker (2012) on the other hand proposes that restoration of encoding capacity is due to a process of hippocampal-neocortical dialogue as indexed by frontal fast sleep spindle activity. We therefore predicted in our fourth set of hypotheses an association between the number of frontal fast spindles and cued recall and the overnight emergence of lexical competition effects.

We did not find a significant relationship between the slow-wave sleep measures and the test tasks. This could suggest that slow-wave sleep is not the key neural mechanism associated with restoration of encoding capacity. However, before accepting such a conclusion we need to consider an important feature of the synaptic homeostasis hypothesis. The theory postulates a slow-wave activity control loop during sleep whereby increasing daytime learning results in increasing synaptic strength, which in turn results in increasing slow-wave activity during sleep. If it is the case that our participants engaged in little new learning before taking the nap, then there would be little need for reducing synaptic strength and little variability across participants in slow-wave activity during the nap. This might limit our ability to detect effects involving slow-wave sleep. This point further highlights a potentially important difference between our study and that of Mander et al. (2011). In

Mander et al., participants encoded a large set of face-name stimulus pairs before napping. This ensures that episodic memory capacity is saturated before sleep and that restoration processes are engaged. In our study we did not give participants a memory task before the nap, and it is therefore possible that for some participants, restoration processes were not needed or were engaged to a limited degree. Furthermore, whilst our wake condition was designed to roughly match with the nap condition in linguistic input without letting participants sleep, it may be the case that the wake condition due to its passive nature allowed some degree of slow-wave activity that may have diluted the benefit of sleep. Brokaw and colleagues (2016) showed that even a brief period of wakeful rest allowed memory consolidation of a previously learned short story to occur, and that this consolidation was associated with slow oscillatory EEG activity. Our participants may have benefitted from such activity while watching the video, in contrast to Mander et al. (2011) who allowed participants to complete usual daily activities during wake.

Turning to our predictions concerning sleep spindles, we did not find a significant correlation between sleep spindles and cued recall performance. However, we did find a positive correlation between fast frontal spindles and overnight increase in the lexical competition effect, as predicted. This supports the view put forward by Mander et al. (2011) and Saletin and Walker (2012) that hippocampal-neocortical dialogue shifts newly encoded memories from hippocampal to cortical-dependence, thus restoring hippocampal episodic learning capacity. It is not clear however why spindles were only associated with emerging lexical competition and not cued recall (although Figure 6 suggests both associations are in the same direction). One possibility is that the observed association is a proxy for overnight memory consolidation. Sleep spindle activity is trait-like such that individuals show stable spindle activity levels from night to night (e.g., De Gennaro et al., 2005). Furthermore, Mylonas et al. (2020) recently showed that there is a strong correlation between an

individual's spindle density observed during a daytime nap, and spindle density during overnight sleep. Therefore, participants with more spindles during the nap in our study may also have had greater levels of spindles during their nocturnal sleep. Tamminen et al. (2010) showed that spindle activity overnight is correlated with larger increases in the lexical competition effect. If spindle activity during a nap is predictive of spindle activity overnight, the association we observed could reflect overnight emergence of lexical integration rather than an impact of the nap on lexical integration. Because this aspect of our experimental design is correlational, further research is needed to clarify possible causal relationships.

The pre-registered analyses targeted fast sleep spindles. This choice was motivated by Mander et al. (2011) who found a role for fast but not slow spindles in restoring encoding capacity. Furthermore, in the memory consolidation literature it has been argued that fast spindles are involved in consolidation processes while slow spindles may not be (e.g., Mölle et al., 2011). In exploratory analyses we sought to establish whether the fast spindle specificity observed by Mander et al. (2011) was also present in our data. We found no statistically significant association between slow spindles and overnight increase in lexical competition effects after Bonferroni correction, suggesting the association we found may be specific to fast spindles. However, the numerical size of the association was very similar for fast (r = 0.40) and slow (r = 0.39) spindles, and it therefore seems premature to rule out a role for slow spindles in restoration of encoding capacity based on our data.

The trait-level aspect of sleep spindles (e.g., De Gennaro et al., 2005; Mylonas et al., 2020) also suggests that it is important to consider individual differences as well as group difference in analyses. Our analyses revealed an association between individual differences in sleep architecture and performance that would have been masked had we restricted our analyses to group comparisons alone. Specifically, the number of fast frontal spindles during Stage 2 sleep of the nap was a significant predictor of changes in lexical competition from the

immediate to the delayed test. The importance of individual differences has also been observed in experimental manipulations of sleep architecture. For example, Ong et al. (2018) found that participants with greater increases in slow oscillations following acoustic stimulation showed better memory recollection. As in our study, Ong et al.'s (2018) group-level analyses did not indicate an impact of slow-oscillation stimulation on cognitive performance. Therefore, future studies may benefit from investigating the individual characteristics that predict the ability of sleep to restore encoding capacity, even when group-level differences are less evident. Such research will require larger sample sizes and perhaps more sensitive measurement.

Other sleep stages may also be of interest in future research. For example, in the current study, less than half (44%, $N = 20$) of participants achieved REM sleep, with the average duration of REM being 12 minutes for these participants. The low levels of REM sleep in the current study shows that most participants did not achieve a full sleep cycle during the nap. Therefore, different effects might emerge after a longer sleep duration, allowing a full sleep cycle. On average our participants spent less time in REM than participants reported in Mander et al. (2011), with our participants reaching an average of 5.5 minutes in REM while Mander et al. reported an average REM duration of 17.5 minutes. The average duration of the other sleep stages is very similar across the two studies. While we cannot exclude the possibility that our results were weakened by low levels of REM sleep, we note that none of the current theories of how sleep restores memory encoding capacity ascribe a role for REM sleep.

We also note that our sample of participants varied in their napping habits although the majority (71%) were habitual nappers. We are not aware of published literature examining whether habitual and non-habitual nappers might differ in the benefits they gain from a daytime nap respective of subsequent learning. A recent study however has shown that

habitual and non-habitual nappers both benefit from a daytime nap occurring after learning (Leong et al., 2021) Even if habitual napping does not influence the memory benefits of a nap, it can influence sleep architecture during a daytime nap and therefore introduce statistical noise to the PSG correlations we were investigating. McDevitt and colleagues (2012) showed that habitual nappers have more Stage 1 and 2 sleep and less slow-wave sleep than non-habitual nappers. We conducted exploratory analyses to check whether habitual and non-habitual nappers in our sample differed in our key measures of sleep architecture. Habitual napping was not significantly associated with the duration of slow-wave sleep or average number of fast frontal Stage 2 sleep spindles, suggesting that this is unlikely to be a substantial confound in our study.

In other exploratory analyses, we examined the effects of test time and napping on recognition memory. Due to mixed results in previous studies using old-new recognition, we were unable to make strong predictions for this task. Some studies have found consolidation benefits for recognition memory (e.g., Tamminen et al., 2010, using the same word learning paradigm) as well as benefits of pre-encoding sleep (e.g., Mander et al., 2011). Our results show significantly better recognition (higher d') in the immediate compared to the delayed test of the old-new recognition task suggesting there was no consolidation benefit in the delayed test. When examining reaction times towards accurate responses, participants responded faster in the delayed test than the immediate test in both conditions, possibly indicating some degree of consolidation demonstrated by speed of recognition. Alternatively, the decreases in accuracy accompanied by faster responses could reflect a speed-accuracy trade-off in the delayed test. This could possibly be underpinned by a change in response strategy whereby participants in the immediate test relied more on recollection than familiarity resulting in slow but accurate responses, but vice versa in the delayed test resulting in fast but inaccurate responses due to both old and new items now being familiar to

a degree. Contrary to Mander et al (2011), we did not find a significant benefit of pre-encoding sleep on recognition memory. However, the paradigms and number of items used differ between our studies. Participants in Mander et al.'s research were presented with 100 encoded face-name pairs and 50 novel pairs during testing whereas we use 22 trained non-words and presented 22 similar sounding foils. Larger numbers of trials, and a different ratio of "old" to "new" trials in Mander et al.'s (2011) study may allow for greater sensitivity than our recognition test, which perhaps allowed for significant effects to emerge – as seen in our exploratory analyses using graded cued recall accuracy. Additionally, greater levels of training both in regard to depth and breadth of materials studied may be beneficial in future research to reduce floor effects, and thus increase sensitivity.

Since we also measured existing vocabulary knowledge, we were able to conduct exploratory analyses investigating associations between existing vocabulary knowledge and ability to learn new words, and between vocabulary knowledge and any nap benefits. We found no significant correlation between vocabulary knowledge and any of the tests of word learning. Although these exploratory analyses were not motivated by a hypothesis, the lack of significant correlation contrasts with experimental studies that show an association between existing vocabulary knowledge and subsequent word learning (e.g., Ricketts et al., 2011). More generally, the lack of an association seems somewhat at odds with the notion of "Matthew effects" (Stanovich, 1986) in vocabulary development, whereby greater existing vocabulary knowledge is associated with faster growth in vocabulary, or the 'rich getting richer'. However, Matthew effects are not consistent across studies (Pfost et al., 2014) and may diminish after childhood (Ricketts et al., 2020). In addition, research considering Matthew effects tends to focus on the long-range vocabulary growth that is observed across months or years of development, rather than the short-range vocabulary learning that occurs in carefully controlled experimental paradigms. Future research carefully manipulating the

nature of existing vocabulary knowledge, and studying these relationships across the lifespan, would be better equipped to examine these effects (e.g., James et al., 2021; Ricketts et al., 2020).

Lastly, we conducted exploratory analyses correlating a proxy for sleep restriction calculated from total self-reported sleep time during weeknights compared to weekends and graded measure of cued recall accuracy. Relevant to this is the social jetlag hypothesis, which suggests that weekday sleep restriction occurs due to a clash between the timing of social commitments (e.g., lectures) and biological sleep preferences (e.g., circadian rhythms) and therefore sleep is longer on the weekends to compensate (Wittmann et al., 2006). Sleep restriction has been related to lower performance across multiple cognitive domains, including declarative memory (e.g., Lowe et al., 2017). Based on such findings we speculated that sleep restriction might predict participants' ability to encode new words. This is of particular interest, as daytime napping has been found to compensate for the effects of adolescent sleep restriction (Cousins, van Rijn et al., 2019). Adolescent participants were allocated 6.5-hours nocturnal time in bed, or 5-hours with a 1.5-hour daytime nap (i.e., a split sleep schedule). Participants with a split sleep schedule were found to have improved memory for pictures encoded after the daytime nap. The groups did not significantly differ in regard to memory for information learned in the morning, prior to the intervention. These results suggest that a daytime nap may compensate for insufficient nocturnal sleep duration. We did not find any significant relationships between sleep restriction and test tasks in these analyses, however. Notably, we did not see a large difference between weekend and weekday sleep in our university student sample and this might have obscured any effects. Most of our data collection took place during spring and summer 2021 when teaching and exams took place online to mitigate COVID-19 risks. It is well documented that social jetlag dramatically reduced during the pandemic (e.g., Blume et al., 2020) thanks to the flexibility of home

working, and it is likely that our participants benefitted similarly. Alternatively, the habitual (past week) nature of the sleep diary used may have masked effects of night-to-night sleep variability during the week, which has been associated with poorer academic grades in university students (Phillips et al., 2017). Such night-to-night variability may impact learning and memory in different ways than weekday/weekend sleep restriction. Future research may benefit from night-to-night recordings for a week prior to the testing sessions, for example using actigraphy and daily diaries to explore these relationships, which are outside the scope of the current study. Examining the relationship between daytime napping and cognition in university students or older adolescents is of particular interest. University students will likely have increased independence and responsibility for their sleep schedules during the transition of moving away from the family home. Lower levels of parental influence over sleep schedules (i.e., greater bedtime autonomy) has been associated with later bedtimes, which mediated the relationship between increasing adolescent age and shorter (often insufficient) sleep durations (Tashjian et al., 2019). As such, older adolescents at college, university, or in the workplace, may be particularly susceptible to the influences of sleep restriction or variable sleep patterns, which may have impacts on cognitive functioning.

When young adults (mean age ~25 years old) were sleep restricted, a 30 minute nap was shown to have benefits for memory encoding following the nap, whilst this was not seen for 10 or 60 minute naps; however, all naps were associated with increased positive affect and greater alertness (Leong et al., 2023). Therefore, perhaps different nap durations would observe different relationships in the current word-learning paradigm. Recent research has used the same non-word learning paradigm to explore memory consolidation during a 90 minute daytime nap for children (10-12-years-old) and adults and found that a daytime nap protected memory for newly learned words, but this was not significantly associated with slow-wave sleep or sleep spindles (van Rijn et al., 2022). Van Rijn et al. (2022) did not

observe any lexical integration occurring over the nap consolidation period, which may suggest that different sleep durations are required, or longer time between testing, to observe lexical competition effects.

In conclusion, our pre-registered analyses did not find a significant main effect of a daytime nap restoring ability to encode novel words. Exploratory analyses, using a graded measure of cued recall accuracy revealed that napping may benefit encoding in the immediate test, but that this benefit is not sustained at the next day test. We found no evidence that napping prior to encoding benefited lexical integration. Despite clear predictions made by the synaptic homeostasis hypothesis (Tononi & Cirelli, 2014), we found no significant evidence that pre-encoding slow-wave activity restored encoding capacity nor ability to integrate new and old information. Yet, we found that a greater number of Stage 2 frontal fast sleep spindles predicted greater overnight improvements in integration of new and old information from the immediate to the delayed test. We conclude that further research is needed to verify whether a daytime nap restores capacity to encode new words. Such research needs to measure episodic memory recall in a graded manner to increase sensitivity and reveal potential effects. The neural mechanisms underpinning cued memory recall were not identified in our current study. However, our spindle data suggest that Stage 2 fast sleep spindles are a likely neural mechanism for restoring the capacity to integrate new words in the mental lexicon. As discussed in Chapter 2, the timing of social commitments, such as education, is often in direct opposition to adolescents preferred timing determined by their biological circadian rhythms. Therefore, these findings also have potential practical importance. Some researchers have suggested that daytime napping should be incorporated into educational settings such as schools and universities, due to the benefits of napping for learning and memory (Cousins et al., 2019). As our research shows some support for the benefits of a pre-encoding nap in language learning in a university sample, future research of

these relationships in applied educational settings and across development (e.g., with children and adolescents) may prove valuable. Chapter 4 and Chapter 5 also consider the role of sleep prior to encoding, through self-reported measures in online studies.

Chapter 4: Sleep, affect, and emotional memory biases

4.1. Introduction

4.1.1. Sleep and memory

There is much evidence that sleep is important for learning and memory (Rasch & Born, 2013). The majority of previous research has focused on the consolidation of memories learned prior to sleep (e.g., Section 1.2.1.). However, recent research, including the nap study described in Chapter 3, has investigated how sleep supports subsequent learning and memory by restoring the brain's encoding capacity (Antonenko et al., 2013; Ong et al., 2020). When slow-wave oscillations were stimulated during a daytime nap, participants subsequently encoded more word-pairs after the nap than participants who received sham stimulation (fewer slow-wave oscillations; Antonenko et al., 2013). The study by Antonenko et al. (2013) suggests that slow-wave oscillations are the mechanism behind the benefit of sleep for restoring the brain for subsequent learning, as predicted by the synaptic homeostasis hypothesis (Tononi & Cirelli, 2003, 2014). The daytime nap was found to restore encoding capacity for a range of declarative memory tasks (e.g., memory for images, word-pairs and word lists) but not for procedural memories, suggesting these relationships may be specific to hippocampal memories. However, other research has failed to find a correlation between encoding abilities after a nap and slow-wave sleep or slow-oscillation power (Ong et al., 2020). Instead, participants who took a daytime nap had greater levels of hippocampal activity when encoding new word-pairs than participants who did not nap, suggesting sleep functioned to restore the hippocampus for subsequent learning. For participants that napped, sleep spindle count was significantly associated with subsequent increases in hippocampal activity during encoding (after the nap) and improved encoding performance after sleep (Ong et al., 2020). This study demonstrates the benefits of sleep, rather than wake, in restoring the

brain for subsequent learning. The study described in Chapter 3 also finds that a higher density of sleep spindles prior to encoding is associated with greater improvements in overnight lexical integration. Additionally, a daytime nap prior to encoding was associated with significantly more accurate recall of the non-words, than wake prior to encoding, however, the neural mechanisms underpinning this association were not identified in Chapter 3. Regardless of the differing neural mechanisms, both Ong et al. (2020), Antonenko et al. (2013), and our study in Chapter 3 support the hypothesis that sleep functions to restore the brain for subsequent learning.

The current research adds to the developing literature regarding the impact of sleep before learning, testing the prediction made by the synaptic homeostasis hypothesis that higher quality and longer sleep before learning is associated with improvements in episodic memory. Whilst this study is conducted online (and therefore, unable to test the neural mechanisms behind the relationship), it is novel in exploring naturalistic sleep patterns regarding the restoration of learning capacity. Measuring longer-term sleep patterns is important as they may contribute to restoring encoding capacity, learning and memory beyond what is found in single night studies. For example, in Chapter 2, we demonstrate greater evidence for sleep restriction in measures of multi-night sleep (e.g., sleep diaries and actigraph data) than single night recordings (e.g., polysomnography).

To date, no research to our knowledge, has investigated the role of naturally occurring long-term sleep patterns in restoring subsequent encoding capacity. However, one study investigated experimental sleep restriction over multiple nights, and found that adolescents (15-18-years-old) had poorer memory encoding when sleep-restricted in comparison to participants allowed a healthy sleep duration (Cousins et al., 2018). Participants were sleep-restricted (allowed only five hours in bed) or well-rested (allowed nine hours in bed) for five days, before encoding pictorial stimuli. After three nights of recovery sleep (nine hours per

night in bed), participants were tested on their memory for the stimuli they encoded whilst sleep restricted. Participants who were sleep restricted before encoding had poorer memory than participants who were well-rested before encoding. This supports the hypothesis that extended wakefulness and lack of sleep is detrimental to encoding abilities (e.g., synaptic homeostasis hypothesis), demonstrating the importance of a healthy sleep duration for learning and memory, and the consequences of insufficient sleep.

In a similar multi-night experimentally manipulated sleep study, participants who were allowed a daytime nap (one-and-a-half hour) after nocturnal sleep restriction (five hours in bed at night) had better memory than participants who were only allowed nocturnal sleep and no nap (six-and-a-half hours in bed at night; Cousins, van Rijn, et al., 2019). This sleep manipulation was designed to experimentally represent a school week, in that sleep restriction occurred for five days, similarly to a school week. Pictorial stimuli were encoded at the end of the "week" and recalled after a two-day "weekend" of recovery sleep. Three additional days of experimental sleep restriction followed, with factual stimuli encoded each day. This was followed by a two-day "weekend", with retrieval tasks on the first recovery night. The nap group performed better than the group that only had nocturnal sleep. However, encoding was only improved for stimuli learned after the nap (in the afternoon, not in the morning). This is theoretically important as the results may be driven by a second period of desaturation, or synaptic downscaling, in the brain during the nap, thus supporting the suggestion made by the synaptic homeostasis hypothesis that sleep functions to support subsequent learning.

Multi-night studies of sleep restriction are important for examining the cumulative influences of sleep on learning and memory. Sleep restriction refers to achieving sleep but of an insufficient duration, whilst total sleep deprivation refers to having no sleep at all. Most studies measure the impact of a single night of total sleep deprivation, which does not reflect

the more realistic issue of the cumulative influence of multiple nights of sleep restriction (Cousins et al., 2018). The multi-night studies of sleep restriction that do exist typically use experimentally manipulated sleep (e.g., Cousins et al., 2018; Cousins, Wong, et al., 2019), rather than naturalistic sleep patterns. However, the current study used the Pittsburgh Sleep Quality Index to measure self-reported habitual sleep that occurred naturally during the past month without experimentally manipulating sleep. Global measures of sleep such as the Pittsburgh Sleep Quality Index, which include component scores related to multiple sleep variables such as sleep duration and subjective quality, may explain some additional variance in the relationship between sleep and memory, which is not explained by sleep duration alone. Subjective sleep quality measures how well a person feels they have slept, which is a unique sleep phenomenon that cannot be captured with objective measures such as polysomnography (Dewald et al., 2010). In naturalistic, correlational research, poor sleep quality has been associated with poorer academic attainment (Curcio et al., 2006; Dewald et al., 2010). Since academic attainment requires the acquisition of factual information and rule-based learning (Gomez Fonseca & Genzel, 2020), sleep quality may be related to learning and memory performance.

Therefore, the current study investigates predictions made by the synaptic homeostasis hypothesis that better sleep before learning functions to improve subsequent memory performance. This study is novel in the inclusion of naturalistic self-reported sleep quality and sleep duration measures used to test this hypothesis. Additionally, this study addresses a gap in the literature by investigating the impact of both last night's sleep and habitual sleep patterns over the past month, which has not yet been investigated regarding the restoration of encoding capacity.

4.1.2. Sleep and emotional memory

As discussed in Section 4.1.1 above, sleep before encoding is associated with better memory. However, the strength of this relationship may be altered by the type of stimuli encoded. For example, one property that may influence the relationship between sleep and memory is the emotional valence of the new memories. Unpublished research by Walker (as cited by Walker, 2009) has found that sleep deprived participants had poorer memory than well-rested participants, but this differed with emotional valence of the stimuli being encoded. Deficits in memory were particularly seen for positive and neutral stimuli, but memory remained relatively intact for negative stimuli. Emotional memory biases, in relation to sleep deprivation, are theorised to be related to increased amygdala reactivity towards emotional stimuli due to lower prefrontal cortex inhibition as a result of an uncalibrated noradrenergic system (Goldstein & Walker, 2014; Walker, 2009). As described in Section 1.3.1., healthy sleep (particularly rapid eye movement [REM] sleep), is supposed to recalibrate the noradrenergic system. When a person is sleep-deprived the lack of noradrenergic calibration is theorised to result in reduced medial prefrontal cortex control over the amygdala; thus resulting in amygdala over-reactivity (Goldstein & Walker, 2014). In essence, shorter sleep duration and poorer sleep quality will predict sub-optimal sensitivity towards emotional information, due to imbalances in the inhibitory control of the amygdala.

As discussed in Section 1.3.2. affective tagging at encoding may also influence selective consolidation. Harrington et al. (2017) proposed that poor sleep results in negative stimuli being 'tagged' relative to positive or neutral stimuli. Thus, negative stimuli are later consolidated to a greater extent, at the expense of positive memories, resulting in a negative memory bias (e.g., observed in depression). Therefore, the current study considers affect at encoding, which may be associated with 'tagging' and could predict immediate memory for emotional stimuli. The present study was also repeated again one-year later (outside of

lockdown conditions). Short sleep duration has been found to predict sub-optimal affect. Affect broadly refers to feelings, and encompasses emotions (transient forms of affect) and mood (longer lasting affect; e.g., A. D. Ong et al., 2017). The empirical literature appears to demonstrate a robust relationship between short adolescent sleep duration and lower positive affect (Short et al., 2020). After multiple nights of experimental sleep restriction, 15-19-year-olds had lower positive affect than well-rested participants (Lo et al., 2016). However, Lo et al. (2016) did not find a significant relationship between sleep restriction and negative affect, which they say may be due to their participants reporting that negative affect items such as "guilty" and "afraid" on the Positive and Negative Affect Scale were not applicable to them. In the current study, conducted during the COVID-19 pandemic with a larger sample, these negative affect items are more likely to be applicable.

Another possible explanation for Lo et al.'s (2016) null findings in relation to negative affect, is that the relationship between short sleep and positive affect has been found to have larger effect sizes than negative affect, depression, and anxiety (Short et al., 2020). Thus, Lo et al.'s (2016) study, using a between-subjects design with 56 participants, may have been underpowered to detect influences of sleep on negative affect. In another multi-night experimental sleep restriction study, using a more highly powered within-subjects design with 50 participants aged 14-17-years-old, the relationship between sleep and negative affect was observed (Baum et al., 2014). Participants were instructed to restrict their time in bed to 6.5 hours for five nights in one experimental week, and 10 hours in bed for five nights in another experimental week. Adherence to the protocol was monitored by actigraphy watches and self-report. When participants were sleep restricted, they had higher negative affect and poorer emotion regulation than when they had a healthy sleep duration. However, Baum et al. (2014) did not investigate affect changes on a daily basis, whereas Lo et al. (2016) found that lower positive affect was only seen on the second night of sleep restriction

onwards, suggesting that a single night of sleep restriction may not be sufficient to alter affect. Therefore, it is possible that last night's sleep may only influence affect if habitual sleep is also restricted, emphasising the importance of studying multiple-nights or habitual sleep patterns. Further, Baum et al. (2014) did not measure positive affect, and suggests that in future research, positive affect is also considered. The lack of research investigating sleep and positive affect is a current gap in the literature. Additionally, there is a current lack of theory explaining how sleep may influence positive affect. Goldstein and Walker's (2014) theory may explain the relationship between sleep restriction, poorer emotion regulation, and higher negative affect. However, amygdala over-reactivity may not be a sufficient explanation of the relationship behind sleep and low positive affect, due to the prediction that the amygdala would be over-reactive to all salient information. The Affective Tagging and Consolidation model (Harrington et al., 2017) aims to explain the differences in positive and negative memory biases via consolidation of negative over positive memories. However, it is important to consider factors that may influence 'affective tagging' and immediate memory at encoding.

As such, the relationship between sleep and affect may also have relevance for emotional memory biases. The associative network theory (Bower, 1981) suggests that memories are more likely to be recalled if they are congruent with emotional states (i.e., mood-congruent memory). If a participant is high in negative affect, they are more likely to recall negative information than positive or neutral information. In Vargas et al.'s (2019) study, participants encoded negative or neutral pictorial stimuli in the evening before being randomised to sleep deprivation (26 hours awake) or rest. The following morning, participants were tested on their memory, after a consolidation opportunity. It was found that participants had worse affect when sleep-deprived, and also had poorer neutral memory compared to well-rested participants, but equal (or preserved) memory for negative stimuli.

Therefore, Vargas et al. (2019) suggested that their results may be due to mood-congruent memory effects. Together, with the finding that sleep restriction is associated with lower positive affect (Lo et al., 2016) and higher negative affect (Baum et al., 2014), sleep may predict emotional memories via affect in accordance with associative network theory (Bower, 1981), and mood-congruent memory mechanisms (Vargas et al., 2019). In essence, good sleep will be correlated with improved memory for positively valenced word-pairs, and this is predicted to interact with increased positive affect. However, the influences of sleep on negative memory may have a shallower slope, due to negative mood in poor sleepers increasing mood-congruent memory for negative affect. Thus, sleep will not be found to have a significant impact on negative memory, due to mood-congruence.

However, the study by Vargas et al. (2019) is unable to distinguish between the potential influences of sleep on emotional memories (e.g., Walker, 2009) and mood-congruence effects (e.g., Bower, 1981). The mood-congruence hypothesis, as described in the associative network theory (Bower, 1981), would also suggest that participants high in positive affect will have greater recollection of positive relative to neutral or negative stimuli. However, Vargas et al. (2019) did not include positively emotionally valenced stimuli. Additionally, Vargas et al. (2019) examined the effects of sleep on the consolidation of emotional memory. Yet, the role of sleep before encoding emotional memories has been largely understudied (Tempesta et al., 2018). Another study, using a sleep deprivation paradigm and positively valenced stimuli, also found an interaction between sleep and valence (Tempesta et al., 2016). In Tempesta et al.'s (2016) study, participants were sleep deprived prior to encoding, and allowed recovery sleep prior to recall. Well-rested participants were found to have better recognition of neutral and positive images from film stimuli than the sleep deprived group, but both groups had comparable memory for negative stimuli, supporting the emotional memory bias found in consolidation research (e.g., Vargas

et al., 2019). Tempesta et al.'s (2016) study allowed a period of two nights recovery sleep before recall. Study designs with recovery sleep opportunities minimise the potential impacts of fatigue (following sleep deprivation) on memory recollection. However, this design also provides an opportunity for memory consolidation processes. Selective consolidation is theoretically relevant for the development of mental health difficulties (e.g., ATaC model, Harrington et al., 2017) and is of interest for further research. However, an additional test immediately after encoding may have been useful to examine the impact of sleep on initial encoding without overnight sleep consolidation. Furthermore, habitually poor sleepers may not have opportunity for sufficient recovery sleep. Therefore, memory processes (such as encoding and recall) under fatigued conditions may be of relevance.

In summary, better sleep before learning is hypothesised to predict better subsequent memory (e.g., synaptic homeostasis hypothesis; Tononi & Cirelli, 2003, 2006,2014). This study is novel in using naturalistic, self-reported sleep quality and duration to test this hypothesis; investigating both the role of sleep on the night immediately prior to the study in addition to habitual sleep patterns. The relationship between sleep and memory may differ for emotional stimuli. Sleep deprivation has been associated with lower positive and neutral memory, but equal negative memory in comparison to well-rested participants (Tempesta et al., 2016; Walker, 2009). The present study investigated two potential mechanisms behind the interaction between sleep and emotional valence in predicting memory. Emotional memory biases may be uniquely related to sleep, and imbalances in the inhibitory control of the amygdala (Goldstein & Walker, 2014; Walker, 2009). However, this theory also predicts over-reactivity to all emotional stimuli, including positive, particularly rewarding stimuli, such as money or highly calorific foods (Goldstein & Walker, 2014). Therefore, this theory may be insufficient at explaining the relationship between short sleep and low positive affect (Lo et al., 2016; Short et al., 2020) and the relationship between sleep deprivation and low

positive memory (e.g., Tempesta et al., 2016). Instead, mood-congruent memory mechanisms (e.g., Bower, 1981; Vargas et al., 2019) may explain this relationship, and may be the mechanism for 'affective tagging' in the Affective Tagging and Consolidation model (Harrington et al., 2017). Whilst this has been investigated to some degree in previous research (e.g., Tempesta et al., 2016), the current study is novel in exploring the independent roles of positive and negative affect, in relation to memory for emotionally valenced stimuli. The current study further fills a gap in the literature by investigating sleep before encoding of emotional stimuli, which is largely understudied. Additionally, data are collected during COVID-19 lockdowns, and again one-year later, to examine any changes in these relationships. Analyses with the data at one-year follow-up were exploratory, and are discussed later in Section 4.4.9.

4.1.3. Hypotheses:

1) Hypothesis 1: On the basis of the synaptic homeostasis hypothesis (Tononi & Cirelli, 2014), we hypothesised that sleep (last night, habitual) would predict performance such that worse sleep (higher PSQI scores) prior to encoding would be associated with lower overall recollection for word-pairs (Antonenko et al., 2013; Cousins et al., 2018; Ong et al., 2020).
2) Hypothesis 2: One explanation for a negative impact of sleep on memory is that a lack of sleep impedes memory for neutral and positive stimuli (Tempesta et al., 2016; Walker, 2009). Therefore, we predicted an interaction between sleep and the emotional valence of the word-pairs: participants with worse sleep would recall fewer neutral and positive word-pairs than participants with better sleep. but sleep would not predict the recall of negative word-pairs, or that the association with negative word-

pairs will be significantly weaker (Tempesta et al., 2016; Vargas et al., 2019; Walker, 2009).

3) Hypothesis 3: The relationship between sleep and emotional memories may also be related to the association between sleep and affect.

   a) Affect may predict emotional memory biases in accordance with associative network theory (Bower, 1981). Therefore, we predicted an interaction between affect and emotional valence of the word-pairs in predicting memory. In essence, we predict that positively valenced word-pairs will be recalled more by participants who are high in positive affect compared to those who are low in positive affect. Similarly, negatively valenced word-pairs will be recalled more by participants high in negative affect than those who are low in negative affect.

   b) Additionally, affect may be a mechanism linking sleep and emotional memory biases. Poor sleep has previously been associated with lower positive affect and higher negative affect (Short et al., 2020). Therefore, sleep may interact with the relationship hypothesised between affect and emotional memory biases (Hypothesis 3a), in that poorer sleepers will have worse affect than good sleepers and greater memory for negatively valenced stimuli compared to positive or neutral valenced stimuli. Therefore, a three-way interaction between sleep, affect, and emotional valence is hypothesised to predict memory.

4.2. Methods

4.2.1. Participants

Two-hundred and fifty participants took part in this study via Prolific. Participants were compensated £4.50 for their time (equivalent of £9 per hour for a 30-minute task). Participants were eligible to take part if English was their first language, they were aged between 18-24 years, and were resident in the UK at the time of the study. All participants

passed the necessary attention checks at high levels (above 50%) and therefore no data were removed due to lack of attention. The data were collected on 24th June 2020, during COVID-19 lockdowns. Replacement data for six participants who reported cheating or did not complete the study (e.g., due to technical errors) were collected on 26th June 2020. Participants provided informed consent by reading an information sheet and ticking the consent boxes. Ethical approval was granted by Royal Holloway University of London. Data collection was repeated at Timepoint 2 (one-year follow-up, outside of COVID-19 lockdown conditions). These data and exploratory analyses are discussed later in Section 4.4.9.

4.2.2. Materials

4.2.2.1. Demographics

Participants reported their age, gender, and ethnicity to characterise the sample. All participants were selected to be aged 18-24-years-old and in the UK at the time of the study.

4.2.2.2. COVID-19

Questions relating to the COVID-19 pandemic were included in this study. These questions are not included in the analyses, but instead are used to characterise the sample. Participants were asked to report their employment status before COVID-19 using a tick-box checklist with the options "full- time employed", "part-time employed", "unemployed", "self-employed", "home-maker", "student", "other (please specify)". They were then asked what their working status was for the majority of lockdown with the options: "furloughed", "key worker", "sick leave", "working from home", "other (please specify)". The same question was repeated but asked participants their current working status, with the same options as "for the majority of lockdown" but with the additional option of "returning to work". Three questions were asked on a slider (visual analogue scale): "How stressful have

you found the lockdown measures?" (Not very stressful – Extremely stressful), "How much has the COVID-19 pandemic affected your life?" (Not at all – Extremely), and "Have the effects of the pandemic been positive or negative for you?" (Completely negative – Completely positive).

4.2.2.3. Affect

The 20-item Positive and Negative Affect Scale (PANAS; Crawford & Henry, 2004) was used to measure participants' current mood ("How do you feel right now?"). The PANAS includes 10 positive items (e.g., "interested", "happy") and ten negative items (e.g., "guilty", "sad") for different mood states, which participants rate on a 5-point Likert scale ("Very slightly or not at all" to "Extremely"). Responses produce scores between 10 (low) and 50 (high) for positive and negative affect separately. The PANAS has previously been determined to be a valid and reliable measure (Crawford & Henry, 2004). One attention check was present in this questionnaire, which stated: "Select "A little" (This is an attention check)".

4.2.2.4. Sleepiness

The Stanford Sleepiness Scale (Hoddes et al., 1972) was used. As described in Section 3.2.2.5., higher scores indicated greater sleepiness.

4.2.2.5. Sleep

The Pittsburgh Sleep Quality Index (PSQI; Buysse et al., 1989) was used to measure participants' habitual sleep (described in Section 4.2.2.5.1. below). An amended version was used to measure sleep the night prior to the study (described in Section 4.2.2.5.2.). The PSQI is a self-report measure of sleep quality and duration. Two attention checks were included in

the sleep section. The attention check within the habitual sleep questions stated "Select "Three or more times a week" (This is an attention check)". The attention check within the last night's sleep questions asked participants to "please type "last night" – this is an attention check".

4.2.2.5.1. Habitual sleep

The PSQI was used to measure habitual sleep during the past month (for this study, that month was during lockdown restrictions). All 17 items used to calculate the global PSQI score were included in the habitual sleep questions; bed-fellow ratings were removed as they are not needed to calculate the global PSQI score and would not provide information relevant to the current research questions. One item ("how often have you had trouble staying awake while driving, eating meals, or engaging in social activity?") was changed from "driving" to "studying" due to the reduction in driving during the pandemic and the age range of the participants. These amendments are in line with previous research making the PSQI suitable for use with adolescents (de la Vega et al., 2015). Items were scored according to the original PSQI guidance, using seven components (subjective sleep quality, sleep latency, sleep duration, habitual sleep efficiency, sleep disturbances, sleep medication, and daytime dysfunction). Each component was given a score of 0-3 with low scores indicating good sleep and high scores indicating poor sleep. The component scores were then summed to generate a global PSQI score, with possible scores ranging from 0 (good sleep) to 21 (poor sleep).

An additional habitual sleep question was created. Participants were asked to respond on a visual analogue scale as to how regular their sleep patterns were in the past month, from "extremely regular" to "extremely irregular" as a measure of intraindividual night-to-night sleep variability.

4.2.2.5.2. Last night's sleep

The PSQI was amended to create a shorter, but comparable, questionnaire to measure participants' sleep the night prior to participating in the study (see Appendix 4.1.). The previous night's sleep questionnaire included the same wording as the original PSQI except "in the past month", which was changed to "last night" and some tense changes. Items from component one (subjective sleep quality), item two from component two (sleep latency), component three (sleep duration), and component four (sleep efficiency) were included in this questionnaire. For component two, question 5a referred to how often a participant "cannot get to sleep within 30 minutes", which is redundant for measuring last night's sleep when question two (the other item from component two, which was included in this questionnaire) asks "how long did it take you to fall asleep?". Components five (sleep disturbances), six (sleep medication), and seven (daytime dysfunction), which all ask for responses regarding frequency of experiencing the items within the past month, were not included within this questionnaire. The included questions produce a range of possible scores from 0 (good sleep) to 12 (poor sleep) for last night's sleep.

One additional question asked participants whether they had taken any naps during the day of the experiment ("no naps", "Less than 1 hour", "1-2 hours", "2-3 hours", "3-4 hours", "4+ hours").

4.2.2.6. Chronotype

The five-item reduced Morningness-Eveningness Questionnaire (MEQr; Adan & Admirall, 1991) was used in the present study. The MEQr can be used to quickly assess if someone is more of a morning-type (e.g., "early bird") or evening-type (e.g., "night owl"). Chronotype can influence sleeping patterns, mood, and timings of cognitive peaks (Zerbini & Merrow, 2017). The MEQr includes questions such as "during the first half hour after having

woken in the morning, how tired do you feel?". Items one, three, and four are scored on a five-point Likert scale from 1 – 5. Item two is scored 1 – 4 and item 5 is scored 0, 2, 4, or 6. These responses produce a total possible range of scores from 4 – 25. Higher scores indicate a greater inclination towards morning-type, and lower scores indicate a greater evening-type preference.

4.2.2.7. Word-pairs

Three lists of 10 emotionally valanced, semantically unrelated word-pairs were created for this study (see Appendix 4.2.). One list was positively valanced, one negatively valanced, and one neutral, based on the Affective Norms for English Words (ANEW; Bradley & Lang, 1999).

The word lists were matched on numerous variables, the means and standard deviations for the matched variables for the different lists can be found in Table 4.1. below. Pairwise semantic relatedness was measured by the cosine semantic distance between the word-pairs based on a large corpus (http://meshugga.ugent.be/snaut-english/), which included both information from film and TV programme subtitles and information from another corpus based on web-crawling (Ferraresi et al., 2008, as cited by Mandera et al., 2017). Pairwise semantic relatedness was not significantly different, $F(2, 27) = 2.014$, $p = .149$, between the negative, positive, and neutral lists. None of the words in a single word-pair list (positive, negative or neutral lists) were associated with each other based on free association data (Nelson et al., 2004). Information regarding frequency, imageability, familiarity, number of letters, phonemes and syllables were all gained from N-Watch (Davis, 2005). As some words are more frequently written or spoken, and we use written presentation of stimuli, the lists were matched on both total (spoken + written) and written only frequency. The total CELEX frequency did not differ, $F(2, 57) = 0.002$, $p = .998$, nor did written CELEX

frequency, $F(2, 57) = 1.146$, $p = .325$. Imageability was matched, $F(2, 57) = 0.684$, $p = .509$, between negative, positive, and neutral lists, as was familiarity, $F(2,57) = 1.146$, $p = .325$. The number of letters, $F(2,57) = 0.267$, $p = 0.767$, phonemes, $F(2,57) = 0.358$, $p = .701$, and syllables, $F(2,57) = 0.685$, $p = .508$, also did not significantly differ between the lists. Based on ANEW data, all the lists had significantly different emotional valence from each other, $F(2,57) = 1524$, $p < .001$. Emotional arousal was significantly different between the lists $F(2,57) = 36.45$, $p < .001$, with positive and negative lists matched on arousal ($p = .938$) and both significantly more arousing that the neutral lists ($p < .001$). See supplementary material on Open Science Framework for data on individual words used in the lists in regard to matched variables.

Table 4.1. Descriptive statistics for the emotionally valanced word-pair lists matched on different variables

Matched variable	Negative		Positive		Neutral	
	Mean	SD	Mean	SD	Mean	SD
Valence	2.12	0.35	7.89	0.34	5.21	0.30
Arousal	5.84	0.87	5.75	0.97	3.93	0.47
CELEX_total	56.3	82.3	55.1	51.3	56.2	62.6
CELEX_written	58.6	84.4	56.6	52.0	56.2	62.6
Imageability	460	172	485	104	438	155
Familiarity	521	55.6	545	44.1	529	52.5
Semantic relatedness	0.86	0.05	0.83	0.47	0.89	0.08
Letters	5.7	1.72	5.35	1.60	5.5	1.19
Phonemes	4.9	1.71	4.8	1.44	4.5	1.50
Syllables	1.8	0.70	1.7	0.80	1.55	0.51

Note: Valence and arousal from ANEW are measured on a rating scale from 1-9 (Bradley & Lang, 1999). Regarding N-Watch data (Davis, 2005), CELEX frequencies refer to the frequency of occurrences per million words. Imageability and familiarity ratings from N-Watch are based on subjective ratings from 1 (very unfamiliar/unimageable) to 7 (very familiar/imageable), which are then scaled to produce a range of 100-700. Semantic relatedness scores range from 0-1, with smaller values indicating a shorter cosine semantic distance (i.e., more closely semantically related) and higher values demonstrating a larger semantic distance (i.e., more semantically unrelated; Mandera et al., 2017).

4.2.2.8. Mental health and wellbeing

The 21-item (reduced) version of the Depression Anxiety and Stress Scale (DASS-21) is a freely available measure of psychological wellbeing, which separates depression, anxiety and stress based on tripartite theory and has previously been examined in combination with the PANAS (Henry & Crawford, 2005). Participants report how much an item applied to

them in "the past week" on a 4-point Likert scale from (0) not at all to (3) most of the time (Gomez, 2016). There are seven items for each subscale (depression, anxiety, and stress) meaning that possible scores range from 0 – 21 per subscale. One attention check was included in this questionnaire asking participants to "Please select "Most of the time" (this is an attention check)". The data from DASS-21 are not included in the analyses of this study, but instead were used to characterise the sample.

4.2.3. Procedure

The tasks and questionnaires for this study were programmed on Gorilla Experiment Builder (Anwyl-Irvine et al., 2019). Participants were recruited via Prolific (https://www.prolific.co/) and completed the study online. The study was restricted to computers only, excluding mobile phones or tablets, to prevent any issues with screen size or touchscreens. Participants were informed that the study would take approximately 30 minutes to complete. They were asked to complete the study in one sitting at a time where they had no distractions. Participants were also asked not to use external memory aids (e.g., writing or photographing the word-pairs). At the end of the study, participants were given the opportunity to provide feedback, self-report the use of external memory aids (no consequences for participant payment), and report what memory aids they used, before being debriefed and exiting the survey. All blocks were self-paced, apart from the presentation blocks for the word-pair encoding task, which was timed as described below.

The study was piloted prior to data collection on Prolific via Gorilla, with 20 participants, using the same inclusion criteria (e.g., age, UK resident) used for the main study. Piloting determined that 30 word-pairs repeated with two blocks of presentation and recall was appropriate for avoiding ceiling effects on the third repetition and reducing the

total study completion time, allowing time for the addition of the DASS-21 questionnaire in the final study.

4.2.2.9. Questionnaires

All questionnaires were presented in the order they are described in the materials section. The items within the PANAS were presented in a randomised order for each participant. All other questionnaires were presented in the fixed order from their original design. Questionnaires were completed at participants own pace and some questionnaires included attention check items as described in the materials section.

4.2.2.10. Word-pairs

Two types of trials were used for the word-pair task, participants first encoded the 30 word-pairs (encoding trial), and then recalled the 30 word-pairs (cued recall trial). These two trial types were repeated over two blocks, meaning that participants encoded and recalled each word-pair twice. Before encoding, participants were instructed to memorise the words as best they could and to imagine the word-pairs in an image together to support memory. Encoding trials began with a fixation cross presented for 500 milliseconds in the middle of a blank screen. For each trial a word-pair was presented with one word above where the fixation cross was presented (cue word) and the other word below the fixation cross position (target word) on the screen for 3000 milliseconds. The top word was later used as a cue during the cued recall trials and participants were asked to type in the bottom word. The recall phase was not time restricted, allowing participants to take as long as they would like to recall the bottom word. Word-pairs were presented in a random order during encoding and recall.

Twelve attention checks were randomly presented throughout encoding sessions, six per block (three left and three right in each block). A red dot appeared to the left or right of

the screen for 500 milliseconds, participants then made a button press judgement as to which side the dot was positioned. Word-pair responses were scored as correctly recalled or incorrectly recalled. Misspelled or mis-typed responses were scored as correct. The display and recall of word-pairs was repeated twice, and therefore, participants are given separate scores for the two trials.

4.3. Analysis plan

4.3.1. Justification of analysis plan

A logistic generalised linear mixed effects model (logistic GLMM) was used to analyse the data. GLMMs are able to model participant and item-level random variance at the same time, which is not possible in other statistical analyses such as ANOVAs (Brown, 2020). Modelling this variance can account for dependence in the data. For example, in the word-pair memory task, each participant completes multiple-trials. Therefore, responses cannot be assumed to be independent (an assumption in ANOVAs) due to individual differences in memory abilities (Meteyard & Davies, 2020). As the current study was conducted online, within a 30-minute window, it was not possible to include multiple individual difference background measures, such as general memory tests. Instead, GLMMs can statistically account for individual differences via participant-level random effects modelling. As the word-pairs used in the memory task are novel to the current study, it is also important to identify any ceiling or floor effects for individual items, which can be examined via the correlations between random-effects (Brown, 2020). For example, if by-participant slopes and intercepts are negatively corelated, it could demonstrate that the word-pair task was too easy (ceiling effects), in essence, memory performance was so high in general that the influence of valence is not observed. Traditional ANOVA or regression statistical approaches are also limited in regard to data types. ANOVAs require categorical predictors

and continuous outcomes; standard regression models typically use continuous predictors and continuous outcomes, with the exception of binary variables in logistic regression. However, this study uses a combination of categorical (e.g., valence: positive, negative, and neutral) and continuous predictors, with a binary outcome (correct or incorrect recall). Whilst a continuous outcome could be derived (e.g., percentage accuracy) this would require at least three separate outcomes for positive, negative, and neutral word-pairs. GLMMs can include emotional valence as a categorical predictor, and correct and incorrect recall as the outcome. This approach is more parsimonious, and reduces the inflation of familywise error and issues associated with multiple comparisons in separate analyses. A logistic GLMM using a binary accuracy outcome is also a more sensitive approach, which does not reduce the data to means and can include the recall of each word-pair over two-trials without losing any data.

4.3.2. Sample size and design

This study used a crossed design, in which all participants were exposed to all items (i.e., positive, negative, and neutral word-pairs). Data were collected for 255 participants, the final sample after data screening and cleaning is reported in the results section. Data were collected within a time-limit, due to fast changing COVID-19 guidelines in the United Kingdom. Therefore, no a-priori power analyses were conducted. Additionally, there is currently no agreed upon method for calculating power within linear mixed effects models (Meteyard & Davies, 2020).

4.3.3. Data screening and cleaning

According to best practice recommendations in conducting and reporting linear mixed effects models, the current study explains all data cleaning procedures, including justifying data removal and any transformations of the data (Meteyard & Davies, 2020).

4.3.3.1. Data removal

As GLMMs can handle missing data well, participants were not excluded listwise due to missing data as they would have been in ANOVAs. However, participants that had a total recall accuracy rate of zero were excluded, as they had no relevant outcome measures. The suspected number of outliers was estimated from visualising the data using box-plots; outliers were then statistically analysed using Rosner's test, which can detect multiple outliers (Soetewey, 2020). If outliers were detected, they would be removed one-by-one. If the removal of an individual outlier was found to change the slopes, demonstrating high leverage, then the outlier would be permanently removed. We would also compare the final models with and without outliers to examine if inclusion or exclusion of outliers significantly altered the model predictions or significance. Any removed participants are reported in the results section. The total number of data points and the final sample size after removal of data is also reported.

4.3.3.2. Scoring

The questionnaire data collected in the current study used means, or aggregated scores, as intended within the original questionnaire designs. These aggregated scores were treated as continuous data, provided that the distribution of scores was adequate. The Pittsburgh Sleep Quality index for habitual sleep was used to create a global score with a possible range of 0 (good sleep) to 21 (poor sleep) based on summing the seven component scores (Buysse et al., 1989). The amended PSQI used to measure last night's sleep was similarly summed by adding the component scores to produce a potential range of 0 – 12. The Positive and Negative Affect Scale produces scores of 10 (low) to 50 (high) for positive and negative affect separately (Watson et al., 1998).

The mental health (DASS-21) and COVID-19 questionnaires are not included in the current analyses, except to characterise the sample. The control variables included within this analysis were: chronotype, age, sleep variability, daytime napping, and sleepiness. Chronotype, as measured by the morningness-eveningness questionnaire, has previously been validated as being on a continuum (Natale & Cicogna, 2002) and thus, it is appropriate to treat it as a continuous variable. The reduced morningness-eveningness questionnaire used in the current study produces a range of scores from 4 – 25, with lower scores indicating a higher inclination towards evening-type preferences (Adan & Admirall, 1991). Sleep variability was measured on a single item visual analogue scale and can be used as a continuous variable. Age was also measured on a single item (18-24-years-old) that can also be used as a continuous variable. Napping and sleepiness were measured on single-item Likert scales.

The outcome variable, memory accuracy for word-pairs, was used as a binary outcome. Each trial was scored as (1) correct or (0) incorrect. GLMMs can handle binary outcomes such as accuracy per item, without the necessity for aggregated scores such as mean accuracy, which would not account for item-level variance. Due to the use of a binary outcome, normal distribution cannot be assumed, hence the use of *generalised* linear mixed effect modelling, specifically, a logistic regression model (Brown, 2020). Instead, the binary outcome is assumed to be distributed binomially, and by default, R uses the logit link to transform data in the binomial distribution family (Brown, 2020).

4.3.4. Plan for analysis

The main effect of valence (positive, negative, or neutral) on memory accuracy for the word-pairs was investigated to test whether the manipulation of emotional valence categories was sufficient. If the manipulation check is not significant, strong conclusions about the

influence of emotionality of the word-pairs cannot be drawn. If there is a negative correlation between by-participant intercepts and slopes, this may indicate ceiling effects, suggesting that the word-pair memory task was too easy, and therefore may mask the influence of valence (Brown, 2020).

The analysis took a nested approach. An initial "empty" model was created with only random effects (intercepts and justified slopes). This was compared to the "control model", which was an identical model, with the addition of the control variables. If the control model was a better fit than the empty model, then the control model would be compared to the predictor models. The parsimonious approach, to prevent over-loading the model and potential convergence issues, is to remove the control variables if they are not a better fit than the empty model. The removal of the control model assumes that if the control variables do not improve fit from the random effects model, then they are unlikely to have a strong influence on the dependent variable. A "predictor model" was then created, which was identical to the best fit model (either the empty model, or the control model), with the addition of the predictor variables fixed effects, to examine if the predictor variables explain variance above the control variables and/or random effects alone. Models were compared using Likelihood Ratio Tests (LRT).

4.3.4.1. Random effects

First an "empty" model with only the random effects was created. The maximal justified random effects model for the design was included (Barr et al., 2013). Participant ID and word-pair item were included as random effects. Their intercepts were modelled, as participants may randomly differ in their general memory performance (e.g., higher/lower than the average participant), and word-pairs may differ in their memorability (e.g., more/less memorable than the average word-pair). The inclusion of by-participant random slopes may

be justified, whilst by-item slopes may not (e.g., see Brown, 2020). As the current study uses a crossed design, all participants saw all items (i.e., all word-pairs) across all emotional valences (positive, negative, and neutral). Each participant may differ in their perception of the emotional valence of the word-pair items. Therefore, it makes intuitive sense to include by-participant random slopes for the word-pairs task. Whilst the word-pairs were selected to be either positive, negative, or neutral emotionally valenced, based on ANEW data (Bradley & Lang, 1999), the perception of different items by different individuals is likely to still vary, especially under pandemic circumstances, in which words such as "ambulance", or "infection" may be more salient. However, by-item random slopes are not justified within the current design, as each item is *either* positively, negatively, or neutrally emotionally valenced, and does not occur in more than one of the manipulated valence categories (e.g., the word-pair "fat-ambulance" is negatively valenced, not positive or neutral and therefore will not be influenced by the valence category). These random effects were modelled with the control variables in the initial model and again in the fixed effects model.

4.3.4.2. Fixed effects

Three categories of predictors were used in this analysis, (1) sleep (habitual and last night's sleep), (2) affect (positive affect and negative affect), and (3) emotional valence of the word-pairs (positive, negative, or neutral). When interpreting the effects of valence, negative valence was used as a reference, due to the prediction that memory for negative valence will be equally high between participants, regardless of predictor variables such as sleep. Relationships between the predictors were examined as follows (a) sleep as a predictor of total word-pair recollection (regardless of emotional valence), (b) the interaction between sleep and emotional valence in predicting the recollection of word-pairs, (c) a three-way

interaction between sleep, affect, and emotional valence in predicting memory for the word-pairs. Each of these predictors were included as fixed effects within the model.

The hypothesised models were identical to the empty model (or the control model if it was a significantly better-fit than the empty model) with the addition of fixed effects for the predictor variables. (H1) Global PSQI scores for last night's sleep and habitual sleep were used as predictors of word-pair recall accuracy. (H2) Sleep variables used in H1 and emotional valence (positive, negative, and neutral) of the word-pairs were used to predict memory accuracy. Interactions between habitual sleep and valence, and last night's sleep and valence were examined. (H3a) Positive affect and negative affect scores, and emotional valence of the word-pairs were included as predictors of recall accuracy. Interactions between positive affect and valence, and negative affect and valence were also added. (H3b) Sleep variables in H1, positive affect, negative affect, and emotional valence of the word-pairs were entered as predictors. Interactions between habitual sleep, positive affect, and valence; and habitual sleep, negative affect, and valence were examined. Table 4.4. in the results section describes the fixed effects included in each model.

All predictors were entered together and compared to the control (described below) or empty model (described in Section 4.3.4.1.) for better fit. If the models were a better fit, predictors were then removed individually to investigate the significance of the separate predictors in the model. If removal of a predictor variable reduced model fit, this is evidence that the predictor significantly adds value to the model. If it is a significant predictor, it was kept in the final model, if the predictor was not significant it may be excluded from the final model. We reported the removal of any predictor variables from the final model.

However, first, control variables were entered into an initial model without any predictor variables. If the control model was a better fit than the random effects alone, then the control model would be compared to the predictor model, to investigate the unique

contribution of the predictor variables above the control variables. Chronotype has been associated with affect, sleep restriction and cognitive performance (Díaz-Morales & Escribano, 2014). Whilst participants were all aged 18-24 years old, age was included as a crude proxy for pre-frontal cortex development, which can continue up until 24-years-of-age (e.g., Arain et al., 2013), as pre-frontal cortex development has theoretical relevance for the relationship between sleep, emotional valence and memory (Goldstein & Walker, 2014). Additional sleep variables were controlled for such as sleepiness ratings (Dewald et al., 2010) and sleep variability (Phillips et al., 2017), which have been found to predict poorer academic performance, and therefore may affect the relationship between sleep and memory recall; and daytime napping, which could compensate for negative impacts on cognition caused by prior sleep restriction (Cousins, van Rijn, et al., 2019).

4.3.4.3. Handling convergence issues

If convergence issues arose, the first option would be to check for misspecification of the model. The predictor variables may also need to be centred or scaled to handle potential issues with collinearity or differing measurement scales (e.g., Brauer & Curtin, 2018; Meteyard & Davies, 2020). Control parameters may also be added, which would allow for editing the estimations; the control parameter that would be amended first is the optimizer, that is, the method the model used to identify the optimal solution, in line with suggestions made by Brown (2020). If any control parameters were used, this would be reported in the final results. Additionally, the maximum number of iterations may have been increased if the convergence warnings indicated this (Barr et al., 2013; Brauer & Curtin, 2018). The checklist, provided by Brauer and Curtin (2018), could also be used to further remedy convergence issues.

Continuous variables may need to be dichotomised. Skew and kurtosis were examined, and if a variable had a particularly narrow distribution, a dichotomous variable could be created using a median split. For example, a dichotomous variable for sleep could separate participants into "good sleepers" and "poor sleepers". It is best to avoid dichotomising continuous variables where possible (Meteyard & Davies, 2020). Therefore dichotomisation would only have been considered after the options above for resolving convergence issues were exhausted. At a last resort, the random effects portion of the model may need to be simplified if none of the above amendments solved the convergence issues. Every amendment in simplifying the model is explained in terms of rationale and which model equations failed to converge (Meteyard & Davies, 2020)

4.4. Results

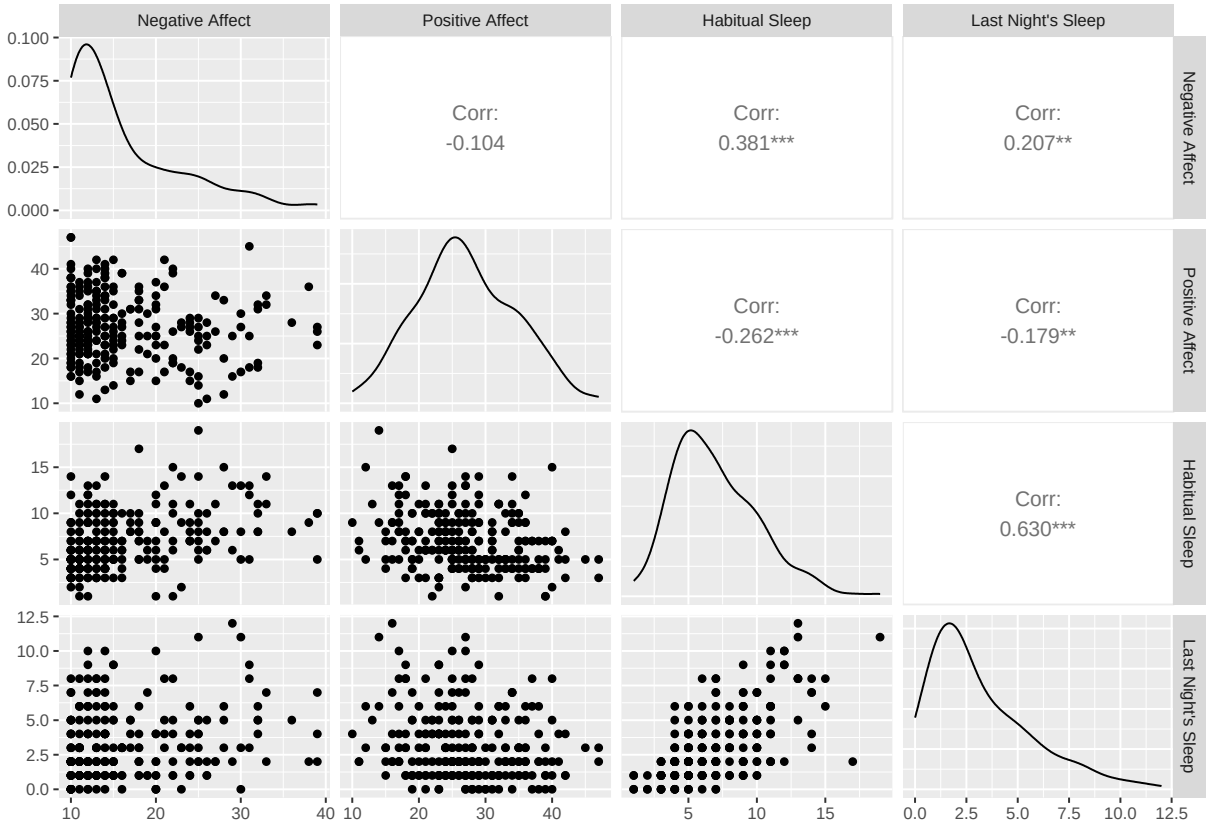
All data were analysed using R Studio. Data cleaning and analysis scripts can be found on the open-science framework, along with the raw data sets (https://osf.io/jvgzb/?view_only=e3beef485b674c6b97aaad610807d097). They are freely available to all researchers who wish to validate the results of the present study, or to run their own analyses with the data. All fixed effects included in the models are outlined in Table 4.4.

4.4.1. Assumptions

Figure 4.1. below shows the distribution of the questionnaire data. The positive affect subscale from the PANAS was normally distributed. Negative affect was skewed towards lower negative affect. Habitual sleep was roughly normally distributed with a tendency toward good sleep. The distribution for the amended version of the PSQI to measure last night's sleep was skewed toward good sleep. Convergence issues arose (see Section 4.4.5.)

which recommended rescaling the data. Therefore, the skewed data (negative affect, last night's sleep) were z-transformed and centred.

Figure 4.1. Distribution of data for predictor variables



4.4.2. Data cleaning

Due to low performance, the first block was removed for all participants and only the second block (trials 31-60 were included as the outcome variable). The first block can therefore be considered a familiarisation task and further training. Data were collected with 255 participants. Four participants reported that they had used external memory aids and were therefore removed from the analyses. A further two participants scored zero on the recall of all word-pair task and were subsequently removed. Two participants had data that could not be sufficiently cleaned for analyses (e.g., due to unclear reporting or improbable scores for sleep), and so were removed from the analyses. No statistical outliers were

identified using Rosner's test. The total remaining sample comprised 247 participants with 7410 observations (30 trials).

4.4.3. Sample characteristics

The mean age of the sample was 21.08 years (*SD* = 2.02, range = 18 – 24 years). Just over half of participants were female (55.87%, *N* = 138), one participant reported a non-binary gender. The DASS-21 was used to characterise the sample, the maximum score per subscale is 28. The mean response for the depression subscale was 14.17 (*SD* = 5.05, range 7-28), 11.57 for anxiety (*SD* = 3.82, range 7 - 23), and 14.07 for stress (*SD* = 4.29, range = 7 - 25). Ethnicity and employment status are reported in Table 4.2. and Table 4.3. respectively below.

Table 4.2. Ethnicity of the final sample

Ethnicity group	Number of participants (*N* =)	Percentage of the sample (%)
White British	189	76.52
Indian	10	3.66
Pakistani	7	2.83
African	7	2.83
White and Black Caribbean	6	2.43
White and Asian	5	2.02
Chinese	4	1.63
Caribbean	4	1.62
Bangladeshi	3	1.12
Arab	1	0.40
White and Black African	1	0.40
Other	10	4.05

Table 4.3. Employment status of the final sample (N = number of participants, % = percentage of the final sample)

	Current working status		Employment status for the majority of the pandemic	
	N =	%	N =	%
Student	85	34.41	99	40.08
Working from home	55	22.27	50	20.24
Not working	47	19.03	42	17.00
Furloughed	29	11.74	40	16.19
Working out of the home	17	6.88	14	5.67
Returning to work	8	3.24	NA	NA
Sick leave	1	0.40	1	0.40
Other	5	2.02	1	0.40

4.4.4. Reliability

In this study, the positive affect subscale of the Positive and Negative Affect Scale had good reliability (Cronbach's alpha = .89, [95% *CI* = .86 – .91]), and the negative affect subscale had excellent reliability (Cronbach's alpha = .92, [95% *CI* = .90 – .93]). The Pittsburgh Sleep Quality Index (PSQI) used to measure habitual sleep had good reliability (Cronbach's alpha = .72 [95% *CI* = .67 – .78]), and the amended version of the PSQI used for sleep the night prior to the study also had good reliability with the global score included (Cronbach's alpha = .78 [95% CI = .74 – .82]), but this reduced if the global score was removed (Cronbach's alpha = .68).

4.4.5. Convergence issues

Convergence issues arose when testing the model for Hypothesis 3b. The convergence error message suggested rescaling variables. Due to having non-normal distributions (see Figure 4.1. above), last night's sleep and negative affect were subsequently z-transformed and scaled. A second convergence issue arose, which suggested the model failed to convergence within the maximum number of iterations. These convergence issues were dealt with by adding a control parameter that increased the number of iterations "control=glmerControl(optimizer="bobyqa", optCtrl=list(maxfun=100000))". This line of code was added to every model to ensure all models were equal and allow fair comparison.

Table 4.4. Fixed effects included in the models

Model	Fixed effects
Empty	None
Control	Chronotype, age, sleepiness, daytime napping, sleep variability
Manipulation	Emotional valence
H1	Last night's sleep, habitual sleep
H2	Last night's sleep, habitual sleep, emotional valence *Last night's sleep * emotional valence* *Habitual sleep * emotional valence*
H3a	(1) Positive affect, emotional valence *Positive affect * emotional valence* (2) Negative affect, emotional valence *Negative affect * emotional valence*
H3b	(1) Habitual sleep, negative affect, emotional valence *Negative affect * emotional valence* *Negative affect * habitual sleep* *^Habitual sleep * negative affect * emotional valence* (2) Last night's sleep, negative affect, emotional valence *Negative affect * emotional valence* *Negative affect * last night's sleep* *^Last night's sleep * negative affect * emotional valence*

Note: The random effects structure was identical in all models, random intercepts were modelled for participant ID and word-pair items. The outcome in each model is accuracy (0/1) for recalling word-pairs on the second recall block (the first recall block was excluded from analysis). H3a and H3b deviate from pre-registration as the models are sub-divided for ease of interpretation. *Fixed effects in italics indicate interaction effects that were tested in the models.* ^ Interactions in H3b marked with "^" indicate that models were compared with and without this three-way interaction.

4.4.6. Control variables

The control model did not significantly differ from the empty model including only random effects, $\chi^2(4) = 1.95$, p = .744. None of the control variables significantly predicted memory. There was no significant association between memory accuracy and chronotype, *B* = -.02, *SE* = .03, *z* = -.81, *p* = .417, age, *B* = -.04, *SE* = .05, *z* = -.83, *p* = .404, sleepiness, *B* = -.05, *SE* = .09, *z* = -.55, *p* = .583, or daytime napping, *B* = -.11, *SE* = .25, *z* = -.44, *p* = .657. Therefore, the empty model, rather than the control model, was subsequently compared to the hypothesised models.

4.4.7. Manipulation check

Emotional valence (positive, negative, or neutral) of the word-pairs was added to the empty model as the only predictor of memory. Emotional valence was not found to significantly improve model fit in comparison to the empty model, $\chi^2(2) = 1.04$, *p* = .595. There was no significant difference between neutral and positive, *B* = .15, *SE* = .40, *z* = .38, *p* = .704, negative and neutral, *B* = -.41, *SE* = .40, *z* = -1.02, *p* = .309, or negative and positive, *B* = -.25, *SE* = .40, *z* = -.64, *p* = .524, valence in predicting memory. As there was not a significant effect of valence on memory for word-pairs, this may suggest that the emotional valence manipulation of the word-pairs was not strong. Therefore, caution should be given in drawing conclusions regarding the effects of emotional valence on memory in any of the hypothesised models.

4.4.8. Hypothesis testing

In the pre-registration, the planned outcome variable was correct or incorrect recollection of individual word-pairs across the two blocks. However, given that memory performance was much lower in the first block (31.92% accuracy) compared to the second

block (61.16% accuracy), it was decided to consider the first block as additional training and practice on the task. Therefore, the analyses deviate from the pre-registration in that only the second block (recall trials 31 – 60) are included as the outcome variable. See Table 4.5. below for means, standard deviation and range of scores for the variables included in the subsequent analyses.

Table 4.5. Mean, standard deviation, and range of scores for the variables included in the analyses

Variable	Mean	SD	Range
Habitual Sleep	6.98	3.11	1-19
Last Night's Sleep	3.17	2.49	0-12
Positive Memory	6.11	2.69	0-10
Negative Memory	6.55	2.50	1-10
Neutral Memory	5.98	2.89	0-10
Negative Affect	16.23	6.80	10-39
Positive Affect	27.05	7.37	10-47

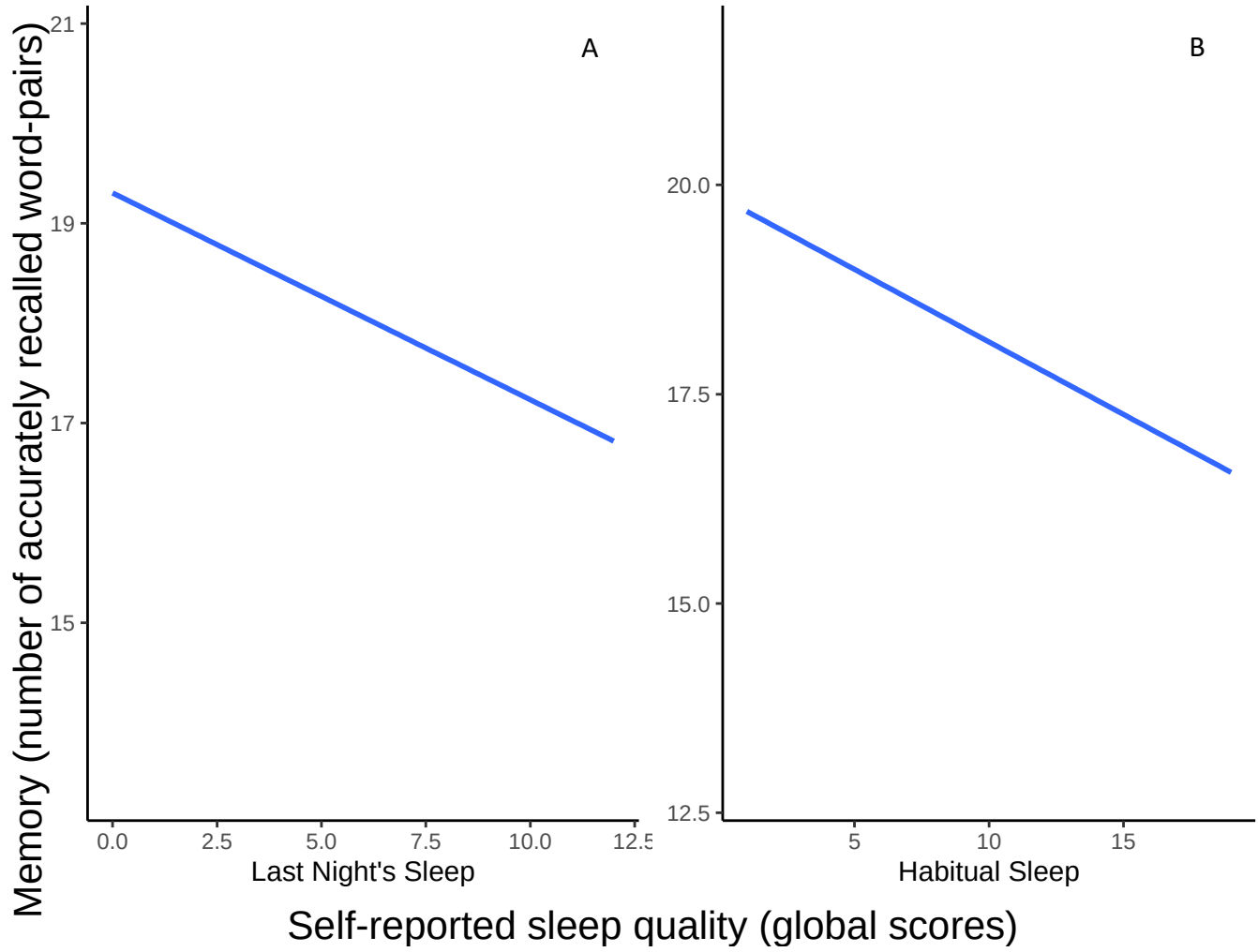
Note: Habitual sleep was measured through the PSQI (see Section 4.2.2.5.1.) and last night's sleep through an amended version of this questionnaire (see Section 4.2.2.5.2.). For both of these measures, higher scores indicating poorer sleep. Memory accuracy is calculated from the accurate recall of 10 word-pairs per affective category, with higher scores indicating better memory. Positive and negative affect were measures on the PANAS (Section 4.2.2.3.) with higher scores indicating greater experience of that affect.

4.4.8.1. Hypothesis 1: sleep and memory

The first hypothesis predicted an association between sleep (last night, habitual) and memory accuracy for the word-pairs when collapsed across valence categories. The relationships between sleep and memory can be seen in Figure 4.2. below. To test this

hypothesis a logistic mixed-effects model with recall accuracy as the outcome measure and habitual sleep and last night's sleep with fixed effects was fitted. Participant ID and word pair were included as random effects. A likelihood ratio test indicated that the inclusion of the two sleep predictors was not a better fit than the empty model, χ^2 (2) = 1.49, p = .475. Looking at the direction of each effect in turn showed that last night's sleep was not a significant predictor of total memory accuracy on the word-pairs task, B = -.05, SE = .13, z = -.36, p = .721. Habitual sleep was also not a significant predictor of memory, B = -.03, SE = .04, z = -.68, p = .494.

Figure 4.2. Associations between last night's sleep (Panel A) or habitual sleep (Panel B) and Memory accuracy

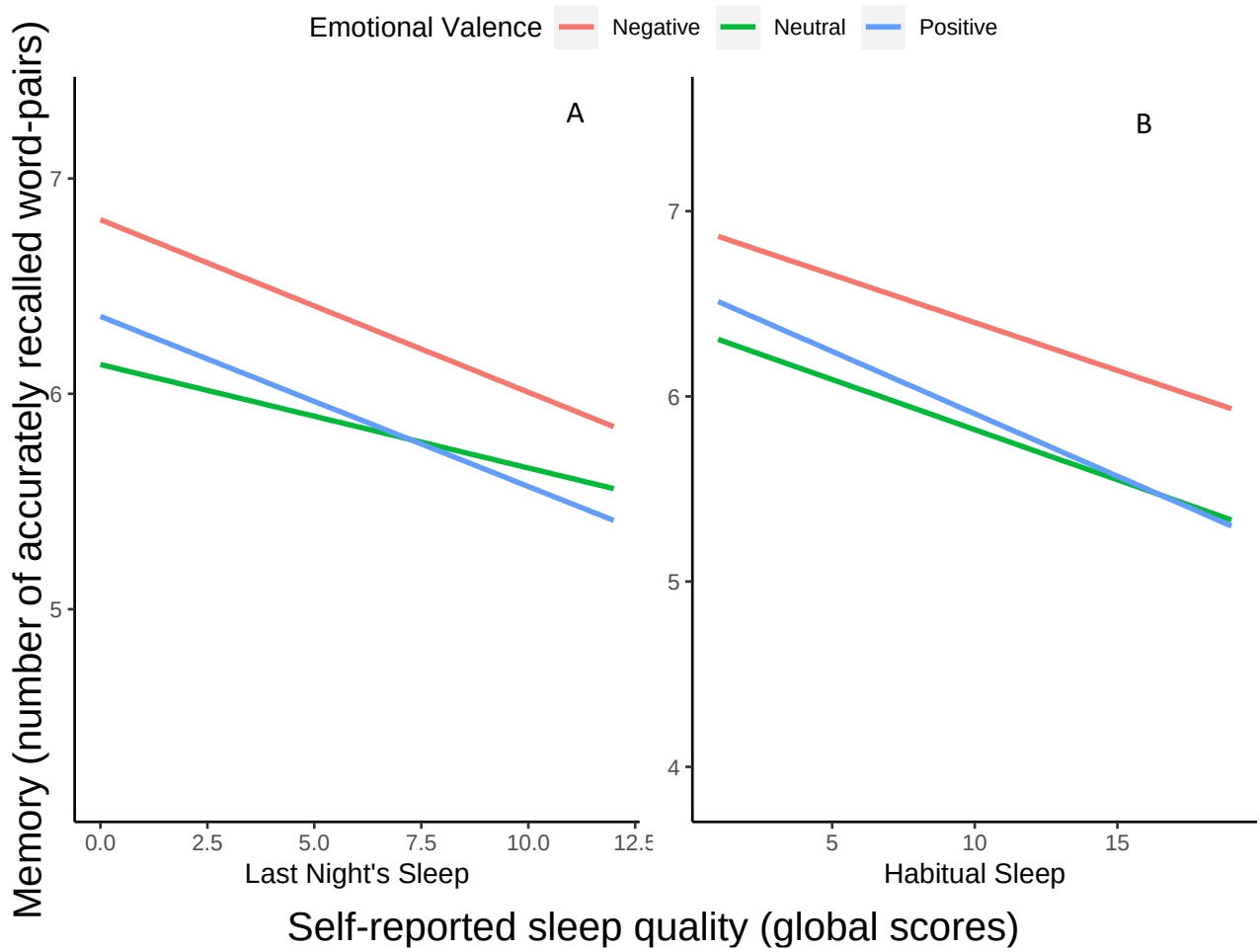


Note: Higher scores for sleep indicate poorer sleep

4.4.8.2. Hypothesis 2: sleep and emotional memories

The second hypothesis predicted that an association between sleep and memory would only be observed for positive and neutral valence categories, and that negative valenced word-pairs would be recalled equally high regardless of sleep. To test this hypothesis, a model including emotional valence as a fixed effect, and the interactions between emotional valence and each of the sleep variables separately was fitted. This was compared to the model for hypothesis 1 that included only sleep fixed effects and was not found to be a significantly better fit, χ^2 (6) = 2.39, p = .880, nor was it a significantly better fit than the empty model, χ^2 (8) = 3.88, p = .868. Habitual sleep was not significantly associated with memory for negative, B = -.16, SE = .04, z = -.35, p = .727, positive, B = -.03, SE = .04, z = -.75, p = .454, or neutral, B = -.04, SE = .04, z = -.79, p = .429, word-pairs (see Figure 4.3., Panel B below). Last night's sleep was also not significantly associated with memory for negative, B = -.09, SE = .14, z = -.66, p = .513, positive, B = -.05, SE = .14, z = -.39, p = .699, or neutral, B = .003, SE = .14, z = .02, p = .984 (see Figure 4.3., Panel A below).

Figure 4.3. Memory for emotional word-pairs as a function of sleep



Note: Higher scores for sleep indicate poorer sleep

4.4.8.3.Hypothesis 3

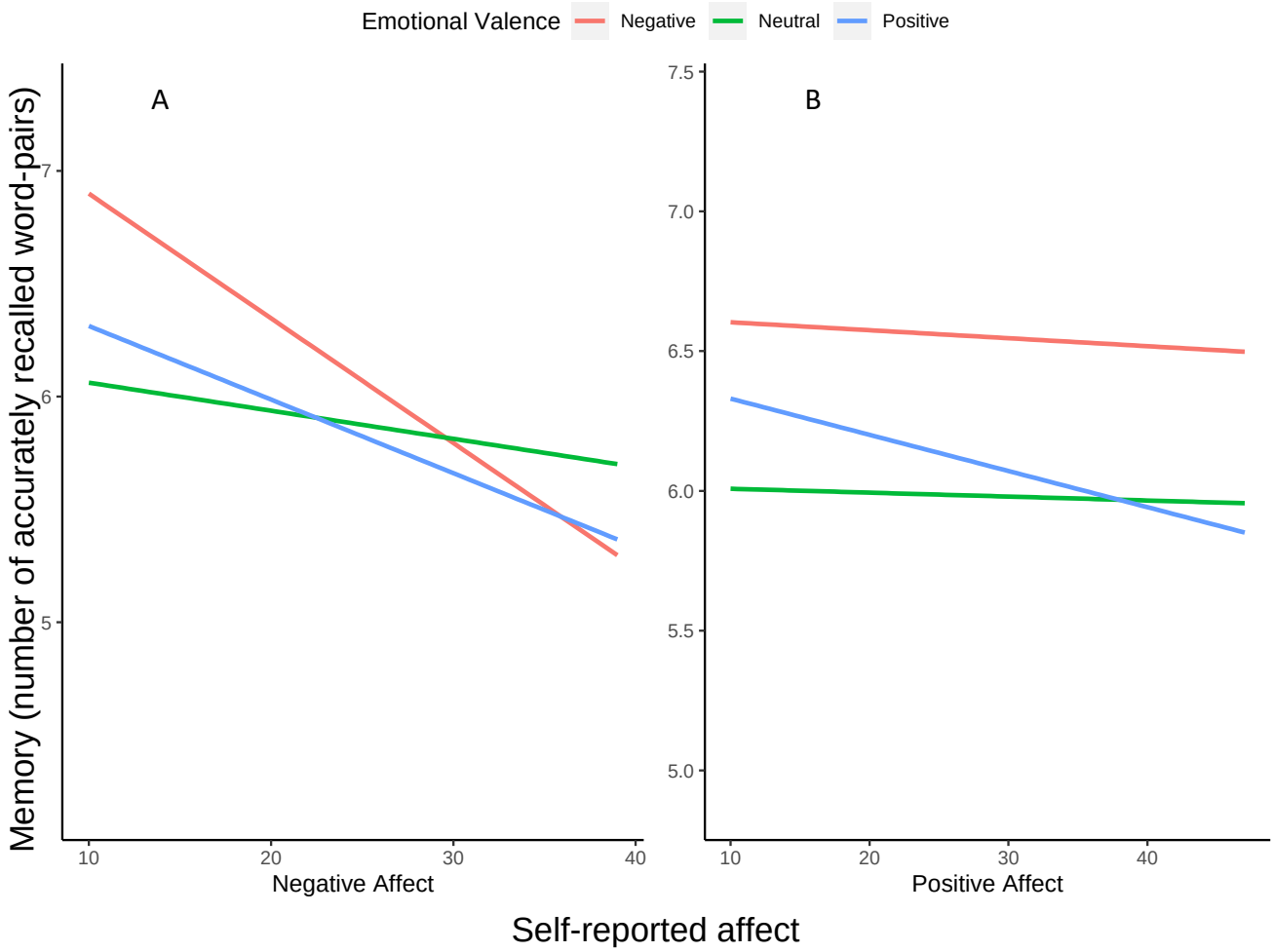
4.4.8.3.1. H3a: Affect and emotional memories

Hypothesis 3a predicted that memory for negative word-pairs is predicted by negative affect, and memory for positive word-pairs is predicted by positive affect. Originally positive and negative affect were pre-registered to be analysed in the same model, however due to difficulties in interpretation, the influences of positive and negative affect on emotional memories were examined separately as this is a more straightforward test of the hypothesis.

Figure 4.4. Panel A below shows the relationships between negative affect and emotional memory. The negative affect model was not a significantly better fit than the

empty model, χ^2 (5) = 9.57, p = .088 nor the hypothesis 2 model, χ^2 (3) = 0.00, p = 1.00. In the negative affect model, there was a significant negative relationship, in that participants with low negative affect recalled significantly more negative word-pairs, B = -.23, SE = .11, z = -2.13, p = .033 (see Figure 4.4. Panel A below). This is contrary to predictions based on mood-congruent mechanisms that would suggest a positive relationship. Negative affect did not significantly predict memory for neutral, B = -.04, SE = .11, z = -.37, p = .715, or positive, B = -.14, SE = .11, z = -1.28, p = .202, word-pairs. The slope of the interaction with negative affect significantly differed between negative and neutral valenced word-pairs, B = .19, SE = .07, z = 2.63, p = .008. There was not a significant difference between the slopes of the interaction with negative affect between negative and positive valenced word-pairs, B = .09, SE = .07, z = 1.26, p = .207, or neutral and positive valenced word-pairs, B = -.10, SE = .07, z = -1.36, p = .174.

Figure 4.4. Relationship between negative affect (Panel A) or positive affect (Panel B) and memory for emotionally valenced word-pairs



When examining the effects of positive affect on emotional memories, the model was not a significantly better fit than the empty model, χ^2 (5) = 1.83, p = .873, or the hypothesis 2 model, χ^2 (3) = 2.06, p = .561. Positive affect did not significantly predict memory for positive, B = -.01, SE = .01, z = -.64, p = .525, negative, B = -.003, SE = .01, z = -.21, p = .836, or neutral, B = -.002, SE = .01, z = -.13, p = .895, word-pairs (see Figure 4.4. Panel B above).

4.4.8.3.2. H3b: Sleep, affect, and emotional memories

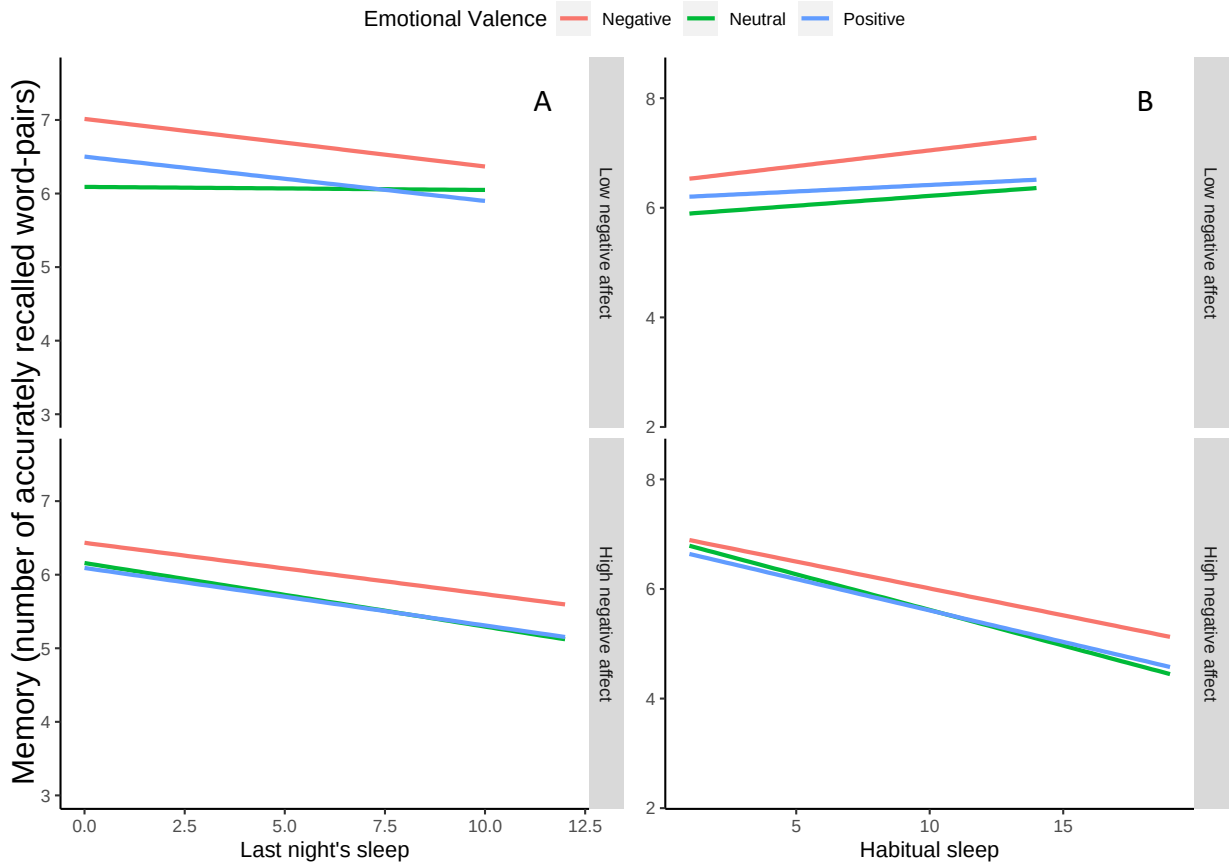
A three-way interaction was included in model 3b to examine whether sleep modulated the relationship between affect and emotional memories, as previous research has

demonstrated a relationship between sleep and affect. The analyses for model 3b deviate from the original pre-registered analysis plan for ease of interpretation. As positive affect was not a significant predictor of emotional memory biases, positive affect was not included in subsequent analyses. Model comparison methods were used to examine whether the three-way interaction significantly adds to model fit. Last night's sleep and habitual sleep were also examined in separate models.

For habitual sleep (Figure 4.5., Panel B below), there was no significant difference between the models with or without the three-way interaction, χ^2 (2) = .202, p = .904. Compared to the empty model, the model with the three-way interaction was not a significantly better fit χ^2 (9) = 13.91, p = .126. This suggests that habitual sleep did not significantly modulate the interaction between negative affect and emotional valence in predicting memory. In the model which included the three-way interaction, the interaction between negative affect and habitual sleep did not significant predict negative, B = -.06, SE = .04, z = -1.77, p = .077, neutral, B = -.07, SE = .04, z = -1.92, p = .055, or positive, B = -.06, SE = .04, z = -1.61, p = .107, memory.

For the model with last night's sleep (Figure 4.5., Panel A below), the three-way interaction was not a significantly better fit than the model without the three-way interaction, χ^2 (2) = 2.28, p = .319, nor was it a better fit than the empty model, χ^2 (9) = 12.90, p = .167. In the model with the three-way interaction, the interaction between negative affect and last night's sleep did not significantly predict negative, B = .10, SE = .10 , z = 1.02, p = .310, neutral, B = .01, SE = .10, z = .11, p = .915, or positive, B = .08, SE = .10, z = .83, p = .406, memory. See Figure 4.3. below.

Figure 4.5. Plot showing the three-way interaction between (median split) negative affect, last night's sleep (Panel A) and habitual sleep (Panel B) with emotional memory biases



Note: A median split was used to define 'high' (>14) and 'low' (</=14) negative affect for the Figure above. Higher scores for sleep indicate poorer sleep.

4.4.9. Exploratory analyses: outside of lockdown

Throughout 2020 and 2021, the United Kingdom was repeatedly under varying government restrictions and guidelines regarding the COVID-19 pandemic. The original data (Timepoint 1) above were collected online during the June 2020 lockdown, prior to the easing of lockdown restrictions on 4th July 2020. A second round of data (Timepoint 2) were collected from the same participants within a window from 23rd June to 6th July 2021, when there were no COVID-19 lockdown restrictions in the United Kingdom. Participants were

paid a higher rate (£6.45 rather than £5) at Timepoint 2 to compensate their time in both parts of the study and to encourage participants to return and complete the study.

Repeated data were collected from 128 participants (50.20% of the original sample). Within the Timepoint 2 data, one participant was identified by Rosner's test as an influential outlier for negative affect. A further four participants were identified as outliers for last night's sleep. All five outliers were removed due to the requirement for scaling these two variables. At Timepoint 2, no participant reported using external memory aids, and no participant had zero for word-pair accuracy. One person had uncleanable data for habitual sleep and another had uncleanable data for last night's sleep, these participants were removed. Therefore, the total analysed sample at Timepoint 2 was 121.

Of these 121 included participants, 118 had matched data included in Timepoint 1 (during lockdown) and Timepoint 2 (outside of lockdown) analyses. At Timepoint 1, one of the participants included in Timepoint 2 analyses reported using external memory aids, another had no outcome, and lastly one participant did not have cleanable data for last night's sleep. The data from the 118 participants with matched data are used to characterise the samples across the two timepoints. The same models were re-run as in Timepoint 1. Participants completed the same questionnaires and experimental tasks as they did during Timepoint 1. Table 4.6. below shows descriptive statistics for the sample with data at both time points.

Table 4.6. Descriptive statistics of questionnaire data and emotional memories for participants with data across both timepoints (N = 118)

	Timepoint 1 (2020)			Timepoint 2 (2021)		
Variable	Mean	SD	Range	Mean	SD	Range
Habitual Sleep	6.95	2.91	1 – 19	6.12	2.73	1 – 16
Subjective quality	1.30	0.64	0 – 3	1.12	0.63	0 – 3
Latency	1.48	0.89	0 – 3	1.26	0.86	0 – 3
Duration	0.43	0.57	0 – 2	0.52	0.58	0 – 2
Efficiency	0.57	0.87	0 – 3	0.48	0.81	0 – 3
Disturbances	1.48	0.58	0 – 3	1.33	0.52	0 – 2
Medication	0.26	0.68	0 – 3	0.14	0.49	0 – 3
Day dysfunction	1.42	0.73	0 – 3	1.26	0.72	0 – 3
Last Night's Sleep	3.24	2.39	0 – 11	2.47	1.94	0 – 9
Subjective quality	1.23	0.67	0 – 3	1.03	0.71	0 – 3
Latency	0.81	0.93	0 – 3	0.61	0.81	0 – 3
Duration	0.60	0.82	0 – 3	0.45	0.62	0 – 3
Efficiency	0.59	0.96	0 – 3	0.39	0.73	0 – 3
Positive Memory	6.08	2.61	0 – 10	6.38	2.76	0 – 10
Negative Memory	6.43	2.52	1 – 10	6.98	2.45	1 – 10
Neutral Memory	5.99	2.84	0 – 10	6.51	2.77	0 – 10
Negative Affect	16.25	6.59	10 – 39	15.47	6.68	10 – 38
Positive Affect	26.58	7.10	10 – 47	25.81	7.42	12 – 44

Note: SD = standard deviation. Habitual sleep was measured using the PSQI questionnaire (Buysse et al., 1989), last night's sleep was measured using an amended version of this questionnaire. Positive and negative affect were measured using the Positive and Negative Affect Scale (Watson et al., 1998). Memory refers to the number of correctly recalled word pairs.

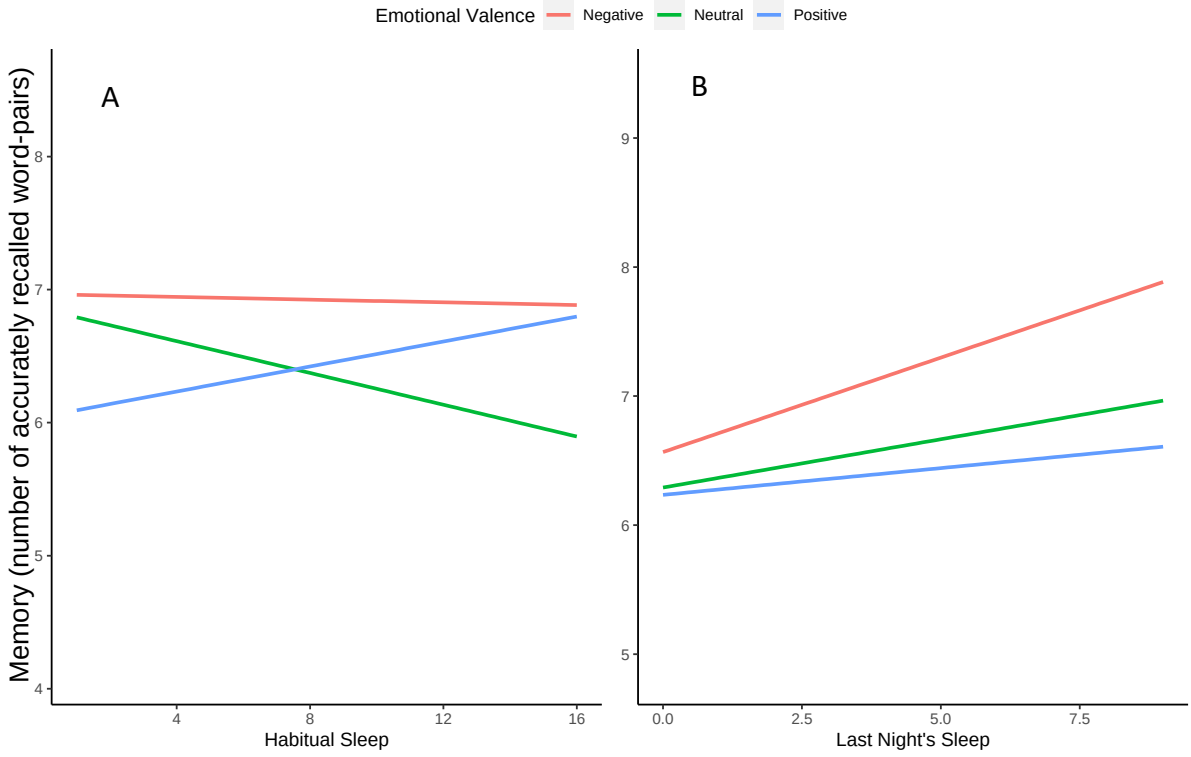
4.4.9.1. Timepoint 2 results

The control model was not a significantly better fit than the empty model, $\chi^2(4) = 3.21$, $p = .523$. The manipulation check for emotional valence was also not significant. Overall, there was not a significant difference for memory for positive compared to negative, $B = -.36$, $SE = .43$, $z = -.85$, $p = .396$, neutral compared to negative, $B = -.35$, $SE = .43$, $z = -.83$, $p = .408$, and positive compared to neutral, $B = -.01$, $SE = .43$, $z = -.02$, $p = .981$, word pairs.

In the Hypothesis 1 model, neither habitual, $B = -.03$, $SE = .06$, $z = -.52$, $p = .601$, nor last night's sleep, $B = .16$, $SE = .18$, $z = .89$, $p = .372$, were significant predictors of overall memory, as in Timepoint 1.

When re-running the model for Hypothesis 2 at Timepoint 2, habitual sleep predicted a significant difference in the slopes for neutral compared to positive memory, $B = -.10.$, $SE = .04$, $z = -2.45$, $p = .015$, but not negative compared to positive, $B = -.07$, $SE = .04$, $z = -1.76$, $p = .079$, nor negative compared to neutral memory, $B = .03$, $SE = .04$, $z = .66$, $p = .511$. Last night's sleep did not predict a significant difference in negative compared to neutral, $B = .06$, $SE = .11$, $z = .57$, $p = .568$, positive compared to neutral, $B = -.15$, $SE = .111$, $z = -1.39$, $p = .165$, nor negative compared to positive, $B = .22$, $SE = .11$, $z = 1.91$, $p = .057$, memory. Figure 4.6. below shows the relationships between habitual sleep (Panel A) and last night's sleep (Panel B) and emotional memories.

Figure 4.6. Emotional memory as a function of habitual sleep (Panel A) and last night's sleep (Panel B) at Timepoint 2 (outside of lockdown)



Note: higher scores for sleep indicate poorer quality sleep.

In the Hypothesis 3a model at Timepoint 2, negative affect did not significantly predict negative memory, B = -.12, SE = .17, z = -.67, p = .500, nor differences between negative and positive, B = .11, SE = ,11, z = 1.08, p = .279, nor neutral, B = .04, SE = .11, z = .36, p = .718, memory. Positive affect did not significantly predict positive memory, B = -.02, SE = .02, z = -1.00, p = .318, nor significant differences between positive and neutral, B = .01, SE = .01, z = .84, p = .402, nor positive and negative, B = -.01, SE = .01, z = -.69, p = .491, memory.

When re-running the Hypothesis 3b models at Timepoint 2, there was no significant difference between the habitual sleep model with or without the three-way interaction, χ^2 (2) = .59, p = .746. None of the relationships within the three-way interaction model were

significant. There was also no significant difference between the last night's sleep model with or without the three-way interaction, χ^2 (2) = 1.05, p = .592. None of the relationships within this three-way interaction model were significant either.

4.5. Discussion

4.5.1. Sleep and memory

The current study did not find significant support for the influence of self-reported habitual sleep or sleep the night before the study on memory, when collapsing across negative, positive, and neutral words, contradicting the first hypothesis. Emotional memory biases are discussed in the next section (Section 4.5.2.). The null relationship between sleep and memory does not support predictions made based on the synaptic homeostasis hypothesis (Tononi & Cirelli, 2003, 2006, 2014) that good sleep before learning improves encoding capacity. Despite the null finding, this research adds to the currently sparse literature investigating the role of sleep prior to learning. The contradiction between the results of this study and the hypothesised relationship may be due to the novel inclusion of subjective sleep quality and habitual sleep patterns in the current study, as most of the previous research investigating sleep preceding learning uses experimental sleep manipulation and focuses primarily on sleep duration. Daytime napping is a commonly used paradigm for investigating whether sleep preceding learning restores encoding capacity (e.g., Antonenko et al., 2013; Mander et al., 2013; Ong et al., 2020; Chapter 3). Both Antonenko et al. (2013) and Ong et al. (2020) used word-pairs, similarly to the current study. Therefore, it may be the sleep manipulation rather than the task explains the differences between the present results and previous research. Alternatively, our word-pair task may have been underpowered to detect perhaps more subtle relationships associated with habitual sleep quality and duration. For example, Antonenko et al. (2013) examined differences in slow-wave sleep when all

participants napped, but they used 100 word-pairs, and other declarative memory tasks such as pictorial stimuli, word-list learning, and procedural memory tasks (e.g., finger-tapping). As the current study was conducted online, only thirty word-pairs were included, in order to keep the study under 30 minutes and maintain participant attention in the absence of a researcher.

Online data collection also prevents the ability to measure neural underpinnings of the relationship between sleep and memory, and therefore cannot test hypotheses made by the synaptic homeostasis hypothesis regarding the importance of slow-wave sleep for learning and memory. In Antonenko et al.'s (2013) study, both groups of participants napped for the same duration, suggesting slow-wave sleep, not sleep duration was important. Sleep duration and subjective sleep quality measured in the current study may not be a sufficient proxy for the neural processes in the brain, which may underpin the restoration of encoding capacity. Instead, sleep architecture may be a better predictor of the restoration of encoding capacity, for example the synaptic homeostasis hypothesis emphasises the role of slow-wave sleep, whilst the nap study in Chapter 3 finds a potential role for sleep spindles. It is important to note that the synaptic homeostasis hypothesis predicts that the relationship between lack of sleep and cognitive detriments is due to reduced opportunity for slow-wave sleep to allow synaptic downscaling. However, even in the presence of sleep restriction, slow-wave sleep is found to be relatively preserved (e.g., Ong et al., 2016; Chapter 2). Therefore, even if the participants in the current study experienced restricted sleep duration, they may have still received sufficient slow-wave sleep duration. In future research, mobile technology could be used to monitor sleep architecture over multiple nights of naturalistic sleep preceding learning, for example by using Dreem headbands or ambulatory polysomnography. Additionally, future research could include measures of subjective sleep quality and sleepiness in a daily diary to support the objective sleep measures, as the combination of

objective and subjective sleep measures may be the best methodological approach (Dewald et al., 2010).

Perhaps, the results of the present study may be due to participants reporting overall good sleep, thus not demonstrating potential detrimental influences of poor sleep on memory. Good sleep may be more prevalent in this study, as the original data were collected during lockdown, when 87% of the sample (N = 216) were likely able to set their own sleep schedule; for example, due to furlough (12%), working from home (22%), studying (likely from home; 34%), or not working (19%). However, data were also collected outside of lockdown and analysed in exploratory analyses. These data demonstrate some improvements in sleep outside of lockdown (in June – July 2021) compared to during lockdown (see Table 4.5.). Other research, in three European countries, found that over a 6-week period of COVID-19 restrictions, sleep duration increased, and social jetlag decreased; however, sleep quality also decreased and was related to perceived burden and decreases in physical and mental wellbeing (Blume et al., 2020). Our data, albeit subjective scores, suggest all sleep parameters showed some improvement outside of lockdown, including sleep duration. Perhaps different results could be found in participants with higher negative affect and poorer sleep such as participants with mental health disorders. Questions relating to sleep and mental health are explored later in Chapters 5 and 6.

4.5.2. Emotional memory biases and mood congruence

It is important to note that the manipulation check did not demonstrate a significant main effect of emotional valence of the word-pairs on memory, suggesting that perhaps the emotional valence manipulation was not sufficient. Therefore, interpretation of results regarding emotional valence should be taken with caution. An alternative explanation would

be that emotional valence alone does not reliably predict memory accuracy. Suggestions for considering emotional valence in future research are made later in Section 4.5.3.

Regarding emotional memory biases, sleep deprivation research demonstrates that participants who are well-rested or sleep deprived have equal negative memory, but sleep-deprived participants have lower neutral (Vargas et al., 2019) and positive (Tempesta et al., 2016; Walker, 2009) memory than well-rested participants. Therefore, the prediction for Hypothesis 2 was that negative memory would be equally high across participants. The second hypothesis was not supported at Timepoint 1 in that sleep was not significantly associated with differential memory for positive, negative, or neutral word-pairs. At Timepoint 2 (one year later, outside of lockdown conditions), exploratory analyses demonstrated that participants reporting good habitual sleep demonstrated higher neutral and lower positive memory, whilst the opposite was found for participants with poorer sleep, in a perpendicular relationship (see Figure 4.6., Panel A). Negative memory was not significantly associated with habitual sleep, suggesting it was equal across good and poor sleepers. However, whilst the relationship not significant (p = .057), poor sleep on the night preceding the study was associated with a larger gap between negative (higher) and positive (lower) memory (see Figure 4.6., Panel B), which may be significant with more emotionally salient stimuli. The implications for exploring habitual sleep compared to last night's sleep are discussed further in the next section (Section 4.5.3).

It is useful to note that sleep and memory for the word pairs improved from Timepoint 1 to Timepoint 2 (see Table 4.6.). The increased memory at Timepoint 2 may be associated with increased training opportunities or long-term consolidation, as the same word pairs were used at Timepoints 1 and 2. As the relationships between sleep and emotional memory biases were only observed at Timepoint 2, it may suggest selective consolidation occurred, as proposed by the Affective Tagging and Consolidation model (Harrington et al.,

2017). Participants with poorer habitual sleep may have selectively consolidated the positive word-pairs from Timepoint 1 and may have recognised them to a greater extent at Timepoint 2. This could have also supported further training at Timepoint 2. Higher positive than neutral memory associated with poor habitual sleep may be explained by an over-reactivity to all emotional stimuli in the presence of poor habitual sleep, resulting in better memory for negative and positive stimuli at the expense of neutral stimuli, supporting Goldstein and Walker's (2014) suggestions. However, Goldstein and Walker's (2014) model suggests that sleep functions to prepare the brain for next-day emotional functioning. As such, these relationships should also be seen in association with last night's sleep. However, last night's sleep demonstrated a trend toward greater negative than positive memory. This suggests that last night's sleep and habitual sleep may have different relationships, perhaps associated with encoding, recall, and consolidation. Last night's sleep may influence encoding or perhaps recall given that these relationships were only seen at Timepoint 2 and not initial encoding at Timepoint 1. Whilst habitual sleep could potentially indicate influences of consolidation, hence the different relationships than last night's sleep.

At initial encoding (Timepoint 1), the relationship between affect and memory was the opposite to predictions (i.e., mood-congruent memory) made by associative network theories (Bower, 1981). When considering the influence of positive and negative affect on emotional memories, the predictions made by mood-congruent memory mechanisms were not supported. Positive affect was not a significant predictor of emotional memory biases. Whilst negative affect did predict emotional memory biases, this was in the opposite direction to the hypothesised relationship (Hypothesis 3a). Instead, high negative affect predicted significantly *lower* negative memory. This may suggest that participant with high negative affect may inhibit negative memories, potentially as an affective-regulation strategy. This relationship was only seen at Timepoint 1, suggesting that negative affect at initial

encoding and recall may be more important than after a consolidation opportunity. However, this model was not a significantly better fit than the empty model with only random effects. Perhaps this suggests there may be other individual differences in this relationship worth considering in the future. Negative affect predicted significant differences in the slopes between negative and neutral memory, in that those participants high in negative affect had fewer negative memories and more neutral memories, whilst the opposite was found for participants low in negative affect (see Figure 4.4. Panel A). Contrary to predictions negative affect did not predict significant differences between negative and positive memories. Again, this may suggest an affective-regulation strategy in that memory is not biased towards negative information at the expense of positive stimuli in the presence of high negative affect for the present sample. However, the relationships between affect and emotional memory biases may have some additional noise due to the differences between mood and emotion that were not accounted for in the present study. The implications for separating mood and emotion in future research is discussed further in Section 4.5.3.

4.5.3. Future research

Whilst the current study benefits from a large sample size (N = 118 – 246) due to the use of online recruitment and data-collection, this limited the type and number of stimuli that could be used. As such the sample size may be highly powered, whilst the design (e.g., stimuli, training repetitions) may be underpowered. A laboratory-based study would allow for more training iterations to increase encoding strength, which has demonstrated stronger relationships with sleep consolidation (e.g., Walker et al., 2019; Berres & Erdfelder, 2021) and may also have relationships with sleep prior to encoding. Given that the manipulation check for the emotional valence of the word-pairs was not significant, laboratory-based research would also allow for the use of more emotionally salient stimuli, with which, more

significant relationships may emerge. Future research could also consider measuring individual ratings of the emotional valence of the word-pairs, rather than assuming that the word-pairs will have similar emotional valence across participants. This may explain a larger proportion of variance associated with individual differences (e.g., in the random effects model). For example, neutral ratings for stimuli (e.g., in the International Affective Picture System; Schneider et al., 2016) may be masking mixed positive and negative ratings, rather than being affectively neutral. This may also extend to neutrality ratings for the valence of the words selected from the Affective Norms for English Words database (ANEW; Bradley & Lang, 1999) used in the present study. Goldstein and Walker (2014) also suggest that sleep and mental health may influence participants perception of the saliency of stimuli, for example, interpreting neutral stimuli negatively, or showing reduced perceptions of the saliency of positive stimuli. As such, different participants may have different affective responses to the stimuli, which requires further investigation in future research. Salient information outside of emotionality could also be used to further test predictions made by Goldstein and Walker (2014), and may be more easily manipulated than emotional valence, for example using rewarding stimuli or risk-based paradigms. Alternatively, other more ecologically valid stimuli could be used, such as movie clips (e.g., Tempesta et al., 2016). It could also be of interest to consider participants' confidence in their ability to remember the stimuli, which may be associated with the emotional valence of the stimuli (Zimmerman & Kelley, 2010).

In the current study, affect was used as an umbrella term encompassing both mood and emotion. However, the Positive and Negative Affect Scale (Watson et al., 1998) can also be used to examine moods and emotions separately by changing the wording of the instruction. Emotions have been found to fluctuate with sleep restriction (Baum et al., 2014), suggesting that sleep may influence emotional regulation. However, moods may be more related to

habitual sleep, in that moods are longer lasting than emotions and therefore potentially a product of cumulative disruptions to sleep. Future research distinguishing between moods and emotions may be able to further address questions regarding daily sleep and habitual sleep is predicting the relationship between affect and memory biases. The present sample also reported low negative affect and good quality sleep overall. As such, the relationships examined in the present study may differ in participants with higher negative affect and poorer quality sleep, such as participants with mental health disorders. The implication of sleep and affect for mental health are explored further in the next chapter (Chapter 5).

In regard to the current sample, data were only collected with older adolescents (18-24-years-old; Sawyer et al., 2018). It may be of interest for future research to consider the developmental impacts involved in these relationships using a wider age range. The recruitment of older adolescents was decided based on the availability of established online platforms for recruitment of participants over the age of 18-years-old (e.g., Prolific). There were not any established platforms available, to our knowledge, for online research with participants under the age of 18-years-old. Instead, online studies with younger participants typically relied on recruitment within schools, which were closed during the initial pandemic lockdown conditions, or social media, which is more time-consuming for recruitment. As recruitment was time-limited, due to regular changes in COVID-19 restrictions, data collection with 18-24-year-olds was the pragmatic option for the present study. Further, there are ethical barriers to participant payment or informed consent in the recruitment of samples under 18-years-old online, creating issues in compensating participants for their time. However, in the next chapter (Chapter 5) we established an efficient method for quickly collecting data online with younger participants and compensating them for their time. Future research may be able to use this method to recruit participants spanning the developmental

period of adolescence (10-24-years-old) to examine the relationships outlined in the present study across adolescent development.

All R scripts used for data cleaning and statistical analyses are shared on the Open Science Framework. This includes scripts used to clean the Pittsburgh Sleep Quality Index (PSQI) questionnaire, which was previously laborious to score, particularly with large samples recruited online. With these R scripts, future researchers can quickly and efficiently use the PSQI online via platforms such as Gorilla Experiment Builder and score and clean it in R using amended versions of the code provided. Additionally, the scripts to score the adapted version of the PSQI to measure sleep on the night prior to the experiment can also be found on the Open Science Framework and can be freely used and adapted in future research. The last night's sleep questionnaire (Appendix 4.1.) could be used in daily diaries and compared to habitual sleep questionnaires. Currently, there is not a well-established or widely-used self-report for measuring sleep on a nightly basis, making meta-analyses or other research syntheses difficult, due to heterogenous methodologies.

As discussed previously, a combination of objective and subjective sleep measures, and considering the role of sleep architecture (e.g., via polysomnography, Dreem headbands) may be of interest for future research investigating the relationships between sleep and emotional memory biases. Furthermore, daytime naps are more commonly used to investigate the restoration of encoding capacity (e.g., Mander et al., 2011; Antonenko et al., 2016; March et al., 2023 described in Chapter 3). However, using different sleep methodologies may show different relationships. For example, a recent meta-analysis has demonstrated greater effect sizes for sleep and episodic memory for naturalistic overnight sleep than daytime naps (Berres & Erdfelder, 2021). As such, investigating sleep quality and multiple nights of sleep may show different relationships. There may also be interactions between sleep prior to

encoding and sleep related consolidation processes after learning, which require further investigation.

4.5.4. Conclusion

The current study adds to the sparse literature considering the role of sleep prior to encoding and considers the role of positive affect and positive memory biases which are often understudied. Self-reported sleep duration and quality both habitually and on the night prior to the study were not significant predictors of memory in generally, contrary to hypotheses that better sleep would restore encoding capacity (e.g., explored in Chapter 3). However, we are unable to examine sleep architectural correlates in an online study to fully test hypotheses based on different theories (e.g., Tononi & Cirelli, 2003, 2004, 2016; Saletin & Walker, 2012). Contrary to predictions regarding mood-congruence memory biases, positive affect was not found to significantly predict positive, or any emotional memory biases. In stark contrast to mood-congruent predictions, high negative affect was significantly associated with lower negative memory at initial encoding. This may suggest an inhibitory mechanism that for affective regulation. At Timepoint 2, habitual sleep was associated with a perpendicular relationship between positive and neutral memories, in that poor sleepers had higher positive than neutral memories. This may be explained by selective consolidation or higher reactivity to all emotional stimuli in the presence of poor sleep. This is contrary to sleep deprivation research which demonstrates lower positive and neutral memory in the presence of poor sleep (e.g., Tempesta et al., 2016). Future research could examine these relationships in samples with poorer sleep and higher negative affect than in the present study (e.g., participants with mental health disorders). The role of inhibition for affective wellbeing can also be directly investigated, for example in Go/No-Go paradigms, as this may have implications for theories and interventions regarding the role of sleep for emotional regulation and mental health (explored in Chapter 5).

Chapter 5: Adolescent sleep, mental health, and mediators

5.1. Introduction

5.1.1. Adolescent development

Almost every mental health disorder is accompanied by disrupted sleep (Baglioni et al., 2016). Adolescents are particularly vulnerable to the onset of mental health disorders, with three-quarters of mental health disorders onsetting prior to the age of 24-years-old (Fusar-Poli, 2019). Changes in sleep timing, resulting in insufficient sleep duration (e.g., discussed in Chapter 2), may increase adolescent vulnerability to mental health. In this chapter, the relationship between sleep and mental health is explored in relation to two existing frameworks discussed in Section 5.1.1. (Goldstein & Walker, 2014) and Section 5.1.2. (Harrington & Cairney, 2021). Both of these frameworks emphasise a key role of the prefrontal cortex, which matures throughout adolescence, until the age of 24-years-old (a potential biological marker of the end of adolescence; Arain et al., 2013; Sawyer et al., 2018). Therefore, vulnerability to mental health disorders during adolescence may also be associated with brain maturation. Such changes in sleep and brain development may impact cognitive abilities associated with emotion regulation. For example, the maturation of the prefrontal cortex has been associated with the concurrent development of affective control abilities which contribute to emotion regulation (Schweizer et al., 2020). Disrupted emotion regulation may result in lower positive affect and higher negative affect, with potential consequences for mental health. The specific role of positive and negative affect for the development of mental health difficulties are discussed in Section 5.1.3.

5.1.2. Sleep, inhibition, and emotion regulation

Prefrontal cortex maturation during adolescence may have implications for the relationship between sleep and cognitive biases associated with mental health difficulties. Cognitive biases can be seen in many mental health disorders. For example, depression has been associated with difficulties distinguishing between neutral and negative stimuli, whilst anxiety has been associated with an over-sensitivity to threatening or fearful stimuli (e.g., Goldstein & Walker, 2014). The framework proposed by Goldstein and Walker (2014) suggests that when sleep (particularly REM sleep) is disrupted, the noradrenergic system is not sufficiently recalibrated. Specifically, they propose that in the case of insufficient (REM) sleep, the concentration of noradrenaline is not effectively regulated, resulting in greater levels of noradrenaline the next day. The dysregulation of the noradrenergic systems reduces the ability of the prefrontal cortex to regulate amygdala activity through top-down inhibition, due to high tonic and phasic activity that results in binding to receptors that disengage the pathway between the prefrontal cortex and amygdala. This results in greater amygdala reactivity to stimuli the next-day, subsequently influencing emotion-related processing (e.g., greater difficulty distinguishing neutral and emotionally salient information, greater negative affect in relation to lower-level stressors as discussed in Section 5.1.3. below).

In the case of major depression, Goldstein and Walker (2014) suggest that REM sleep disruptions, such as shorter latency to REM onset and longer REM sleep duration, at the expense of other sleep stages, results in low levels of noradrenaline. As such the amygdala and prefrontal cortex become less sensitive in their ability to discriminate between emotional (positive or negative) and neutral stimuli, resulting in blunted emotions. In anxiety and post-traumatic stress disorder insufficient or disrupted REM sleep results in a greater anticipatory response from the amygdala that is not inhibited by the prefrontal cortex, resulting in neutral or ambiguous stimuli being perceived as threatening (Goldstein & Walker, 2014). Reactivity

toward emotional stimuli can be understood as 'affective control', which refers to the ability to control behavioural reactions towards emotional stimuli, and regulate attention paid to affective stimuli (e.g., Schweizer et al., 2020). Affective control develops alongside, but at a slower rate than, general cognitive control abilities (e.g., inhibiting attention to distractions and focusing on goal-relevant stimuli) during adolescence (Schweizer et al., 2020).

5.1.3. Sleep and thought control

Inhibitive processes, described in Section 5.1.2., such as cognitive and affective control may also extend to the ability to inhibit intrusive thoughts. Recent research has found that participants were less able to inhibit intrusive thoughts in a think/no-think task following sleep deprivation than when they were well rested (Harrington et al., 2021). A new model suggests that poor sleep not only reduces prefrontal cortex control over the amygdala, but also over the hippocampus, giving rise to intrusive thoughts and unwanted memories, for example in post-traumatic stress disorder (Harrington & Cairney, 2021). This framework suggests that poor sleep results in an inability to inhibit recollections of intrusive memories, and difficulties in downregulating negative affect associated with these unwanted memories. Specifically, Harrington and Cairney's (2021) model proposes that the right dorsolateral prefrontal cortex is responsible for inhibitory control of motor responses and suppression of unwanted memories and thoughts. When sleep deprived, the right dorsolateral prefrontal cortex is less active. As such, top-down inhibition of the amygdala and hippocampus is less effective, and more intrusive thoughts and memories occur inducing greater emotional salience and distress. Thus, sleep is further disrupted due to greater emotional distress and more intrusive thoughts and memories at night, exacerbating this cycle between poor sleep, inability to control thoughts and memories, and increased negative affect.

Ability to control thoughts and the retrieval of memories are considered fundamental skills for successful emotion regulation. Such skills are employed in many emotion regulation strategies, such as cognitive reappraisal (Engen & Anderson, 2018). Some emotion regulation strategies also rely on thought substitution, which Harrington and Cairney (2021) propose rely on different neural mechanisms (e.g., left-lateralised caudal and ventromedial prefrontal cortex) than memory suppression (e.g., right dorsolateral prefrontal cortex). As such, thought substitution may be less impacted by sleep deprivation. However, Goldstein and Walker (2014) suggest that the ventromedial prefrontal cortex is also affected by sleep deprivation. Nonetheless, emotion regulation strategies can function to downregulate negative affect as well as upregulate positive affect. Harrington and Cairney's (2021) model claims to apply to a range of mental health disorders, including depression. However, this model focuses on the inhibitory control and downregulation of unwanted memories and subsequent negative affect but does not consider the potential influence of upregulating positive affect. When considering the tripartite model (discussed in Section 5.1.4. below), positive affect plays an important role in depression symptoms (e.g., anhedonia) and therefore, would be an important inclusion in frameworks aiming to explain the relationship between sleep and depression. The present study aims to fill this gap by including positive affect in the relationship between sleep and depression.

5.1.4. Positive and negative affect

The tripartite model (Clark & Watson, 1991) suggests that poor mental health is typically characterised by high negative affect, whilst low positive affect is a unique marker of depression. Negative affect refers to emotion and mood states such as sadness, anger, and fear, which may be associated with symptoms of mental health disorders when experienced in the extreme. Positive affect refers to emotions and mood states such as happiness,

enthusiasm, and motivation, which may be lacking in the presence of depression symptoms (e.g., anhedonia). Sleep restricted adolescents are found to have lower positive affect (Lo et al., 2016) and higher negative affect (Baum et al., 2014), suggesting that sleep plays a key role in affective wellbeing, which can subsequently impact mental health. Total sleep deprivation has been associated with greater negative affect following low-stress cognitive tasks, compared to participants that were well-rested; however, both groups reacted equally negatively to high-stress tasks (Minkel et al., 2012). Therefore, poor sleepers may be less adept in handling mild stressors (e.g., low-stress cognitive tasks) than good sleepers.

These suggestions have been supported in a study that investigated the transition of older adolescents and young adults to university (19-29-years-old, M = 22.99 years; Lev Ari & Shulman, 2012). Participants that transitioned more adaptively were found to sleep for longer (eight hours or more per night) and reported higher subjective sleep quality, compared to those that experienced greater difficulties during the transition. Additionally, the group that transitioned well (and had better sleep) had lower negative affect, less stress, and were better able to handle daily stressors compared to the group that had difficulties during the transition. The group that had difficulties during the transition also experienced greater symptoms of depression and anxiety compared to the well-adjusted group. Additionally, another group emerged, which the authors refer to as the "re-adjusted group". The re-adjusted group had the shortest sleep duration at Timepoint 1 (start of the first term of university), but their sleep duration was significantly longer at Timepoint 2 (at the end of the second semester). This suggests that improvements in sleep were associated with improvements in adjustment and coping, although, direction and causality cannot be established. However, negative affect for the re-adjusted group did not significantly differ from Timepoint 1 to 2, despite sleep duration increasing. This suggests that improved sleep did not lead to reduced negative affect, and that negative affect may not have been the mechanism improving adjustment.

Increases in positive affect may have contributed to increased adjustment and coping. For example, higher positive affect may be a potential mechanism explaining the relationship between improved sleep and lower levels of depression (e.g., tripartite theory), which requires further study and is explored in the present study. However positive affect was not considered within Lev Ari and Shulman's (2012) study. The relationship between shorter adolescent sleep duration and lower positive affect has been found to have a greater effect size than the relationship between short sleep duration and increased depression, negative affect, and anxiety (Short et al., 2020). However, there is much less literature investigating positive affect than negative affect or specific mental health disorders. According to a systematic review and meta-analysis of the adolescent sleep and mood literature, more research needs to include a broad range of positive mood states (Short et al., 2020), a gap in the literature that the present study aims to address. It is also of interest that Lo et al. (2016) did not see significant differences in positive mood on the first night of sleep restriction, only on the second night onward. This may indicate that a single night of restricted sleep may not be sufficient to influence mood, and instead may require multiple nights of poorer sleep. The way in which sleep may function to maintain homeostasis over affective processes was discussed in the theories above (Section 5.1.1. and 5.1.2).

5.1.5. The current study, hypotheses, and research questions

The behavioural measurement of responses to affective stimuli can be used to test the hypotheses that poor sleep is related to poorer affective processing (e.g., discriminability) and lower inhibition of responses to affective stimuli, which acts as the mechanism between poor sleep and poor mental health. Reactivity to emotional stimuli can be behaviourally assessed, via an affective Go/No-Go paradigm, which has previously been used with children, adolescents, and adults (e.g., Tottenham et al., 2011). The online format of this study reduces

the opportunity to monitor participant wellbeing in the presence of distressing stimuli. Therefore, the Go/No-Go paradigm used facial expressions, which is more ethical for online research with adolescents than alternatives (e.g., the International Affective Picture System; Lang et al., 1997). Therefore, the Go/No-Go paradigm was used to test predictions based on Goldstein and Walker's (2014) framework.

The current study also aims to investigate the relationship between poor sleep and lower thought control abilities, to test predictions made by Harrington and Cairney's (2021) model. Tasks used to measure intrusive thoughts are often flawed, for example, the 'white bear' effect, which asks participants not to think of a 'white bear' but report when it comes to mind, remembering to report it makes it incredibly unlikely that thought suppression will be successful (e.g., see Engen & Anderson (2018) for review). Other tasks (e.g., the Think/No-Think task) would require participants to take part in a longer study or view potentially distressing images, which is neither practical nor ethical in online research. Instead, the present study uses the Thought Control Ability Questionnaire (TCAQ-20; Williams et al., 2010). In a systematic review, performance on Think/No-Think tasks has been significantly associated with the responses on the Thought Control Ability Questionnaire (Feliu-Soler et al., 2019), suggesting that this questionnaire is a reasonable proxy for intrusive thoughts and memories as measured experimentally. The present study also aims to add to existing theoretical models of the relationship between sleep and mental health, by considering both negative and positive affect, as positive affect is particularly understudied.

Despite both Goldstein and Walker (2014) and Harrington and Cairney (2021) hypothesising a key role of the prefrontal cortex in the relationship between sleep, emotion regulation and mental health issues, neither directly consider the influence of adolescent development and how this may impact the relationships in the frameworks. The prefrontal cortex is maturing up until approximately 24-years-old (Arain et al., 2013). Cognitive

processes such as affective control and cognitive control are also maturing during adolescence (Schweizer et al., 2020). Therefore, the cognitive processes proposed in Goldstein and Walker's (2014) and Harrington and Cairney's (2021) models, which rely on inhibitory control by the prefrontal cortex, may also be maturing alongside adolescent brain development. Further, the majority of mental health disorders onset prior to the age of 24-years-old (Fusar-Poli, 2019). As such, adolescents may have lower inhibitory control from the prefrontal cortex and may be more vulnerable to the relationships between poor sleep and mental health during this developmental period. Alongside the maturation of the brain alongside affective and cognitive control processes, sleep also changes during adolescence. Sleep is delayed (e.g., by circadian shifts) during adolescence often resulting in insufficient sleep duration (e.g., discussed in Section 1.4. and Chapter 2). In combination, these developmental changes occurring during adolescence likely impact upon multiple mechanisms proposed in both frameworks. Given the potential influence of adolescent development on these relationships, each hypothesis was investigated in both 13-16-year-old and 18-24-year-old adolescent samples. The older group was likely further along in brain maturation than the younger group, and as such, may have displayed more advanced inhibitory control than the younger group.

Hypotheses:

1) Poor sleep will be associated with reduced inhibition. In essence, poor sleep will be associated with lower performance on the Go/No-Go task, due to lower discriminability between emotions and difficulty inhibiting responses to affective stimuli (e.g., Goldstein & Walker, 2014). Further, poor sleep will be associated with lower thought control ability as proposed by Harrington and Cairney (2021).

2) Affective control, emotion recognition, and speed/accuracy trade-off in the Go/No-Go paradigm will significantly mediate the relationship between poor sleep and mental health. This may be related to specific emotions or disorders, for example responses to 'happy' stimuli may be uniquely associated with depression and positive affect, whilst responses to 'angry', 'sad', or 'afraid' trials may be associated with negative affect and mental health broadly (e.g., tripartite theory).
3) Lower thought control abilities will significantly mediate the relationship between poor sleep and higher mental health symptoms as proposed by Harrington and Cairney (2021).
4) Low positive affect will mediate the relationship between poor sleep and higher levels of depression; high negative affect will mediate the relationship between poor sleep and higher levels of all mental health symptoms. This is based on the suggestions by tripartite theory (Clark & Watson, 1991) that depression is uniquely characterised by low positive affect, and previous research showing relationships between poor sleep and low positive and high negative affect. This hypothesis is not clearly stated in the pre-registration, but is discussed throughout.

5.2. Methods

The study methods, hypotheses and analysis plan were pre-registered on the Open Science Framework prior to data collection (https://osf.io/9gr7t/?view_only=184e23f3905f4ad3852bfe34d15e1aa8). The study was conducted online within Gorilla Experiment Builder (Anwyl-Irvine et al., 2019). Full ethical approval was obtained from Royal Holloway University of London's ethics committee. The entire study was conducted online. Participants took approximately twenty minutes to complete the study, in a single session. Questionnaires and tasks were completed in the

following order: information and informed consent, demographic information (Section 5.2.2.6.1.), sleep questionnaires (Section 5.2.2.1.), affective Go/No-Go task (Section 5.2.2.2.), Thought Control Ability Questionnaire (5.2.2.3.), mental health questionnaires (5.2.2.4.), Positive and Negative Affect Scale (5.2.2.5), Physical Development Scale (optional, 13-16-year-olds only; Section 5.2.2.6.3.), feedback, debrief.

5.2.1. Participants and recruitment

A total of 128 adolescents aged 13-16-years-old and 145 adolescents aged 18-24-years-old took part in the present study. Final characteristics of the sample, after data cleaning, are described in Section 5.2.3.1. Participants were eligible to take part if they completed the study on a computer (desktop or laptop with working keyboard) and were resident in the UK at the time of completing the study. Participants were recruited online via an anonymous link. 18-24-year-olds completed the study directly within Prolific. As Prolific does not currently allow under 18-year-olds to participate in studies on their platform, a screening questionnaire was sent to parent/carer(s) on Prolific, filtered to ensure that their first child was within the 13-16-year-old age range. At the end of the screening questionnaire, parent/carer(s) could follow a link to a webpage to read information about the study. Parent/carer(s) were invited to share the study link (from the webpage) with their adolescent if they consented to them taking part. The Gorilla study was integrated into Qualtrics for the 13-16-year-olds to prevent the same participant taking part in the study multiple times, or to prevent bots, using the Qualtrics "prevent ballot stuffing" function. Participants were paid £5 for their time directly within Prolific (for 18-24-year-olds) or via a voucher sent to their email address (collected in Qualtrics, separate from the experiment data, for 13-16-year-olds).

5.2.2. Materials

5.2.2.1. Sleep

Habitual sleep quality and duration during the past month was measured using the Pittsburgh Sleep Quality Index (PSQI; Buysse et al., 1989). This measure is described earlier in Section 4.2.2.5.1. Due to near zero variance on the sleep medication item, this was removed from the present analyses. Therefore, the total range of possible scores for this study is 0 – 18, with higher scores indicating poorer sleep.

An amended version of the PSQI measuring sleep on the night directly before participation in the study was also used. This questionnaire was created, and is described, in Chapter 4 (Section 4.2.2.5.2., Appendix 4.1.). The questionnaire measuring last night's sleep had a total possible range of scores from 0 – 12.

5.2.2.2. Go/No-Go paradigm

It has been found that reaction times in online studies using Gorilla Experiment Builder are comparable to those collected within the laboratory (Anwyl-Irvine et al., 2019). Therefore, the present study used Gorilla Experiment Builder to programme the experiment. The current study used a Go/No-Go paradigm with emotional facial expressions. Facial expressions are not distressing or upsetting, and thus are ethical to use within online research. The affective Go/No-Go paradigm (Tottenham et al., 2011) is comprised of eight blocks of 30 trials. We also included a short practice block of eight trials. The entire task took participants approximately 9 – 10 minutes to complete. We used an identical procedure to Tottenham et al. (2011). However, as their facial stimuli were not available, we instead used the freely available KDEF facial stimuli (www.kdef.se). Similarly, to Tottenham et al. (2011), we used front-facing profiles of facial expressions (neutral, angry, afraid, sad, and happy) from ten subjects (five males and five female), which were randomly selected. ID for

the facial expression stimuli was randomly selected using a random number generator. The codes for the selected faces can be found in Appendix 5.1. The faces were edited in adobe express photoshop (https://photoshop.adobe.com/adjust) so that they were greyscale with backgrounds and hair removed by cropping faces into an oval shape.

In the centre of the screen, one face was presented per trial for 500 milliseconds, followed by a fixation cross for 1000 milliseconds to allow participants sufficient time to respond or inhibit their responses. Each block was comprised of neutral expressions and one type of emotional expression (either happy, sad, afraid, or angry). Participants were informed at the start of the block what the "go" stimuli was (either neutral or one of the emotional expressions) and were asked to press the spacebar when they saw the "go" expression. They were asked not to do anything if they saw any stimuli other than the "go" expression.

The task collects measures of cognitive control, affective control, emotion recognition, and emotion regulation, according to Tottenham et al. (2011). The false alarm rate (commission error, i.e., pressing the spacebar for "No-Go" trials) toward all stimuli is considered to be a measure of cognitive control generally. Affective control was measured by the false alarm rate for each of the emotional facial expressions separately, in experimental blocks where emotional faces were the "No-Go" stimuli. False alarm rates and hit rates were used to calculate d-prime scores (d'), which is thought to represent how well the participant recognises different emotional expressions. The d' scores were calculated in Excel, using NORMINV with 0 as the mean and 1 as the standard deviation. One d' score is calculated per participant per block, in essence, two d' scores per emotion (one for "Go" and one for "No-Go" trials). Lastly, a speed/accuracy trade-off was calculated using the ratio of z-transformed (using the scale function in R) reaction times on correct trials when emotional faces were "No-Go" (speed) and false alarm rates (accuracy).

A higher score for affective control indicates greater levels of false alarm rates, demonstrating lower ability to control responses to affective stimuli. Higher d' scores indicate better discrimination between emotional and affective stimuli (i.e., higher emotion recognition). Higher speed/accuracy trade-off ratio indicates slower reaction times were required to respond accurately, suggesting greater difficulty in inhibiting responses to affective stimuli.

5.2.2.3. Thought control

Due to the timing of the online study and the ethics of showing distressing images online without researchers present, we opted to use a questionnaire rather than an experimental (Think/No-Think) task to measure thought control. We used the shorter, 20 item, Thought Control Ability Questionnaire TCAQ-20; Williams et al., 2010) to measure participants' self-reported ability to control their thoughts. Participants responded on a scale of completely disagree (1) to completely agree (5) in response to items such as "I often cannot avoid having upsetting thoughts" (reverse scored) and "it is very easy for me to free myself from troublesome thoughts". A total score with a possible range of 20 – 100 is produced, with higher scores indicating higher ability to control unwanted thoughts.

5.2.2.4. Mental health

The Revised Children's Anxiety and Depression Scale (RCADS) is widely used in mental health practice (Wolpert et al., 2015) and therefore was selected for the current research, which may have implications for clinical practice. The present study used the reduced, 25-item scale (Ebesutani et al., 2012) to save time during the online study. Participants responded on a scale of (0) "Never" to (3) "Always" to items such as "I worry that something bad will happen to me". Ten items are included in the depression subscale, and 15 in the anxiety subscale. This creates a total range of possible raw scores of 0 – 30 for depression,

and 0 – 45 for anxiety, with higher scores indicating greater levels of symptoms. The RCADS provides age and gender normed scores that can be calculated from the raw scores. However, normed data are only available up until 18-years-old. Therefore, only raw scores are used in the current study. Due to the RCADS not being validated for use with older adolescents (18-24-years-old), we also use the Depression Anxiety and Stress Scale (DASS-21), described earlier in Section 4.2.2.8. The DASS-21 produces a possible range of scores from 7 – 28 for each subscale with higher scores indicating greater levels of symptoms.

5.2.2.5. Affect

The Positive and Negative Affect Scale (PANAS; Watson et al., 1998), described in Section 4.2.2.3., was used to measure affect. Higher scores indicate that the participant was experiencing that affect to a greater extent.

5.2.2.6. Additional measures

5.2.2.6.1. Demographics

All participants were asked to report their age (in years), gender, and ethnicity to characterise the samples. Demographics of the final sample are reported in the results.

5.2.2.6.2. SASS-Y

To investigate weekend and weekday sleep differences, data were collected using the SASS-Y (Dietch et al., 2019) described in Section 3.2.2.3. Due to the complexity of the current models (e.g., as a result of limitations in defining higher order constructs), this questionnaire was excluded from the analyses. Difficulties in cleaning and scoring the data for this questionnaire would also have resulted in more missing data and subsequently a smaller sample.

5.2.2.6.3. Puberty

13-16-year-olds participants were given the option to complete the Pubertal Development Scale (PDS; Earis et al., 2000) described in Section 2.2.2.2.5. This questionnaire is only used for exploratory analyses of any developmental differences between the two age-groups' models.

5.3. Analysis plan

5.3.1. Justification of the analysis plan

Structural equation modelling (SEM) is able to test the inter-relations between multiple variables and underlying latent structures whilst also accounting for measurement error. Variance based partial least squares SEM (PLS-SEM) is used to analyse the current data. PLS-SEM is used rather that the more traditional covariance-based SEM (CB-SEM) as PLS-SEM focuses on explained variance, maximising predictive values (R^2) and minimising measurement error terms rather than focusing on best-fit, estimating parameters, and establishing covariance like in CB-SEM (Hair et al., 2014). This makes PLS-SEM more appropriate for the current research questions, which focus on the predictive relationships between the latent variables. Nonetheless, PLS-SEM does have limitations such as biases in parameter estimates and a lack of established goodness-of-fit criteria, yet these are not priorities within the current research questions. CB-SEM is also typically used for testing or comparing existing theoretical models; however, PLS-SEM is more adept at integrating multiple theories and investigating new or exploratory relationships, which is more in keeping with the aims of the current study.

PLS-SEM is also better equipped at handling more complex models than CB-SEM. The predicted model is complex in terms of the number of latent and manifest variables and hypothesised indirect effects. Measurement error, which can be evaluated in SEM, is an important consideration when investigating indirect effects; however, until recently, SEM has

not been equipped in handling such indirect effects (Muller et al., 2005). PLS-SEM is superior to CB-SEM for conducting mediation path analyses within the structural equation model (Nitzl et al., 2016; Sarstedt et al., 2020). Additionally, CB-SEM mainly handles interval level data, whereas PLS-SEM is capable of handling different types of data. Some researchers argue that Likert data is interval level, whilst others argue it is ordinal (Mondiana et al., 2018). Whilst Likert scales have previously been used in traditional CB-SEM (Mondiana et al., 2018), they are commonly criticised for non-normality of data (Leung, 2011). The issues with the normality of Likert scales is more problematic with smaller (e.g., 4-points) scales (Leung, 2011), which were used in the well-established questionnaires selected for this research. As PLS-SEM is a non-parametric statistical analysis, this approach is also better able to deal with non-normally distributed data, unlike CB-SEM.

5.3.2. Data screening and cleaning

Data were cleaned and analysed within R Studio. Analyses were conducted using the semPLS package (Monecke, 2015) and the default centroid weighting scheme. Unfortunately, semPLS does not support mediation analyses within the model, so data were extracted from R and mediation pathways were analysed in 4th version of the SmartPLS software (Ringle et al., 2022), which unfortunately is not freely available, except during a trial period. The same models were used in SmartPLS as in semPLS. The SmartPLS model account for all direct and indirect pathways, including serial mediation (multiple mediators between a predictor and outcome construct). PLS-SEM is a non-parametric statistical analysis, and therefore accounts for skew and kurtosis. An error occurred that made the questionnaire data from the cognitive emotion regulation questionnaire (CERQ-18; Garnefski & Kraaij, 2006) unusable, so this deviates from the pre-registration and was removed from the present study. The analysis plan also deviates from the pre-registration in that the SASS-Y was excluded from analyses (as mentioned in Section 5.2.2.6.2.). Last Night's sleep was also removed from the

analyses to reduce convergence failures in the bootstrap iterations. Instead a separate model was conducted for last night's sleep, which matched the habitual sleep model in regard to structure and pathways modelled. Higher order latent constructs are also not used due to limitations of the software and R packages.

5.3.3. Evaluating the measurement (outer) model

Measurement model evaluation was conducted manually in Excel or within SmartPLS and is reported in Section 5.4.3. The model was predicted to be reflective. Reflective models are evaluated for internal consistency reliability, indicator reliability, convergent validity, and discriminant validity. In the pre-registration, internal consistency was going to be measured via Cronbach's alpha. However, Dillon-Goldstein's rho is a preferred method of measuring internal consistency in PLS-SEM due to lower sensitivity to the number of items in a construct (Hair et al., 2017) and therefore is used instead. If the average variance extracted (AVE) is above 0.5 this indicates convergent validity (Hair, Sarstedt, Pieper, et al., 2012; Olya, 2017). Up to 20% of problematic indicators may be removed if required. Discriminant validity is assessed using the HTMT criteria. This is the preferred method for PLS-SEM, due to greater sensitivity in detecting discriminant validity issues compared to Fornell-Larcker criteria (Henseler et al., 2014). The HTMT criteria states that if correlations between two latent constructs are below .90 (liberal) or .85 (conservative) then discriminant validity is achieved (Henseler et al., 2014). If correlations are above these values, then discriminant validity cannot be established, suggesting the two latent constructs measure the same concept. In this instance, providing it makes theoretical sense, the constructs may be collapsed into a single latent variable, or removed due to redundancy. Any amendments made to the model, were theoretically justified, supported by prior research and/or make intuitive sense. Model amendments or removal or items are explicitly stated alongside the justification for their removal.

5.3.4. Evaluating the structural (inner) model

Non-parametric bootstrapping is needed in PLS-SEM to test the significance of the coefficients. Specifics of the bootstrapping procedures are reported, such as the number of bootstrapping iterations, the bootstrap sample size, and software parameters (Streukens & Leroi-Werelds, 2016). We used the recommended 10,000 minimum bootstrap iterations (Streukens & Leroi-Werelds, 2016). Bias-corrected 95% confidence intervals (BCa 95% CI) were calculated (from bias-corrected and accelerated [BCa] bootstrapping), as BCa corrects for biases caused by skewness and is superior to other bootstrapping methods in terms of power, accuracy, and type one error rate (Streukens & Leroi-Werelds, 2016). BCa 95% CI were used to infer the significance of pathways in the model. If the range of the BCa 95% CI for a pathway did not include zero, then the null hypotheses was rejected, and the pathway was considered significant.

5.4. Results

The 13-16-year-old model converged after 15 iterations, and the 18-24-year-old model converged after 9 iterations. 10,000 bootstrap samples were set and, a fixed seed was used to allow replicability. Bias corrected and accelerated (BCa) 95% confidence intervals (CI) are used to infer significance, as stated in Section 5.3.4. Descriptive statistics for each latent construct per age group are presented in Figures 5.1. (questionnaires) and 5.2. (Go/No-Go paradigm) below.

5.4.1. Data cleaning

7.81% (N = 10) of participants were removed from the 13-16-year-old cohort during cleaning, leaving a total sample of 118 participants. These ten participants were removed due to having outlying data in the speed/accuracy trade-off analyses, which meant the ratios could not be calculated accurately if these data were included. One of these participants also had

uncleanable sleep data. 6.21% (*N* = 9) of participants were removed from the 18-24-year-old group. Six of the removed participants had outlying data in the speed/accuracy trade-off. A further three participants had uncleanable responses in the sleep questionnaire. This left a total of 136 participants in the 18-24-year-old group. The characteristics of the final samples are reported below in Section 5.4.2.

5.4.2. Sample characteristics

The demographics of the 13-16-year-old sample are reported below in Table 5.1. and for 18-24-year-olds in Table 5.2. The mean age of for the 13-16-year-old group was 14.03 years (*SD* = 0.95 years) and the mean age for the 18-24-year-old group was 21.28 years-old (*SD* = 1.94 years). Figure 5.1. and 5.2. below demonstrate descriptive statistics for all measures included in the structural equation model.

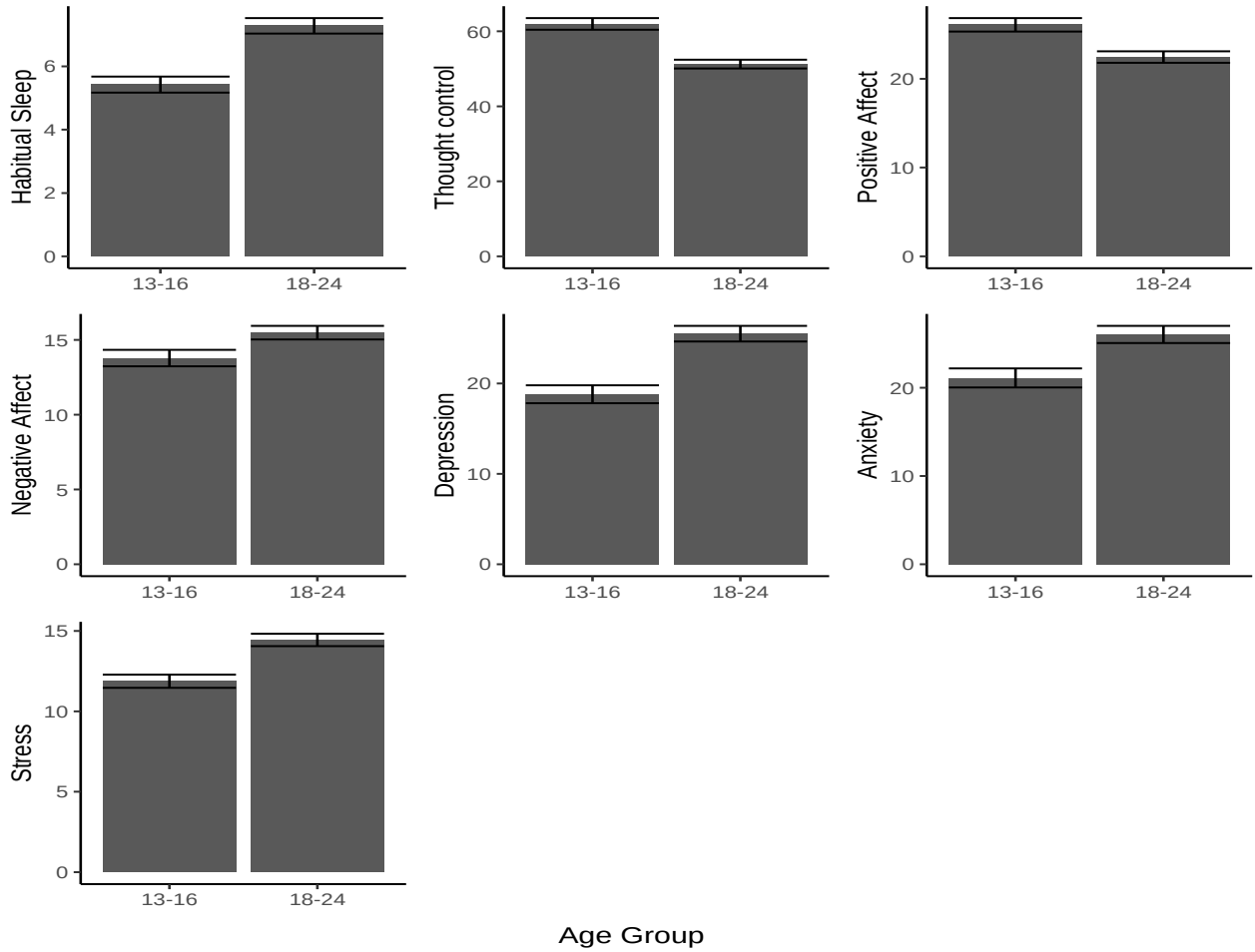
Table 5.1. Demographic variables for the 13-16-year-old group

		N =	% of sample
Ethnicity	White British	101	85.59
	Asian	4	3.39
	Black	4	3.39
	Mixed ethnicity	7	5.93
	Other ethnicity	2	1.69
Gender	Female	59	50.00
	Male	58	49.15
	Non-binary	1	0.85
Age	13-years-old	41	34.75
	14-years-old	42	35.59
	15-years-old	25	21.19
	16-years-old	10	8.47

Table 5.2. Demographic variables for the 18-24-year-old group

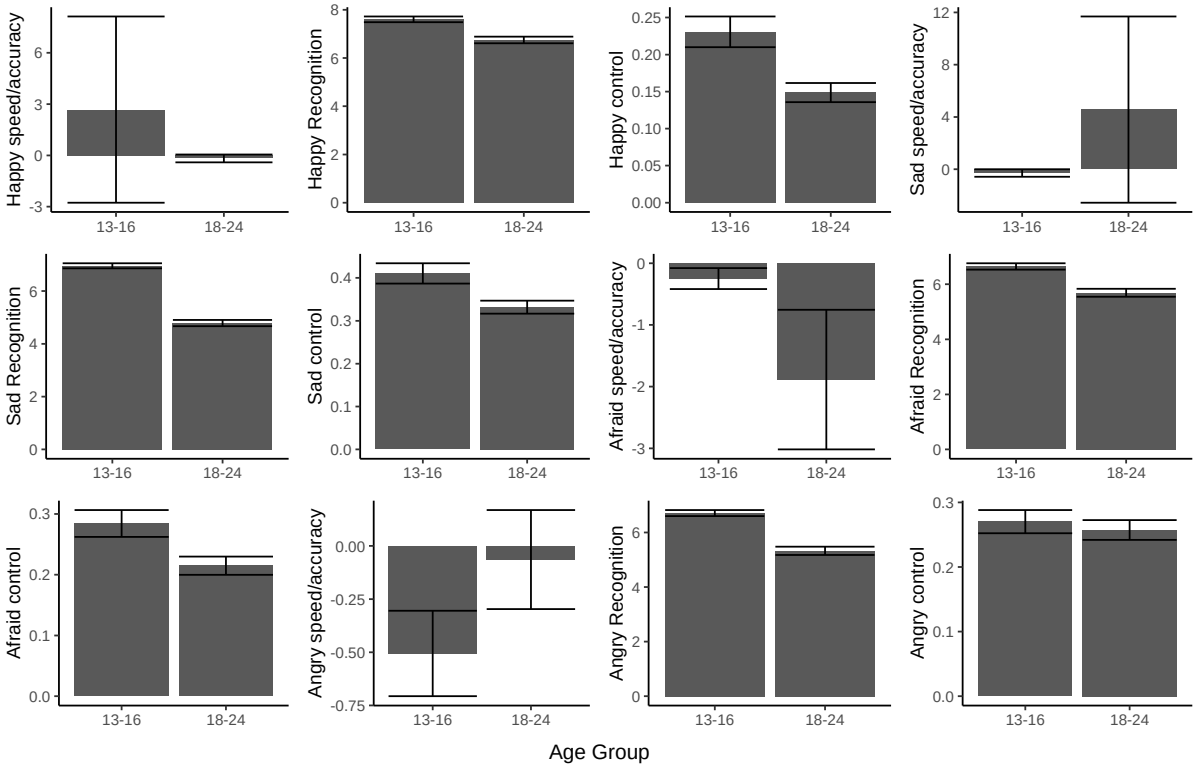
		N =	% of sample
Ethnicity	White British	81	59.56
	Asian	18	13.24
	Black	11	8.09
	Mixed ethnicity	7	5.15
	Other ethnicity	19	13.97
Gender	Female	87	63.97
	Male	45	33.09
	Non-binary	4	2.94
Age	18-years-old	13	9.56
	19-years-old	15	11.03
	20-years-old	25	18.38
	21-years-old	19	13.97
	22-years-old	20	14.71
	23-years-old	20	14.71
	24-years-old	24	17.65

Figure 5.1. Bar graphs showing the means and standard error bars for latent variables measured by questionnaires



Note: Habitual sleep was measured with the PSQI, with higher scores indicating poorer sleep. Thought control was measured with the TCAQ, with higher scores indicating better self-reported ability to control thoughts. Positive and negative affect were measured with the PANAS, with higher scores indicating higher levels of that affect. Depression, anxiety, and stress were measured with the DASS-21, with higher scores indicating higher levels of symptoms associated with these difficulties.

Figure 5.2. Bar charts showing the mean and standard error bars for latent constructs from the Go/No-Go paradigm per age group



Note: For each emotion (happy, sad, afraid, angry) different measures were calculated from participant responses to the Go/No-Go paradigm. The conceptualisation of speed/accuracy trade-off, recognition, and affective control are described in Section 5.2.2.

5.4.3. Measurement model

A reflective model converged for both age groups. Table 5.3. below outlines the manifest variables included in each latent construct in the models. Last night's sleep was removed due to resolve convergence issues, and was analysed in a separate model. One item ('I feel scared if I have to sleep on my own') from the RCADS anxiety subscale was removed due to binary responses in the 13-16-year-old group meaning it was of a different scale than all other responses in the latent construct, and therefore could not be fitted in the SmartPLS model. Removal of this item from all models allowed consistency in the structural model across age groups and direct and indirect analyses.

Table 5.3. Manifest variables included within latent constructs

Latent Construct	Manifest Variables included (N = number of included variables)
Habitual Sleep	Component 1 – 5 and 7 on original PSQI (*N* = 6*)
Positive Affect	All positive affect items from the Positive and Negative Affect Scale (PANAS; *N* = 10)
Negative Affect	All negative affect items from the Positive and Negative Affect Scale (PANAS; *N* = 10)
Thought Control	All items from the reduced version of the Thought Control Ability Questionnaire (TCAQ-20; *N* = 20)
Cognitive control	Total false alarm rate across all Go/No-Go trials (*N* = 1)
Affective Control	False alarm rates per participant per emotion. There were 9 No-Go trials per emotion that were used to calculate a single item false alarm rate per emotion (sad [*N* = 1], afraid [*N* = 1], angry [*N* = 1], happy [*N* = 1])
Emotion Recognition	There were a total of 21 Go and 9 No-Go trials used to calculate each d' score. Two scores are calculated per emotion, one when the emotion was the 'Go' trial and one when neutral was the 'Go' trial and the emotion was 'No-Go' (sad [*N* = 2], afraid [*N* = 2], angry [*N* = 2], happy [*N* = 2])
Speed/Accuracy	A speed accuracy ratio was calculated for each emotion from z-transformed reaction times (for correct 'Go' trials) to false alarm rates (sad [*N* = 1], afraid [*N* = 1], angry [*N* = 1], happy [*N* = 1])
Depression	DASS-21 (*N* = 7) and RCADS (*N* = 10) depression items
Anxiety	DASS-21 (*N* = 7) and RCADS (*N* = 14*) anxiety items
Stress	Stress items from the DASS-21 (*N* = 7)

Note: * As explained in Section 5.4.3., one item from the RCADS and one from the PSQI (habitual sleep measure) was removed from all models. Last night's sleep was also not included within the models to resolve convergence issues

5.4.3.1. Validity

Average variance extracted (AVE) above .50 indicates convergent validity. AVE was calculated in Excel, using a formula from Analysis Inn (2020) and factor loadings extracted from R. In the 18-24-year-old model, AVE was below the threshold for negative affect (.43), thought control (.38), habitual sleep (.41), and anxiety (.39), suggesting these variables did not demonstrate convergent validity. AVE was above the threshold for positive affect (.53), stress (.51), depression (.504), afraid recognition (.68), angry recognition (.70), happy recognition (.66), and sad recognition (.58). For 13-16-year-olds, negative affect (.58), positive affect (.55), stress (.59), depression (.56), afraid recognition (.58), angry recognition (.50), happy recognition (.59), and sad recognition (.53) met the threshold for convergent validity. Thought control (.49), habitual sleep (.41), and anxiety (.48) were again below the threshold. Due to the recommendation that a maximum of 20% of low loading indicators are removed (Hair et al., 2017a, as cited by Ali Memon et al., 2017) and that removing up to 20% of low loadings items did not bring the AVE above the threshold in both models, all indicators were included in the latent construct, with the exception of those stated in Section 5.4.2.

If the HTMT criteria is below .90 (liberal) or .85 (conservative) then discriminant validity is achieved. The HTMT value was calculated in SmartPLS. In the 18-24-year-old model, neither criterion was met for sad and afraid recognition (1.24), sad recognition and cognitive control (1.16), sad and angry recognition (1.10), sad recognition and sad affective control (1.09), sad and happy recognition (.98), happy and afraid recognition (.96), afraid recognition and afraid control (.94), and cognitive control and afraid recognition (.91). The liberal, but not the conservative, criteria were met for happy recognition and happy control (.90), cognitive control and angry recognition (.88) and stress and anxiety (.88). All other

variables met the HTMT discriminant validity criteria. For the 13-16-year-old model, neither criterion was met for sad and angry recognition (1.90), happy and angry recognition (1.51), sad and afraid recognition (1.48), angry and afraid recognition (1.34), sad and happy recognition (1.24), happy and afraid recognition (.996), cognitive control and angry recognition (.98), and stress and depression (.92). The liberal, but not the conservative criteria, were met for stress and anxiety (.89), and depression and anxiety (.87). All other variables met the HTMT criteria. The lack of convergent validity for some of the Go/No-Go variables makes sense, given that they rely on similar variables in their calculations (e.g., false alarm rates). It does not make theoretical sense to collapse the variables, as Goldstein and Walker (2014) propose different relationships for emotion recognition than affective control for instance, and this may also differ based on the emotion. It also does not make theoretical sense to collapse depression, anxiety, and stress into combined latent variables, due to different predictions regarding their relationships with positive affect. However, it is worth noting that convergent validity issues are a limitation in the present study.

5.4.3.2. Reliability

Dillon-Goldstein's rho was calculated in R as a measure of composite reliability. For the 13-16-year-old model, Dillon-Goldstein's rho was below the .70 threshold for angry recognition (.54), and sad recognition (.61). Whereas, stress (.91), depression (.96), anxiety (.95), positive affect (.92), negative affect (.93), happy recognition (.72), afraid recognition (.71), thought control (.95) and habitual sleep (.80) all met the criteria. For 18-24-year-olds, the criterion was met for all latent constructs: stress (.88), depression (.94), anxiety (.93), positive affect (.92), negative affect (.88), sad recognition (.72), happy recognition (.78), angry recognition (.82), afraid recognition (.81), thought control (.92) and habitual sleep (.80).

5.4.4. Structural model

Figures 5.3. (13-16-year-olds) and 5.4. (18-24-year-olds) below demonstrate the pathways included in the structural model.

Figure 5.3. Structural model for 13-16-year-olds with beta-values showing direct pathways

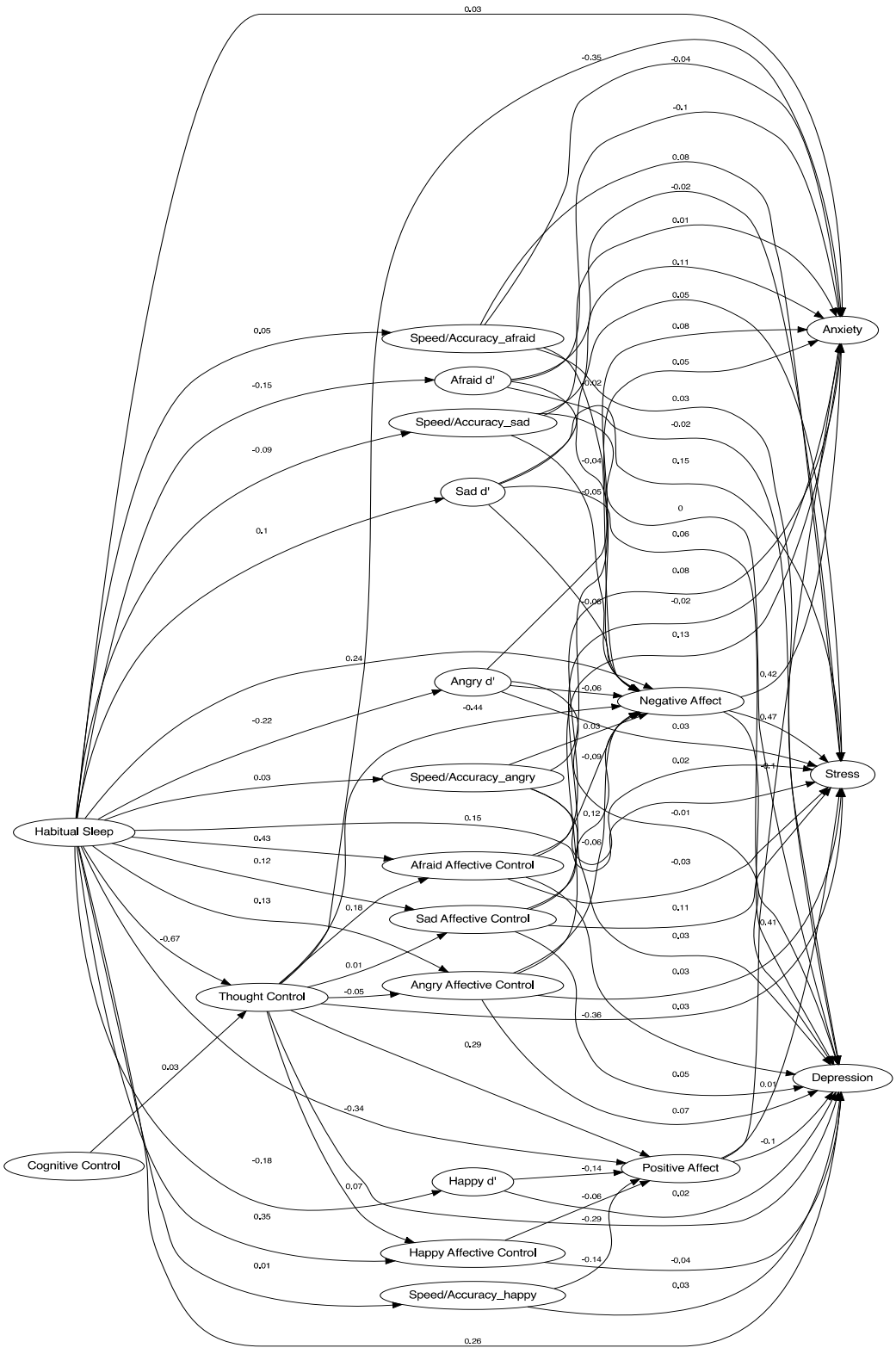
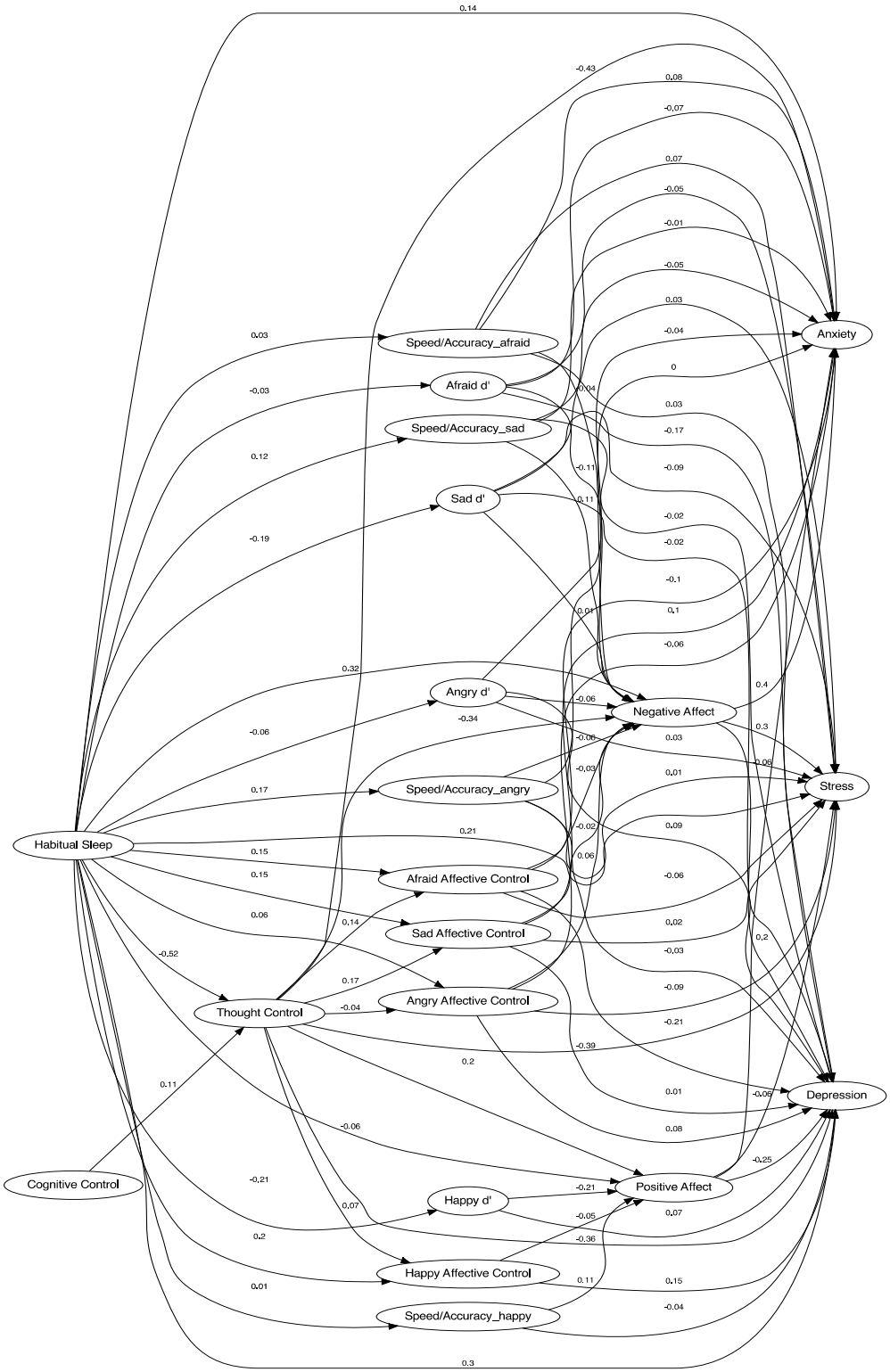


Figure 5.4. Structural model for 18-24-year-olds with beta-values showing direct pathways



5.4.4.1. Direct effects

Table 5.4. below shows that for both 13-16-year-old and 18-24-year-old participants, higher thought control ability was significantly associated with lower negative affect, depression, anxiety, and stress. Higher negative affect was significantly associated with higher levels of depression, anxiety, and stress. Higher positive affect was associated with lower levels of depression only. Poor habitual sleep was associated with lower levels of thought control abilities, as well as greater levels of negative affect and depression.

For 13-16-year-olds only, poorer habitual sleep was significantly associated with more false alarms (worse affective control) for afraid and happy faces, and lower positive affect. For this age group only, higher thought control ability was associated with greater positive affect. Higher d' for sad faces (greater recognition) was also associated with greater levels of stress. For 18-24-year-olds only, poor habitual sleep was significantly associated with higher anxiety and stress. Additionally, fewer false alarms (higher affective control) towards afraid faces were associated with higher levels of depression.

The direct effects in the model for last night's sleep were mostly the same as in the habitual sleep model (see Appendix 5.2.). In the 13-16-year-old model, last night's sleep was not significantly associated with afraid or happy affective control, nor with negative affect, despite these relationships being significant in the habitual sleep model. For 18-24-year-olds, last night's sleep was significantly associated with speed/accuracy trade-off for sad faces, which was not seen for habitual sleep. Last night's sleep was not significantly associated with anxiety, whilst it was associated with habitual sleep. Afraid affect control was also not significantly associated with depression in the model for last night's sleep.

Table 5.4. Beta values and 95% BCa CI for direct pathways in the 13-16 and 18-24-year-old habitual sleep models

			13-16-year-olds			18-24-year-olds		
Path from…	Path to…		ß	Lower CI	Upper CI	ß	Lower CI	Upper CI
Habitual Sleep	**Thought Control**		**-.674**	**-.748**	**-.568**	**-.522**	**-.627**	**-.348**
	Speed Accuracy Trade-off	Sad	-.093	-.325	.100	.123	-.100	.297
		Happy	.009	-.245	.236	.008	-.201	203
		Afraid	.047	-.104	.190	.030	-.174	.202
		Angry	.033	-.208	.239	.169	-.012	.303
	Affective Control	Sad	.120	-.114	.305	.152	-.071	.336
		Happy	**.350**	**.094**	**.520**	.203	-.101	.447
		Afraid	**.435**	**.196**	**.614**	.147	-.141	.354
		Angry	.128	-.110	.350	.063	-.183	.277
	Emotion Recognition	Sad	.104	-.188	.372	-.191	-.360	.096
		Happy	-.177	-.346	.224	-.210	-.400	.083
		Afraid	-.150	-.270	.158	-.034	-.181	.250
		Angry	-.220	-.377	.195	-.062	-.204	.167
	Positive Affect		**-.340**	**-.545**	**-.116**	-.057	-.265	.194
	Negative Affect		**.244**	**.002**	**.470**	**.321**	**.128**	**479**
	Depression		**.258**	**.124**	**.392**	**.300**	**.168**	**.446**
	Anxiety		.035	-.153	.214	**.135**	**.014**	**266**
	Stress		.146	-.000	.300	**.211**	**.049**	**.354**
Positive Affect	**Depression**		**-.096**	**-.190**	**-.016**	**-.250**	**-.350**	**-.131**
	Anxiety		-.098	-.222	.013	-.062	-.145	.053
	Stress		.009	-.112	.105	-.061	-.181	.050
Negative Affect	**Depression**		**.410**	**.027**	**.517**	**.202**	**.054**	**.362**
	Anxiety		**.418**	**.222**	**.581**	**.340**	**.254**	**.533**
	Stress		**.466**	**.310**	**.606**	**.298**	**.157**	**.439**
Cognitive Control	Thought Control		.033	-.085	.188	.114	-.078	.281
Emotion Recognition (Sad)	Negative Affect		-.059	-.224	.158	.005	-.188	.272
	Depression		.062	-.016	.251	-.020	-.140	.146
	Anxiety		.107	-.014	.273	-.048	-.193	.083
	Stress		**.150**	**.072**	**.317**	-.091	-.257	.124
Affective Control (Sad)	Negative Affect		.121	-.086	.333	-.022	-.227	.190
	Depression		.053	-.065	.184	.011	-.129	.144
	Anxiety		-.020	-.165	.149	.104	-.025	.233
	Stress		.106	-.023	.276	.022	-.167	.168
Speed/Accuracy (Sad)	Negative Affect		-.047	-.172	.056	.106	-.057	.271
	Depression		-.002	-.091	.090	-.018	-.144	.096
	Anxiety		.011	-.121	.137	-.005	-.130	.135

	Stress		.051	-.076	.176	.029	-.084	.141
Emotion Recognition (Afraid)	Negative Affect		-.035	-.205	.151	-.110	-.327	.151
	Depression		.019	-.137	.129	-.167	-.325	.068
	Anxiety		-.100	-.285	.077	-.073	-.241	.088
	Stress		-.020	-.203	.153	-.051	-.246	.169
Affective Control (Afraid)	Negative Affect		-.092	-.297	.090	-.030	-.253	.220
	Depression		.026	-.117	.159	**-.214**	**-.411**	**-.018**
	Anxiety		.080	-.107	.250	-.095	-.282	.055
	Stress		-.031	-.196	.125	-.056	-.268	.146
Speed/Accuracy (Afraid)	Negative Affect		-.019	-.141	.099	-.036	-.194	.130
	Depression		.029	-.076	.157	.028	-.076	.122
	Anxiety		-.036	-.153	.116	.082	-.005	.168
	Stress		.075	-.046	.221	.072	-.025	.156
Emotion Recognition (Angry)	Negative Affect		-.061	-.317	.155	-.060	-.337	.155
	Depression		-.013	-.157	.087	.091	-.058	.294
	Anxiety		.080	-.062	.236	-.039	-.182	.223
	Stress		.026	-.116	.160	.025	-.156	.231
Affective Control (Angry)	Negative Affect		-.063	-.307	.137	.058	-.180	.249
	Depression		.073	-.026	.214	.080	-.067	.284
	Anxiety		.126	-.028	.283	-.062	-.199	.106
	Stress		.031	-.091	.178	-.089	-.256	.112
Speed/Accuracy (Angry)	Negative Affect		.033	-.094	.181	-.062	-.236	.103
	Depression		.034	-.073	.127	-.030	-.144	.083
	Anxiety		.049	-.034	.168	.002	-.101	.105
	Stress		.023	-.100	.124	.012	-.097	.122
Emotion Recognition (Happy)	Positive Affect		-.140	-.310	.038	-.209	-.612	.047
	Depression		.022	-.090	.130	.074	-.165	.274
Affective Control (Happy)	Positive Affect		-.065	-.229	.075	-.049	-.460	.179
	Depression		-.044	-.202	.075	.155	-.028	.340
Speed/Accuracy (Happy)	Positive Affect		-.137	-.275	.013	.107	-.084	.296
	Depression		.027	-.062	.098	-.039	-.156	.084
Thought Control	Positive Affect		**.294**	**.093**	**.467**	.197	-.040	.392
	Negative Affect		**-.442**	**-.625**	**-.237**	**-.338**	**-.487**	**-.185**
	Depression		**-.294**	**-.426**	**-.173**	**-.361**	**-.494**	**-.209**
	Anxiety		**-.346**	**-.487**	**-.171**	**-.431**	**-.557**	**-.297**
	Stress		**-.363**	**-.493**	**-.217**	**-.386**	**-.523**	**-.246**
	Affective Control	Sad	.011	-.234	.257	.175	-.059	.370
		Happy	.068	-.148	.279	.066	-.178	.288
		Afraid	.180	-.023	.381	.143	-.087	.339
		Angry	-.053	-.247	.201	-.045	-.276	.198

Note: CI = 95% bias corrected and accelerated confidence interval (95% BCa CI). Pathways in bold are significant according to 95% BCa CI

5.4.4.2. Indirect effects

Given that the direct relationships in the habitual and last night's sleep models were similar (as described in Section 5.4.4.1.), mediation analyses were only examined in one of the models. As the habitual Pittsburgh Sleep Quality Index is widely used in research, the following analyses are conducted using habitual sleep, to increase comparability to previous research.

5.4.4.2.1. Sleep to mental health via affect

According to the p-value calculated in SmartPLS, positive affect did not significantly mediate the relationship between habitual sleep and depression in the 13-16-year-old group, *t* = 1.82, 95% BCa CI = .006 - .083, p = .070. However, the BCa confidence intervals (BCa CI = .006 - .083) suggest the mediation is significant. Positive affect did not significantly mediate the relationship between habitual sleep and depression for the 18-24-year-old group, *t* = .22, 95% BCa CI = -.06 - .07, *p* = .827.

In the 13-16-year-old group negative affect significantly mediated the relationship between sleep and depression, *t* = 2.31, 95% BCa CI = .02 - .18, *p* = .021, anxiety, *t* = 1.96, 95% BCa CI = .02 - .21, *p* = .050, and stress, *t* = 2.14, 95% BCa CI = .02 - .22, *p* = .032. In the 18-24-year-old group, negative affect was also a significant mediator of the relationship between habitual sleep and depression, *t* = 2.04, 95% BCa CI = .02 - .14, *p* = .041, anxiety, *t* = 2.92, 95% BCa CI = .06 - .23, *p* = .003, and stress, *t* = 2.85, 95% BCa CI = .05 - .19, *p* = .004.

5.4.4.2.2. Sleep to mental health via Go/No-Go

In both the 13-16-year-old and 18-24-year-old group, performance on the Go/No-Go task (affective control, emotion recognition, and speed/accuracy trade-off for happy, sad, afraid, and angry faces) did not significantly mediate the relationship between poor sleep and depression, anxiety, or stress. The results of these analyses are reported in Appendix 5.3.

5.4.4.2.3. Sleep to mental health via thought control

In the 13-16-year-old group, self-reported ability to control thoughts significantly mediated the relationship between habitual sleep and depression, t = 4.20, 95% BCa CI = .10 - .29, p = <.001, anxiety, t = 3.73, 95% BCa CI = .11 - .36, p <.001, and stress, t = 4.21, 95% BCa CI = .13 - .35, p <.001. For the 18-24-year-olds, thought control ability was also a significant mediator of the relationship between habitual sleep and depression, t = 4.34, 95% BCa CI = .10 - .28, p <.001, anxiety, t = 5.03, 95% BCa CI = .15 - .34, p <.001, and stress, t = 4.68, 95% BCa CI = .13 - .31, p <.001.

Thought control was also a serial mediator, along with negative affect, for the relationship between habitual sleep and depression, t = 3.70, 95% BCa CI = .07 - .21, p <.001, anxiety, t = 3.02, 95% BCa CI = .06 - .23, p = .003, and stress, t = 3.53, 95% BCa CI = .08 - .24, p <.001, in the 13-16-year-old model. According to the p-value, thought control and positive affect were not serial mediators of the relationship between habitual sleep and depression, however the BCa 95% CI do not contain zero, suggesting the relationship is significant in this sample, t = 1.70, 95% BCa CI = .003 - .05, p = .090. In the 18-24-year-old model, thought control and negative affect were also significant serial mediators of the relationship between habitual sleep and depression, t = 2.13, 95% BCa CI = .01 - .07, p = .034, anxiety, t = 3.54, 95% BCa CI = .03 - .11, p <.001, and stress, t = 2.97, 95% BCa CI = .02 - .10, p =.003. Thought control and positive affect were not significant serial mediators of

the relationship between habitual sleep and depression, t = 1.45, 95% BCa CI = -.01 - .06, p = .147.

5.4.4.2.4. Puberty

13-16-year-olds completed a self-reported questionnaire about pubertal development. An exploratory multiple linear regression was conducted between self-reported puberty and the latent constructs that showed different relationships between the models for the two age-groups. There was no significant association between self-reported pubertal development positive affect, ß = .01, SE = .04, t = .22, p = .824, habitual sleep, ß = -.22, SE = .12, t = 1.81, p = .073, depression, ß = -.02, SE = .05, t = -.32, p = .748, thought control, ß = .01, SE = .03, t = .43, p = .667, happy false alarm, ß = -1.74, SE =1.45, t = -1.20, p =.234, afraid false alarm, ß = 1.12, SE = 1.36, t = .83, p = .410, and stress, ß = .10, SE = .11, t = .95, p = .345. There was a significant positive relationship between sad recognition, ß = -.49, SE = .24, t = -2.00, p = .049, and self-reported pubertal development, in that, participants further into pubertal development recognised sad stimuli more accurately (higher d').

5.5. Discussion

The present study collected data online with 13-16-year-olds and 18-24-year-olds to examine whether adolescent development may affect the relationships between sleep, mental health, and their potential mediators. Each of the mediators examined in the present study are discussed in the sections below. Suggestions for future research and final conclusions are then made. Poor habitual sleep was associated with greater false alarms (lower affective control) for happy and afraid facial expressions only for the younger group and not for last night's sleep. This shows partial support for Hypothesis 1 based on Goldstein & Walker's (2014) framework, but not full support (discussed further in Section 5.5.2.). Inhibition on the Go/No-Go paradigm was also not a significant mediator of the relationship between sleep and

mental health despite predictions made by Goldstein and Walker (2014) and Harrington and Cairney (2021), which informed Hypothesis 2. However, thought control was a significant mediator between sleep and mental health supporting Hypothesis 3, based on Harrington and Cairney's (2021) model (discussed in Section 5.5.3.). Regarding Hypothesis 4, poor habitual sleep was associated with higher negative affect in both age groups, but with lower positive affect only in the younger age group. Additionally, last night's sleep was not directly associated with negative affect for the younger group but was still associated with positive affect. Negative affect was a significant mediator of the relationship between poor sleep and mental health supporting the tripartite theory and the novel inclusion of sleep within this relationship. However, positive affect mediated relationship between sleep and depression only in the younger group. These findings are discussed further in Section 5.5.1. below.

5.5.1. Positive and negative affect

The present study addresses a gap in the literature, identified by Baum et al. (2014) who suggested that future research investigate the relationship between sleep, affect, and tripartite models of mental health. The current research found that poor habitual sleep was associated with higher negative affect in both age groups. However, last night's sleep was directly associated with negative affect only for 18-24-year-olds and not 13-16-year-olds, suggesting that negative affect in younger adolescents may require cumulative effects of poor sleep. Negative affect was a significant mediator between habitual sleep and depression, anxiety, and stress for both age groups. In 18-24-year-olds only, poor habitual sleep was directly associated with higher anxiety and stress. However, in both age groups, higher negative affect was directly associated with all mental health conditions. This supports the suggestions in tripartite model (Clark & Watson, 1991) that high negative affect is universal to mental health disorders. The present study adds to this theory by including sleep as a

contributing factor to these relationships and demonstrating that negative affect is a required mediator for younger adolescents in the relationship between poor habitual sleep and higher anxiety and stress but is directly associated for older adolescents.

In both age groups, sleep was directly associated with depression. Interestingly, better sleep and higher thought control ability were directly associated with higher positive affect only for 13-16-year-olds, yet positive affect was negatively associated with depression in both age groups. This finding shows partial support for Hypothesis 4, based on tripartite theory, in that depression is uniquely characterised by low positive affect; however, these relationships appear to be associated with sleep only for the younger age group. This is contrary to a systematic review that found good sleep is associated with higher positive affect in children, adolescents, and adults (e.g., Ong et al., 2017). For the 13-16-year-olds only, poor habitual sleep and low thought control ability were associated with lower levels of positive affect. Positive affect also acted as a serial mediator with thought control ability between habitual sleep and depression for this age group only. It is important to note that this relationship was only significant according to the confidence intervals, not the p-values, suggesting that larger samples may be required to observe these relationships in detail. Given that these relationships were only seen for the 13-16-year-olds but not the 18-24-year-olds, there may be some developmental changes in the relationships between poor sleep and positive affect, potential related to brain maturation, which are discussed in relation to the other mediators in later sections.

Together, this suggests that positive affect may be more closely related to sleep in younger adolescents than in older adolescents. This may explain why depression, which is characterised by low positive affect, showed a direct association with sleep in the younger sample, whilst anxiety and stress did not show a direct relationship. Sleep was directly related to mental health in older adolescence but appears to require mediation through negative affect

for younger adolescents (for anxiety and stress). A review of the literature suggested that the majority of research examining sleep and emotion regulation often does not separate out the different strategies or cognitive processes underlying the relationships (Palmer & Alfano, 2017). This may particularly be the case, in that positive affect may use emotion regulation strategies aimed at up-regulating and maintaining positive affect whilst emotion regulation of negative affect may focus on down-regulation or inhibition (Carl et al., 2013). Yet positive affect is typically understudied (A. D. Ong et al., 2017; Short et al., 2020). As such, further investigation into the strategies employed across adolescent development to regulate both positive and negative affect may be useful.

Given that depression is characterised by excessive REM sleep (e.g., Goldstein & Walker, 2014), and low positive affect in the present model and tripartite theory, perhaps excessive REM sleep reduces positive affect to a greater extent in younger adolescents. Goldstein and Walker's (2014) framework suggests that REM sleep recalibrates the emotional system in preparation for the next day, as well as reducing the emotional strength of memories and experiences from the previous day (e.g., Sleeping to Forget and Sleeping to Remember model). Typically, this research focuses on aversive emotions, or insufficient sleep duration. The present thesis highlights the need for future research to include positive affect and items with positive valence (e.g., Chapter 4), as well as considering developmental impacts during adolescence (e.g., prefrontal cortex development), and the impact of REM sleep on these relationships.

5.5.2. Inhibition and emotion regulation

The present study found that neither emotion recognition, inhibition of responses to emotional stimuli (affective control), nor speed/accuracy trade-off were significant mediators of the relationship between habitual sleep and mental health symptoms. This is contrary to

predictions made by Goldstein and Walker (2014) that the relationship between poor sleep and depression would be associated with lower discriminability between affective and neutral stimuli. It is worth noting that the Go/No-Go task may not have captured emotion recognition through the d' measure. Instead, the d' in this task may be better interpreted as an emotion detection measure. The Go/No-Go task requires participants to discriminate between neutral and emotive stimuli, rather than recognising differences between different emotions (e.g., scared versus surprised). However, even though d' may capture discriminability rather than emotion recognition in the current study, this measure may be sufficient to test Goldstein and Walker's (2014) hypothesis. Their theory posits that poor sleep is associated with a lower ability to effectively identify neutral stimuli as non-salient. As such, this lower discriminability between neutral and affective stimuli contributes toward the negativity bias in depression, in that people with depression are thought to have a tendency to interpret neutral stimuli as negative. Therefore, according to this hypothesis, the relationship between poorer sleep and higher levels of depression should be mediated by lower ability to discriminate between neutral and negative (e.g., sad, angry) facial expressions, due to a greater likelihood of misinterpreting neutral faces as negatively emotionally salient.

Goldstein and Walker (2014) further proposed that the relationship between poor sleep and anxiety would be associated with lower ability to inhibit attention and reactivity towards fearful stimuli. However, the Go/No-Go task requires participants to be able to both detect the presence of an emotion as well as inhibit behavioural responses, which makes it challenging to separate the influence of emotion detection/recognition or emotion regulation rather than behavioural inhibition. Instead, this paradigm may capture cognitive control, through the ability to inhibit behavioural responses towards stimuli, which may be independent of the influence of affective stimuli. Whilst the overall false alarm rate across all trials was taken as a measure of 'cognitive control' (considered to be a proxy for individual

differences in response inhibition generally), future research may benefit from using an affectively neutral task. For example, a separate Go/No-Go task involving neutral stimuli such as shapes, colours, or perhaps facial identity to measure individual differences in inhibitory control without the confound of the ability to discriminate between neutral and emotive facial expressions. Cognitive control generally shows protracted development throughout adolescence and may contribute toward emotion regulation abilities (see Schweizer et al. [2020] for a review). However, future research investigating emotion regulation may wish to examine other cognitive processes that may be involved in the association between cognitive control and emotion regulation, such as measuring affective control through the ability to shift attention and update working memory in relation to affective stimuli. Physiological responses to emotional stimuli may also be informative of the relationship between sleep, attention and emotion regulation, and mental health. Therefore, future research may wish to include galvanic skin responses, heart rate monitors, or other devices to measure physiological responses to emotional stimuli. Unfortunately, such measures were not practical during an online study with limited time and no physical contact with participants.

The present study is also limited by the use of self-reported global sleep quality, which does not directly examine the influence of REM sleep posited by Goldstein and Walker (2014). Unfortunately, it was not possible to examine sleep architectural and neural processes, due to COVID-19 lockdowns, requiring non-contact (e.g., online) research at the time. Despite the lack of significant mediation in the current study, poorer habitual sleep was directly associated with more false alarms (lower affective control) for afraid and happy faces in the 13-16-year-old model only. Higher d' (greater recognition) for sad faces was also associated with greater levels of stress for this age group only. Whereas fewer false alarms (higher affective control) towards afraid faces was associated with higher levels of depression

for 18-24-year-olds only. In the 18-24-year-old group, the association between affective control for afraid faces and depression may be associated with participants that have comorbid depression and anxiety. It is worth noting that these relationships were only seen in the habitual sleep model. Participants identified as having higher levels of depression may also experience anxiety and so may have biases towards 'afraid' stimuli through the anxiety, masked by comorbid depression. The relationship between habitual sleep and affective control seen for younger participants (13-16-year-olds) could be developmental. For example, adolescents are found to be more risk-taking and have greater sensitivity to rewards when they have insufficient or poor-quality sleep (Telzer et al., 2013). Thus, poorer sleep may be associated with a lower ability to regulate reactivity towards rewarding (e.g., happy) and risky (e.g., fearful) stimuli. In the last night's sleep model, both the older and younger groups showed a significant relationship between sleep and speed/accuracy toward sad stimuli, in that participants with poor sleep required slower reaction times to respond accurately for sad faces. Perhaps poor sleep on the *preceding night only* may affect abilities to regulate and inhibit responses to sad stimuli accurately. Therefore, participants with poor sleep on the previous night may be more reactive to sad stimuli the next day. These suggestions are speculative and require further investigation due to limitations with the current experimental paradigm such as the use of low-level affective stimuli. Future research with rewarding, risky, or more emotionally salient stimuli may have greater sensitivity and indicate different relationships.

Individual emotions were included in separate constructs as this allowed for the different relationships associated with positive (e.g., happy) and negative (e.g., sad, angry, afraid) affect, and different mental health disorders (e.g., anxiety and fear). Originally, it was planned that higher-order latent constructs would be used, which may have addressed some of the discriminant validity issues. Unfortunately, specifying higher order latent constructs

was not possible within the semPLS package. Another R package, SEMinR (Ray et al., 2022), was recently released, which allows specification of higher order constructs. However, SEMinR conducts consistent PLS-SEM analyses that would not converge for both age groups in the current reflective models. Therefore, standard PLS-SEM was used in semPLS and SmartPLS. The Go/No-Go paradigm was also limited by the stimuli available. Facial expression stimuli were taken from the KDEF repository (www.kdef.se). All the faces in the database at the time were young, White participants, and therefore lacked representativeness. However, our samples were also predominantly White and all of them were young (13-24-years-old). Facial expressions may be less emotionally salient than other stimuli, which made this stimuli more ethical for online research. Future laboratory-based studies could use more emotionally salient stimuli (e.g., international affective picture system; Lang et al., 1997) to elicit greater emotional responses from participants, creating wider variability in the data and allowing more sensitive analyses.

5.5.3. Thought control

The Think/No-Think task is an experimental paradigm that works similarly to the Go/No-Go task often using distressing images. Using this task with distressing images would not have been ethical or appropriate for this online study. Including a second experimental task would also be more time consuming for an online study, and therefore is better suited to a laboratory study. The Think/No-Think paradigm has been used with word-pairs with children, and shows improved thought control ability associated with development (Paz-Alonso et al., 2009). Potential developmental changes may be associated with brain maturation. Given that the Thought Control Ability Questionnaire has been associated with performance on Think/No-Think paradigms (Feliu-Soler et al., 2019) it is likely that the results of the present study would also be found under laboratory experimental conditions.

Harrington and Cairney (2021) proposed that poor sleep results in lower thought control ability and greater levels of mental health symptoms. The present study found that lower scores on the Thought Control Ability Questionnaire (TCAQ) significantly mediated the relationship between poor habitual sleep and greater mental health difficulties in both age groups, supporting Harrington and Cairney's (2021) framework. Thought control also acted as a serial mediator with negative affect between habitual sleep and mental health for both age groups, but with positive affect only for 13-16-year-olds (as discussed in Section 5.5.3.). The original framework proposed by Harrington and Cairney (2021) does not directly consider the role of positive and negative affect, instead they consider 'affective disturbances' in relation to thought intrusions as a mechanism further disrupting sleep. Considering positive and negative affect is a novel inclusion in the current study and appears to show developmental differences between the age groups, which may be associated with maturation of the neural networks and cortical areas highlighted in Harrington and Cairney's (2021) framework.

Harrington and Cairney (2021) suggested that the right dorsolateral prefrontal cortex is responsible for inhibiting hippocampal retrieval of intrusive thoughts and the arousal response of the amygdala toward these thoughts. Whereas Goldstein and Walker (2014) suggest a role for the medial prefrontal cortex in regulating amygdala reactivity, including good sleep supporting the connectivity between the amygdala and ventromedial prefrontal cortex. Recent research shows that functional connectivity between the hippocampus and prefrontal cortex demonstrates protracted development throughout adolescence (Calabro et al., 2020). Specifically, the ventromedial (perhaps important for regulating the amygdala), but not dorsolateral (perhaps important for regulating the hippocampus), prefrontal cortex shows increased functional connectivity throughout adolescent development. Additionally, early life stress (e.g., in infancy) has been associated with wider connectivity between

amygdala, hippocampal and prefrontal cortex networks during adolescence (Silvers et al., 2016). Silvers et al. (2016) also found that a stronger connection between the ventromedial prefrontal cortex and the hippocampus was a significant predictor of reduced anxiety in the future. This suggests the ventromedial prefrontal cortex may regulate the hippocampus in regard to anxious thoughts in the context of early life stressors. Given that the present study observed relationships between sleep and thought control in both age groups, the dorsolateral prefrontal cortex may be involved (as suggested by Harrington and Cairney, 2021), as the dorsolateral prefrontal cortex does not show such large changes in connectivity with the hippocampus during adolescence (Calabro et al., 2020). However, the connectivity between the prefrontal cortex, amygdala, and hippocampus may be altered when considering early life stressors (e.g., Silvers et al., 2016). Future research spanning the course of adolescent development (10-24-years-old) and including neural imaging may provide more information regarding developmental changes in these relationships. Suggestions for future research are discussed more broadly in the section below (Section 5.5.4).

5.5.4. Future research

As with any cross-sectional research, including the present study, direction and causality cannot be established. Experimental or longitudinal designs could be employed by future research to examine whether the relationships identified here are directional or causal. Longitudinal protocols may be able to explore the direction of the relationship between sleep, affect, and mental health, in that longitudinal designs can examine whether poor sleep precedes poor mental health, which has been explored to some degree in previous research (e.g., Alvaro et al., 2013; Baglioni et al., 2011). Whilst naturalistic longitudinal studies (e.g., Lev Ari & Shulman, 2012) are more sensitive to directionality, they are less controlled than experimental studies that may be better able to inform about causal mediators and

mechanisms. Instead, combined experimental and longitudinal designs (e.g., in intervention studies) may be better able to inform about directionality and causality. Experimental studies could manipulate sleep restriction (e.g., Baum et al., 2014). Examining sleep restriction would be preferred over total sleep deprivation, as sleep restriction paradigms can examine the cumulative influences of poor sleep, which is likely to be more naturalistic in regard to the sleep issues adolescents experience (e.g., discussed in Chapter 2). Future research could also consider night-to-night changes in sleep over a longitudinal design. Intraindividual night-to-night sleep variability may be more common during adolescence than childhood (Becker et al., 2017). One study found that night-to-night changes in sleep quality and duration were directionally associated with next-day suicidal ideation (Littlewood et al., 2019), which can be considered an intrusive thought. As such, ability to control thoughts may vary in a dose-response relationship with sleep. This may further be influenced by positive and negative affect, based on the relationships established in the current study.

Intrusive thoughts could also be measured experimentally, for example using Think/No-Think paradigms, which have been correlated with reports on the Thought Control Ability Questionnaire used in the present study (Feliu-Soler et al., 2019). However, intrusive thoughts, such as suicidal ideation, are very personal to an individual and may not be effectively captured in general Think/No-Think experimental paradigms. Instead, one study used an experimental design that implemented personal intrusive thoughts and neutral thoughts, and found that personal intrusive thoughts were much harder to inhibit and replace than neutral thoughts (Ólafsson et al., 2014). Using a personally relevant design could be more informative of the relationships associated with sleep and the development of mental health issues. Such a personally salient and potentially distressing design would not have been ethical in an online study where participant wellbeing cannot be monitored effectively, and would be more appropriate for laboratory-based research in the future.

In addition to thought control, the relationships investigated in Chapters 4 (emotional memory biases) and 5 could be examined in people with depression, to add to theoretical frameworks such as the Affect Tagging and Consolidation model (ATaC; Harrington et al., 2017). A better understanding of the theoretical underpinnings of the relationships between sleep and mental health, and the impact of adolescent development on these relationships, could better inform interventions. The majority of therapeutic interventions for adolescents experiencing mental health difficulties involve cognitive behavioural therapy (CBT) as recommended by NICE guidance (e.g., Donnellan et al., 2013). Part of CBT involves cognitive strategies to challenge unhelpful thought, beliefs, and attitudes, and CBT can also be used to treat sleep issues (Cliffe et al., 2020; de Zambotti et al., 2018; Manber et al., 2008). Cognitive strategies (e.g., reappraisal) require thought control abilities (Engen & Anderson, 2018), which the present study has found to be associated with sleep, affect, and mental health. Therefore, the present research suggests that young people who are sleeping poorly may be less able to regulate their emotions via cognitive strategies associated with thought control, resulting in greater mental health issues. As such, the techniques employed in CBT may be less effective when adolescents are not sleeping well. A review of the literature suggests that maturational changes in the connectivity between the prefrontal cortex, amygdala and hippocampus during adolescent development may explain difficulties in treating anxiety disorders and the high rates of recurrence of symptoms after treatment for this group (Zimmermann et al., 2019). However, the review did not consider the impact of sleep and theories regarding the relationship with sleep do not fully consider the impact of adolescent development. As such, future research combining both development and sleep would help to inform theoretical models and interventions to support adolescent sleep and mental health.

5.5.5. Conclusions

In conclusion, the present study found that poor sleep was associated with greater levels of mental health symptoms via negative affect and thought control abilities. This supports Harrington and Cairney's (2021) model. When thought control was considered alongside the tripartite theory of mental health, negative affect acted as a serial mediator with thought control in the relationship between sleep and mental health. Positive affect appears to be important in the relationship between sleep and depression (including as a serial mediator with thought control) for younger, but not older adolescents, suggesting developmental influence within these relationships. Despite predictions made by Goldstein and Walker (2014), the current study did not find that reactivity toward emotional stimuli was a significant mediator between sleep and mental health. However, this may be due to methodological constraints associated with conducting online research. Future research investigating the impact of adolescent development on the relationships between poor sleep and mental health would help to inform the current theoretical models (e.g., Goldstein & Walker, 2014; Harrington & Cairney, 2021). Better understanding of the mechanisms between these relationships would also help to inform mental health and sleep interventions, which may need to be tailored specifically to different stages of adolescent development. The treatment of sleep issues in the Child and Adolescent Mental Health Services (CAMHS) is explored in the next chapter (Chapter 6).

Chapter 6: Sleep problems and sleep medication prescription in the child and adolescent mental health services (CAMHS)

6.1. Introduction

6.1.1. The Child and Adolescent Mental Health Services (CAMHS)

Children and adolescents attending CAMHS are experiencing psychological distress requiring intervention. CAMHS patients are more likely to have greater severity in internalising and externalising symptoms of mental health difficulties than patients attending other services such as community paediatrics (Roongpraiwan et al., 2007). When young people attend CAMHS for an initial assessment of their difficulties, best practice is that they complete Routine Outcome Measures (also known as ROMs), which are questionnaires such as the Revised Children's Anxiety and Depression Scale (RCADS; Chorpita et al., 2000). These questionnaires can facilitate therapeutic engagement and provide information about symptoms that may require intervention. CAMHS patients may have a variety of mental health difficulties, including depression, anxiety, obsessive-compulsive disorder, eating disorders, and psychosis. They may also experience transdiagnostic difficulties such as self-harm, school-avoidance, and sleep issues.

As seen in Chapter 5, poor sleep is associated with higher levels of depression, anxiety, and stress. Almost all mental health conditions have comorbid sleep issues (Baglioni et al., 2016) and young people with insomnia often have comorbid psychiatric disorders (e.g., depression, anxiety) or neurodevelopmental (e.g., Autism, ADHD) conditions (Nunes & Bruni, 2015). Poor adolescent sleep has been associated with increased negative affect (e.g., Chapter 5) and reductions in emotion regulation abilities (Baum et al., 2014). Poor sleep has also been associated with an increased likelihood of adolescents engaging in self-harm (McGlinchey et al., 2017), increased recurrence of self-harm (Asarnow et al., 2020), daily

fluctuations in suicidal ideation (Littlewood et al., 2019) and adolescent psychiatric hospitalisation (Borschmann et al., 2018). The relationship between sleep and suicidality (e.g., self-harm) may be associated with reduced ability to inhibit difficult thoughts (e.g., suicidal ideation) in the presence of poor sleep (e.g., Harrington & Cairney, 2021; Chapter 5). Therefore, a high proportion of CAMHS patients may be experiencing sleep issues, but not everyone in CAMHS will have trouble sleeping.

6.1.2. Predictors of sleep issues

It is important to identify which sociodemographic and service-use factors may be associated with sleep issues. Mental health diagnosis may be a significant predictor of reported sleep issues. Some mental health disorders are characterised by sleep disturbances that form part of their diagnostic criteria (e.g., depression, anxiety, post-traumatic stress disorder; DSM-5, 2013). These mood and anxiety disorders are often the focus of theoretical models of sleep and mental health (e.g., Goldstein & Walker, 2014, Harrington & Cairney, 2021). For example, Goldstein and Walker's (2014) framework suggests that REM sleep functions to reset emotion regulation systems in the brain. Without sufficient REM sleep, excessive release of noradrenaline occurs. This can cause the sensation of imminent danger and hypervigilance (e.g., anxiety, post-traumatic stress disorder). With excessive REM sleep, likely at the expense of other sleep stages and quality of sleep, noradrenaline is much lower (e.g., depression). In both instances, the prefrontal cortex becomes less able to inhibit amygdala reactivity. Excessive amygdala reactivity results in neutral stimuli being perceived as emotionally salient, in particular, with a negative bias (e.g., in depression). This negative emotional memory bias was discussed and explored in Chapter 4. Harrington and Cairney's (2021) framework further considers the cyclical relationship between poor sleep and intrusive thoughts, as a result of poorer inhibition over the hippocampus and amygdala when sleep is

poor. The role of sleep, thought control and mental health was explored in Chapter 5. These frameworks suggest poor sleep is a fundamental factor in the development of mood disorders such as depression, anxiety, and post-traumatic stress disorder. Despite this, the majority of paediatric sleep medication literature (discussed in Section 6.1.3.) focuses primarily on neurodivergence (e.g., ADHD and autism). However, a meta-analysis of controlled polysomnography research suggests that there are not significant alterations in sleep architecture for people with ADHD; but there are sleep alterations evident in a range of other disorders including autism, schizophrenia, depression, anxiety and emotional disorders (Baglioni et al., 2016).

As discussed in Section 6.1.1. above, sleep issues may also be associated with trans-diagnostic symptoms of mental health. Sleep issues have been found to precede adolescent psychiatric inpatient admissions, suggesting poor sleep is a risk factor for inpatient admission (Borschmann et al., 2018). Poor quality sleep and short sleep duration has been found to directionally predict next day suicidal ideation (Littlewood et al., 2019). Further, two-thirds of community CAMHS patients with mood disorders that engaged in self-harming experienced significant regular sleep disturbances (McGlinchey et al., 2017). Therefore, sleep issues may be associated with greater risk (e.g., suicidality) and more complex mental health presentations, potentially resulting in greater levels of service-use. This could include psychiatric inpatient admission or the amount of time spent in CAMHS. Therefore, it is important to understand the characteristics of CAMHS patients that may be more vulnerable to poor sleep.

Adolescent insomnia has been found to be most prevalent in girls and older adolescents (de Zambotti et al., 2018). The relationship between sex, age and sleep issues may be related to pubertal development influencing circadian rhythms. Therefore, mental health issues that are associated with poorer sleep (e.g., depression) may be more common in girls, and older

adolescents in CAMHS. Pubertal development is typically earlier in girls than boys and advances with chronological age (Marshall & Tanner, 1969). Across mammalian species puberty is associated with delayed circadian rhythms (Hummer & Lee, 2016). Specifically, girls' pubertal development is associated with greater eveningness chronotype preferences (Carskadon et al., 1993). Thus, older adolescents and girls will likely be further into puberty and more likely to experience short sleep duration as a function of delayed circadian rhythms. Additionally, increasing adolescent age is associated with later bedtimes and greater likelihood of experiencing week-day sleep restriction, and thus suffering from insufficient sleep durations (Gradisar et al., 2011), which may be a function of greater independence over bedtimes for example (Tashjian et al., 2019). Increased parental restrictions preventing later bedtimes has been found to act as a protective factor preventing the likelihood of a young person developing depression (Gangwisch et al., 2010).

Social factors such as socioeconomic status may also be associated with sleep issues. For adolescents from a low socioeconomic background, sleep duration and timing has been found to be less regular than for adolescents from higher socioeconomic backgrounds (Marco et al., 2011). This may be related to environmental factors such as a higher likelihood of people from lower socioeconomic status sharing a bedroom (Buckhalt et al., 2007). Additionally, pre-sleep worries in children from lower socioeconomic backgrounds has been associated with greater time awake at night and more sleep disruptions (Bagley et al., 2015). The relationship between poor sleep and pre-sleep worries may be explained by inability to inhibit worrying thoughts before sleep, which may further exacerbate poor sleep cyclically (Harrington & Cairney, 2021). Therefore, socioeconomic status may be associated with sleep issues in CAMHS samples.

Ethnicity is another sociodemographic variable considered in the present study. In the USA, one study has found participants of Hispanic, African American, or Other ethnicities

were more likely to report insufficient weekday sleep duration compared to White participants (Perlus et al., 2018). Another study in the USA found that African American students reported greater excessive daytime sleepiness than White or Latinx students (Edens, 2006). Racial discrimination has been associated with the sleep of young adults of Mexican ethnicity living in the USA in different ways depending on their cultural orientation (e.g., bicultural, marginalised; Zeiders et al., 2017). Together, the research in the USA suggests lower levels of sleep issues for White people and higher levels for people of other ethnicities, particularly for the people who are more discriminated against and marginalised. In a large sample of ~96,000 adult participants in the UK, actigraphy monitored sleep was also found to be of lower quality for people of minority ethnicities (e.g., Black, South Asian) compared to White adults (Dawkins et al., 2022).

Given the implications of poor sleep for mental health (e.g., Chapter 5), identifying the characteristics of CAMHS patients experiencing sleep difficulties is important. Therefore, our first research question, aims to identify the demographic and service-use variables that are associated with reporting trouble sleeping, these include gender, age, socioeconomic status, ethnicity, primary mental health diagnosis. active days in South London and Maudsley CAMHS, psychiatric inpatient admissions, and reason for being discharged from CAMHS.

6.1.3. Sleep treatment

Based on Sections 6.1.1. and 6.1.2. above, it is likely that many CAMHS patients experience sleep issues. However, not all of the CAMHS patients experiencing sleep issues receive sleep medication prescriptions. In the previous section, we discussed the demographic and service-use variables that may be associated with sleep issues in CAMHS populations (the first research question). The second research question of this study aims to investigate which of these variables are associated with sleep medication prescription in CAMHS.

Psychological and behavioural interventions can be used to treat sleep problems (e.g., sleep hygiene, cognitive behavioural therapy (CBT) for insomnia). CBT for insomnia is found to also reduce depression symptoms and support the efficacy of other implemented mental health treatments (Manber et al., 2008). Meanwhile, treating adolescent depression with psychological therapy is associated with some decreases in sleep issues (Reynolds et al., 2020). However, in Reynolds et al.'s (2020) study approximately half of adolescents continued to experience residual sleep issues after depression was treated. In a recent service evaluation study of CAMHS, it was found that approximately three-quarters of young people that were prescribed melatonin by CAMHS had parent-led behavioural sleep interventions (e.g., sleep hygiene) prior to commencement of sleep medication (Speedy et al., 2021). However, psychological sleep interventions (e.g., sleep hygiene) can lack consistency and standardisation in their delivery (Irish et al., 2015). This makes non-pharmacological interventions difficult to evaluate on a large scale in secondary data. Additionally, studies of pharmacological interventions for sleep issues in CAMHS populations are particularly sparse (Ramtekkar & Ivanenko, 2015). Therefore, pharmacological sleep intervention is the focus of the current study.

Sleep medication may be prescribed at higher rates to certain groups compared to others. Approximately 50% of Swedish adolescents receiving a prescription of melatonin had a diagnosed mental health or behavioural disorder, with ADHD being the most common (Kimland et al., 2020). This suggests that sleep medication prescription is prevalent in health and social care settings, where mental health or behavioural disorder diagnoses are given, such as in CAMHS. Paediatricians who treated patients with ADHD exhibiting sleep issues were four times more likely to prescribe sleep medication than paediatricians not treating this patient group (Owens et al., 2003). Additionally, for autistic young people (aged 4-20-years-old) admitted to psychiatric inpatient hospitals, sleep medications were the third most

prescribed medication group following anti-psychotics and ADHD medication (Wink et al., 2018). As sleep problems are common in people with ADHD or autism this may explain high prescription rates for these groups (Rzepka-Migut & Paprocka, 2020). However, the prescription of sleep medication may also be given to counteract side effects (i.e., insomnia, sleep issues) of stimulant medications (e.g., ADHD medications). Despite the high prescription rates of sleep medication for young people with ADHD, a systematic review concluded there was still a lack of sufficient evidence regarding sleep medication prescription even within this group (Barrett et al., 2013). Therefore, even in populations with a relatively high rate of sleep medication prescription, still very little is known about the use of sleep medication for children and adolescents.

In a Swedish population study, boys were found to be more likely than girls to be prescribed melatonin sleep medication (Kimland et al., 2020), suggesting there may be sex biases in sleep medication prescription. In a nationwide study conducted in New Zealand with data from 0-18-year-olds, autistic children of European ethnicity were significantly more likely to be prescribed melatonin than autistic children of Asian, Middle Eastern, Latinx, and Pasifika ethnicity (McLay et al., 2021). This suggests that ethnicity may also be predictive of sleep medication prescription. In adult samples, the prescription of benzodiazepines and z-drugs (types of sleep medication) was significantly higher in primary care NHS practices from low socioeconomic areas than from higher socioeconomic areas (Soyombo et al., 2020). However, benzodiazepines were not examined solely for their sedative effect and may be used to treat anxiety or other symptoms too. As the current study used data from South London and Maudsley CAMHS in particular, we considered the role of individual socioeconomic status in its relationship with sleep medication prescription.

6.1.4. Research questions

Sleep issues are likely a common problem within the CAMHS population as poor sleep is associated with mental health problems and sleep changes during adolescence, making young people particularly vulnerable to this relationship (as discussed in Section 6.1.1.). Such sleep issues may be associated with sociodemographic and service-use characteristics such as socioeconomic status, ethnicity, sex, and mental health diagnoses (see Section 6.1.2.). These variables may further be associated with sleep medication prescriptions (as explored in Section 6.1.3.).

The current exploratory study uses a large secondary NHS dataset to answer the following research questions:

1) Within South London and Maudsley CAMHS, what are the characteristics of patients reporting sleep issues? (Based on the item "I have trouble sleeping" from the routinely collected RCADS)
2) What are the characteristics of CAMHS patients receiving pharmacological intervention for sleep issues?

As the models for both research questions share the same variables, a narrative comparison of the two models is discussed in section 6.4.3.

6.2. Methods

6.2.1. Study population

A sample of 7,167 patients were extracted from the Clinical Records Interactive Search (CRIS) based on a range of inclusion criteria outlined in Section 6.2.1.2. Demographics of the extracted sample are reported below in Tables 6.1. and 6.2. Ethical approval for the use of CRIS data for research was granted by South Central – Oxford C Research Ethics Committee (current reference: 23/SD/0257). Consent for the use of the data in the current project (20-033) was granted through NHS information governance procedures.

Table 6.1. Frequency demographics for the extracted sample

Characteristic	Number of patients (N =)	Proportion (%) of whole sample (N = 7167)
Ethnicity		
White – British	2720	37.95
White – Other	489	6.82
Black/Black British – Caribbean	373	5.20
Black/Black British African	443	6.18
Black/Black British – Other	1017	14.19
Mixed race – White & Black African	109	1.52
Mixed race – White & Black Caribbean	435	6.07
Mixed race – Other	278	3.88
Asian/Asian British	298	4.16
Other ethnic group	227	3.17
NULL	778	10.86
Sex		
Female	3683	51.39
Male	3468	48.39
NA/Other	16	0.22
Primary mental health diagnosis		
ADHD or hyperkinetic	657	9.17
Anxiety, stress, and emotional disorders	2016	28.13
Autism or developmental disorders	706	9.85
Depression	627	8.75
Obsessive-compulsive	158	2.20
Oppositional and conduct	366	5.11
Psychosis	49	0.68
Other disorders	915	12.77
No diagnosis	1673	23.34
Inpatient admissions		
None	7000	97.67
At least one	167	2.33
End reason		
Turned 18-years-old	932	13.00
Discharged	4986	69.57
Still in services	1247	17.40
Age group		
12+ years-old	4610	64.32
Under 12-years-old	2557	35.68

Note: Characteristics in bold are used as the reference condition in all analyses

Table 6.2. Descriptive statistics for continuous variables for the extracted sample

Characteristic	Mean	Standard deviation	Min	Max	Median	Missing
SES	28.65	10.89	1.77	73.26	29.39	73
Active days in SLaM CAMHS	650.5	620.02	1	4242	466	11*
Age at index	12.74	2.87	3.46	18.00	13.28	0

Note: *The missing data for active days were patients with no recorded active days (assigned NA, in order to calculate the log). SLaM = South London and Maudsley. Min = Minimum, Max = Maximum

6.2.1.1. Data source

The Clinical Records Interactive Search (CRIS) is a large dataset of clinical records for patients in South London and Maudsley mental health services. CRIS CAMHS data has been used to examine predictors of antipsychotic prescriptions for autistic patients (Downs et al., 2016) and has been joined with other large naturalistic datasets such as the National Pupil's Database (Wickersham et al., 2020) and social care records (Downs et al., 2019). Therefore, this dataset provides a large, rich data source adept at examining many research questions in CAMHS populations.

6.2.1.2. Inclusion criteria

The overall window start date began in 2008 and ended on 1st January 2019 to allow a sufficient opportunity for a twelve-month follow-up, with outputs extending to 1st January 2020. Only face-to-face patients were included, to ensure that referrals had been accepted by the services. However, type of contact (i.e., face-to-face) was not reported as mandatory before 2010, and therefore there might have been some artificially missing data for patients in the services prior to 2010. Therefore, for data before 2010, the RCADS could be conducted before the individual window (i.e., before the first face-to-face session), but from 2010 onwards, the RCADS must be on or after the individual window (i.e., after the first face-to-

face session). Only the 47 item (full) RCADS has been inputted into CRIS, not the reduced versions (e.g., the 25-item version described in Chapter 5, Section 5.2.2.4.). Item 11 ("I have trouble sleeping") on the full RCADS must have been completed for patient data to be extracted for this study, but any other item could be missing. This meant that all the extracted data included information about (self or caregiver) reported sleep issues from the RCADS, but that data about symptom profiles may be incomplete. Inpatients were excluded at the window start date to ensure that all patients were local to the boroughs that South London and Maudsley covers, and not inpatients from external areas, to increase the likelihood of having sufficient data for each patient. If a patient was an inpatient in their first contact with South London and Maudsley CAMHS, these patients would not be included in the sample, and would likely be out-of-area patients rather than local to the South London and Maudsley boroughs. The first ever referral within this window was used in the hopes of understanding sleep profiles and sleep treatments as close to the beginning of service use as possible. Patients were included if they were 16-years-old or younger at the beginning to ensure sufficient time for follow-up before aging out into adult services. The end dates were determined by either reaching adult services (18th birthday), discharged and case closed (with no return), or still in the services on 1st January 2020 when data were extracted. "No return" was included within the discharge criteria to avoid missing data from repeat patients.

6.2.3. Measures

6.2.3.1. Sample characteristics

Socioeconomic status was extracted as an Index of Multiple Deprivation across numerous years. The year of the Index of Multiple Deprivation was selected based on the year of index (first ever Revised Children's Anxiety and Depression Scale [RCADS] questionnaire completed). Lower scores indicate greater levels of socioeconomic deprivation.

Sex and ethnicity data were also extracted. Ethnic groups with fewer than 100 patients were collapsed into larger ethnic groups (e.g., "White Irish" was grouped with "White – Other"). All other ethnic groups with samples larger than 100 patients remained. Age was calculated based on the difference in time between the patient's date of birth and index date, creating "age at index". Primary and secondary diagnoses were extracted from the diagnostic field and could occur outside of the window, given that many conditions likely existed during their time in CAMHS, even if they were not diagnosed until a later date (e.g., ADHD, Autism). Diagnoses were selected that were closest to the index date. Only primary mental health diagnoses were included in the analyses. If the original primary diagnosis was a physical or environmental ICD-10 code, then this was replaced with a secondary diagnosis on the same date as the primary diagnosis, or another primary diagnosis after the initial diagnosis. If no mental health diagnoses were present, then primary mental health diagnosis was assigned NULL. Due to some small group sizes, diagnoses were grouped into the following categories: anxiety, stress, and emotional disorders; autism and developmental disorders; ADHD and hyperkinetic disorders; psychosis; depression; obsessive-compulsive disorders; conduct and oppositional disorders; and no diagnoses. All other diagnoses were grouped into "other diagnoses". These groupings are based broadly on previous research using CRIS data (Downs et al., 2016). The exact ICD-10 codes included in the diagnostic groups can be found in Appendix 6.3. Other items on the Revised Children's Anxiety and Depression Scale were used to characterise the symptom profile of the patients (e.g., social phobia, general anxiety disorder etc.). Service use data were also extracted, this included inpatient admissions, reason for leaving the services, and active days in the services. Patients with any inpatient admission to a psychiatric inpatient hospital during their time in the South London and Maudsley CAMHS were compared to patients with no psychiatric inpatient admissions. Reasons for leaving the services included turning 18-years-old, discharge (or dropped out) with no return,

or end of general window (1st January 2020). Active days in the South London and Maudsley CAMHS was the accumulated number of days a patient was registered in the services (e.g., excluding discharge or drop-out dates, but accounting for patients returning to the services after these periods). Active days in South London and Maudsley CAMHS was log-transformed using the "log()" function in R due to the large range (see Table 6.2.).

6.2.3.2. Sleep issues

Information about sleep is not routinely collected or recorded within the National Health Service (NHS). However, the Revised Children's Anxiety and Depression Scale (RCADS) is routinely collected in CAMHS. Within this questionnaire, a single item asked patients to respond from (0) "Never" to (4) "Always" to the statement "I have trouble sleeping". In this study, this item is used to indicate (broadly) whether the patient has 'any' (sometimes, often, always) or 'no' (never) sleep issues. Dichotomising the sleep variable allowed logistic regressions to be conducted that were comparable to the sleep medication prescription analyses explained below. As such, categorising this sleep variable allowed more coherent narrative comparison of the two models, to identify differences in characteristics of young people. This dichotomisation is similar to that used in other NHS data analyses (e.g., Table 1.7a, Lifestyles Team & NHS Digital, 2021). The RCADS could have been completed by the child or young person themselves, or by a caregiver. We extracted the earliest completed RCADS (index) to maximise sample numbers. However, informant discrepancies between self and caregiver report have been evidenced in the RCADS (Becker et al., 2019). Therefore, analyses were conducted to examine whether self or caregiver reports significantly differed from index reports. If so, analyses would be re-run with caregiver and self-report separately, and if not, analyses would be run with index reports only.

6.2.3.3. Sleep medication

Sleep medication was grouped into three main categories: melatonin (e.g., Circadin), promethazine (e.g., Phenergan) and z-drugs (e.g., zopiclone, zaleplon, and zolpidem). Benzodiazepines and other sleep medications were excluded as they are widely used to treat other disorders (e.g., anxiety) and are less specific to sleep treatment. Therefore, this would be difficult to validate within free-text mentions. Mentions of sleep medication were identified within the free-text or pharmacy data of CRIS. The prescription of sleep medication was validated on a patient level, rather than a mention level, in that multiple mentions per patient were examined to identify whether the mention of the sleep medication prescription met the following rules:

1) The medication must have been prescribed by a professional (including psychiatrist, GP, paediatrician); not purchased over the counter or taken from someone else (e.g., family).
2) The medication must be taken by the patient. In essence, it must appear that the patient actually used the medication and that it was not prescribed for someone else.
3) The mention must be current, not a historical mention of previous prescriptions outside of the current CAMHS.

For Promethazine, which has more uses than just for sleep treatment, the rules also included:

4) The medication must be mentioned in relation to sleep issues or swapping to or from other sleep medications, implying that the medication is taken to support sleep (rather than distress management, anxiety, or antihistamine properties).

Ambiguous mentions were scored by two researchers, and any disagreements were resolved through discussion until a final decision was made. The second researcher also

scored a random sample of 10% of the remaining mentions, which were compared to the original scorings from the first researcher. Agreement was 100% for melatonin, and 88.1% (Cohen's Kappa = 0.66) for promethazine suggesting substantial agreement for promethazine and perfect agreement for melatonin, therefore, no further validations were required. Z-drugs were not included independently due to a very small number of mentions, which could have resulted in patients becoming identifiable. However, to include z-drugs in the analyses, a variable of "any sleep medication" was created if a patient was prescribed either melatonin, promethazine, or any z-drug for sleep.

6.2.4. Statistical analyses

Two multivariable logistic regression models were created to examine independent predictors of any sleep issues or any sleep medication prescription. This follows a similar analysis to previous research examining antipsychotic medication prescription with CRIS data (Downs et al., 2016). Predictor variables included in both models were: sex (male or female), log-transformed active days in South London and Maudsley CAMHS, end reason (discharge versus turning 18), socioeconomic status, ethnicity, primary mental health diagnosis, inpatient (ever or never), age at index, and age group (under 12 versus 12+ years old). Sleep issues were included as a predictor of sleep medication prescription.

Crude odds ratios were calculated by entering each predictor individually into the models without any other predictors accounting for variance in the outcome. Adjusted odds ratios were taken from the multivariate analyses, which included the other predictor variables within the model. Adjusted p-values and adjusted coefficients are reported in the text. Adjusted and crude odds ratios and upper and lower confidence intervals are reported in subsequent tables. All analyses were conducted within R. There was only a small amount of data for 12-month follow-up, and prescription mentions were validated per patient rather than

per mention, which meant they were not time-bound; therefore, follow-up analyses were not conducted within this study.

6.3. Results

6.3.4. Characteristics of CAMHS patients reporting sleep issues

Based on the index (first ever) Revised Children's Anxiety and Depression Scale item "I have trouble sleeping", patients were grouped as having "any sleep issues" if they responded "sometimes", "often" or "always"; or "no sleep issues" if they responded "never". This allowed comparison of the two logistic regression models for reported sleep issues and prescribed sleep medication. The grouping of any sleep issues compared to no sleep issues, is similar to the groupings in other NHS research (see Table 1.7a in Lifestyles Team & NHS Digital, 2021).

Given the potential discrepancies between self and caregiver reports (Becker et al., 2019) analyses were conducted to examine whether index sleep issues significantly differed from each informant report. There was no significant difference between self-reported and index reported sleep issues, $t(12553) = 2.23$, $p = .026$ (Bonferroni corrected $p = .052$). There was also no significant difference between caregiver compared to index reported sleep issues, $t(12454) = -1.94$, $p = .053$. Therefore, we used index reported sleep issues in all further analyses, to maximise the sample size. See Table 6.3. below for the frequencies of reported sleep issues. The model predicting sleep issues was significantly better than an empty model with intercepts only, $\chi^2(30) = 509.43$, $p < .001$; log likelihood -3642.91. Adjusted and crude odds ratios and 95% confidence intervals are presented in Table 6.4. (below).

Table 6.3. Frequency of reporting sleep issues on the item "I have trouble sleeping"

Reported by	Reported sleep issues	Number of patients (N =)	Proportion of sample (%)
Index			
	Always	1616	22.55
	Often	1578	22.02
	Sometimes	2258	31.51
	Never (None)	1715	23.93
	Any	5452	76.07
Self			
	Always	1386	23.61
	Often	1315	22.40
	Sometimes	1866	31.79
	None (never)	1303	22.20
	Any	4567	77.80
	Missing	1297	
Caregiver			
	Always	1218	20.99
	Often	1296	22.33
	Sometimes	1849	31.86
	None (never)	1440	24.81
	Any	4363	75.19
	Missing	1364	

Table 6.4. Multivariate logistic regression model for the reporting of any (compared to no) sleep issues at index

Characteristic (reference)	OR (95% CI)	Unadjusted p-value	aOR (95% CI)	Adjusted p-value
Male (female)	**0.64 (0.57 – 0.71)**	**<.001**	**0.71 (0.63 - 0.80)**	**<.001**
Log active days in SLaM CAMHS	**1.26 (1.20 – 1.32)**	**<.001**	**1.20 (1.14 - 1.26)**	**<.001**
End Reason (Turned 18 years old)				
Discharged	**0.48 (0.39 – 0.58)**	**<.001**	**0.73 (0.58 - 0.90)**	**.005**
Still in services	**0.71 (0.56 – 0.89)**	**<.003**	0.91 (0.71 – 1.17)	.484
Socioeconomic status	**0.99 (0.98 – 0.99)**	**<.001**	**0.99 (0.99 - 1.00)**	**.019**
Inpatient admission ever (never)	**1.65 (1.11 – 2.55)**	**.019**	1.05 (0.68 - 1.68)	.841
Age at index	**1.07 (1.05 – 1.09)**	**<.001**	**1.06 (1.02 - 1.10)**	**.003**
Age group: Under 12 years (12+)	**0.77 (0.69 – 0.86)**	**<.001**	**1.34 (1.07 - 1.67)**	**.010**
Ethnicity (White British)				
Asian	0.96 (0.71 – 1.32)	.813	0.94 (0.69 - 1.31)	.723
White – Other	**0.70 (0.56 – 0.88)**	**.002**	**0.74 (0.58 - 0.94)**	**.012**
Mixed race – White & Black Caribbean	**0.74 (0.58 – 0.94)**	**.014**	**0.77 (0.60 - 0.99)**	**.041**
Mixed Race – White & Black African	0.75 (0.48 – 1.21)	.216	0.83 (0.52 - 1.35)	.425
Mixed Race – Other	0.93 (0.68 – 1.28)	.626	1.00 (0.73 - 1.39)	.994
Black/Black British - Caribbean	**0.38 (0.30 – 0.48)**	**<.001**	**0.43 (0.37 - 0.55)**	**<.001**
Black/Black British - African	**0.36 (0.29 – 0.44)**	**<.001**	**0.42 (0.34 - 0.52)**	**<.001**
Black/Black British – Other	**0.53 (0.45 – 0.63)**	**<.001**	**0.57 (0.48 - 0.67)**	**<.001**
Other ethnicity	0.75 (0.55 – 1.05)	.083	0.86 (0.62 - 1.22)	.394
Missing ethnicity	**0.63 (0.52 – 0.76)**	**<.001**	0.85 (0.69 - 1.03)	.098
Primary mental health diagnosis (ADHD/hyperkinetic)				
Anxiety, stress and emotional	**1.52 (1.23 – 1.87)**	**<.001**	**1.40 (1.11 - 1.75)**	**.004**
Autism or developmental	1.25 (0.97 – 1.61)	.079	1.29 (0.99 - 1.67)	.058
Depression	**3.06 (2.24 – 4.22)**	**<.001**	**2.61 (1.87 - 3.67)**	**<.001**
OCD	1.20 (0.74 – 1.67)	.658	0.96 (0.63 - 1.48)	.844
Oppositional/conduct	0.76 (0.57 – 1.01)	.059	0.85 (0.63 - 1.14)	.282
Psychosis	0.93 (0.49 – 1.86)	.826	0.69 (0.35 - 1.42)	.291
Other	0.98 (0.77 – 1.23)	.836	1.00 (0.77 - 1.26)	.920
None	**0.63 (0.52 – 0.78)**	**<.001**	**0.73 (0.59 - 0.91)**	**.006**

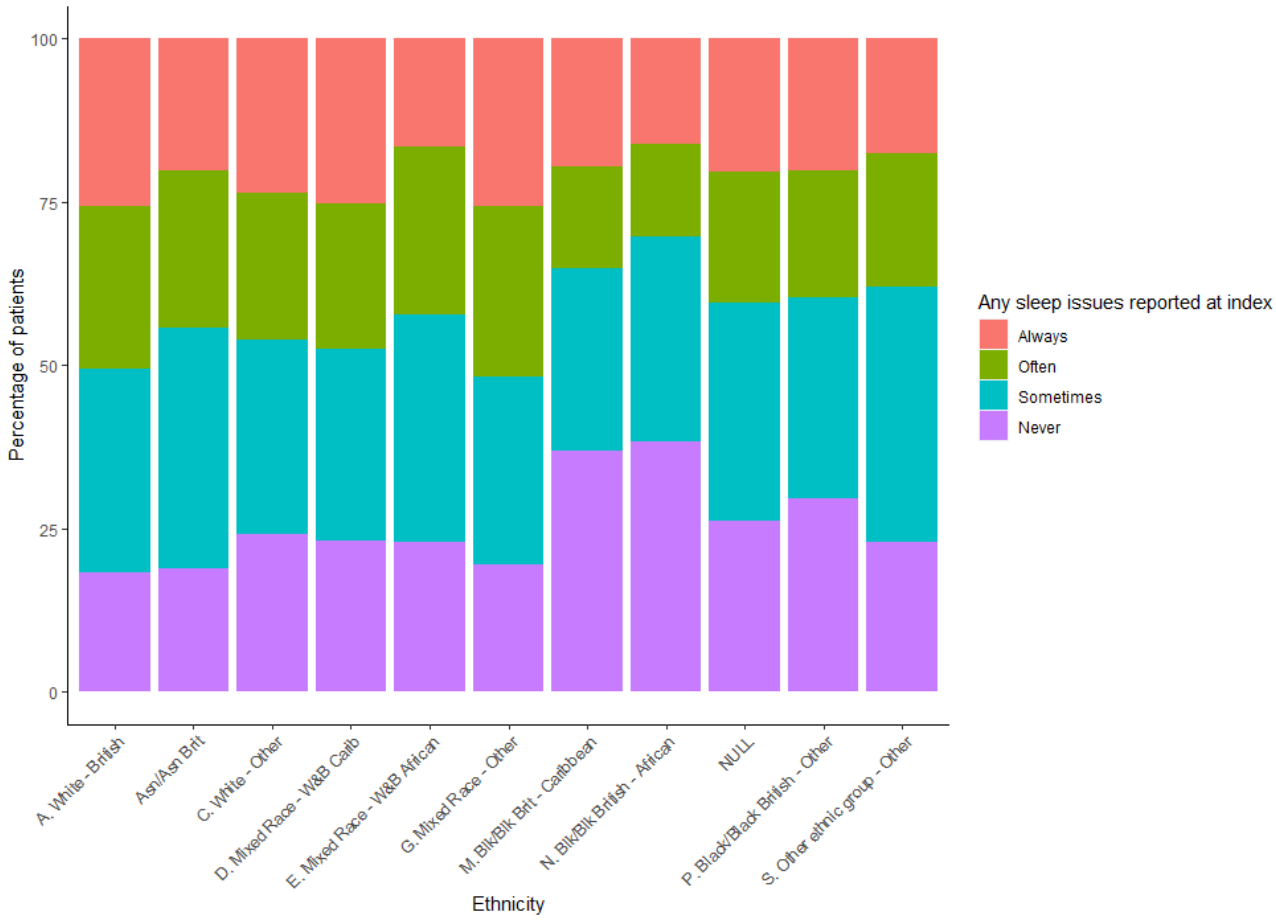
Note: This model uses index Revised Children's Anxiety and Depression Scales (RCADS) sleep issues. OR = odds ratios, aOR = adjusted odds ratio, CI = confidence interval. Significant relationships are in bold.

Girls were significantly more likely to report sleep issues than boys, $B = 0.34$, $p < .001$. Patients spending a greater amount of time in the South London and Maudsley services were significantly more likely to report sleep issues, $B = 0.18$, $p < .001$, than patients spending less time in the services. Patients with reported sleep issues were significantly more likely to age out of the services (turn 18-years-old), compared to patients that did not report sleep issues, $B = 0.32$, $p = .005$. This shows that patients reporting sleep issues were significantly less likely to reach successful discharge before turning 18-years-old, than patients who did not report sleep issues. Patients who were older when they completed the first Revised Children's Anxiety and Depression Scale were significantly more likely to report sleep issues than patients who were younger, $B = 0.06$, $p = .003$. As patients under 12-years-old may receive different services, particularly regarding sleep medication prescription, we also grouped age by under and over 12-year-old. Patients who were under 12-years-old were less likely to report sleep issues than patients aged 12-years-old and above, $B = 0.29$, $p = .010$. Socioeconomic status was also significantly associated with sleep issues, $B = -0.006$, $p = .019$ with lower socioeconomic status (lower Index of Multiple Deprivation score) associated with a greater likelihood of experiencing sleep issues.

Ethnicity was significantly associated with reporting sleep issues (see Figure 6.1. below). White British was the largest ethnic category in the sample and was therefore used as the reference condition. Black/Black British Caribbean, $B = -0.85$, $p < .001$, Black/Black British African, $B = -0.87$, $p < .001$, Black/Black British (other), $B = -0.57$, $p < .001$, Mixed race White and Black Caribbean, $B = -0.26$, $p = .041$, and White (other), $B = -0.31$, $p = .012$ patients were significantly less likely to report sleep issues compared to White British patients. There was not a significant difference in reporting of sleep issues for Asian, $B = -0.06$, $p = .723$, Mixed race White and Black African, $B = -0.19$, $p = .425$, Mixed race (other)

patients, B = -0.001, p = .994, and patients of other ethnicities, B = -0.15, p = .394, or with no reported ethnicity, B = -0.17, p = .098, compared to White British patients.

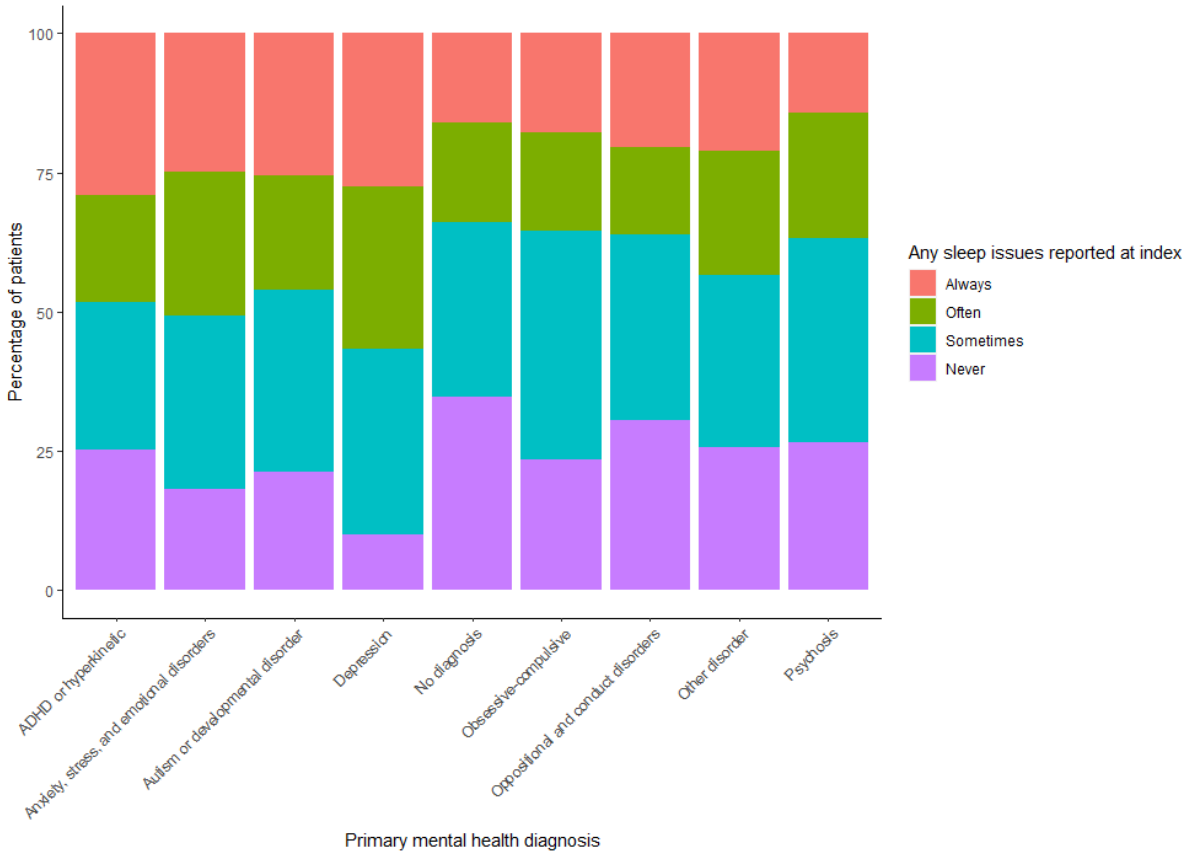
Figure 6.1. Percentage of patients with reported sleep issues at index per ethnicity group



Primary mental health diagnosis was significantly associated with reporting sleep issues. As previous research has suggested that patients with ADHD are likely to be prescribed sleep medication (e.g., Kimland et al., 2020), patients with ADHD (also known as hyperkinetic disorders in ICD-10) were included as the reference condition. Patients with anxiety, stress and emotional disorders, B = 0.34, p = .004, and depression, B = 0.96, p <.001, were significantly more likely to report sleep issues than patients with a diagnosis of ADHD (or hyperkinetic disorders). Patients with no diagnosis, B = -0.31, p = .006, were less

likely to report sleep issues than patients with ADHD. Patients with psychosis, B = -0.38, p = .291, oppositional and conduct disorders, B = -0.16, p = .282, obsessive-compulsive disorders, B = 0.04, p = .844, autism or developmental disorders, B = 0.25, p = .058, or other disorders, B = -0.01, p = .92, did not report significantly different levels of sleep disorders than patients with ADHD.

Figure 6.2. Percentage of patients with reported sleep issues at index per primary mental health diagnosis



6.3.2. Characteristics of CAMHS patients prescribed any sleep medication

Patients with a validated prescription of melatonin, promethazine, or any z-drug were compared to patients who were not prescribed any of these sleep medications (see Table 6.5. for the frequencies of prescriptions). The model predicting any sleep medication prescription

(see Table 6.6.) was a significantly better fit than an empty model ($\chi^2(31) = 1512.87$, $p < .001$, log likelihood = -1925.09). Appendix 6.1. examines the difference between melatonin and promethazine prescriptions.

Table 6.5. Frequency of sleep medication prescription within the sample

Medication	Patients with validated prescriptions (N =)	Patients without validated prescriptions (N =)
Any sleep medication	902	6265
Promethazine	293	6874
Melatonin	756	6411

Note: Prescription of z-drugs is very small and is therefore not reported independently. However, any sleep medication includes z-drugs in addition to promethazine and melatonin.

Table 6.6. Multivariate logistic regression model for any (compared to no) sleep medication prescription from CAMHS

Characteristic (reference)	OR (95% CI)	Unadjusted p-value	aOR (95% CI)	Adjusted p-value
Any sleep issues (none)	**3.74 (2.96 – 4.78)**	**<.001**	**3.48 (2.66 - 4.62)**	**<.001**
Male (female)	**1.18 (1.03 – 1.36)**	**.019**	1.04 (0.87 - 1.25)	.658
Log active days in SLaM CAMHS	**3.83 (3.45 – 4.26)**	**<.001**	**2.89 (2.55 - 3.28)**	**<.001**
End reason (turned 18-years-old)				
Discharge	**0.24 (0.20 – 0.29)**	**<.001**	**0.67 (0.52 - 0.86)**	**.001**
Still in services	1.16 (0.95 – 1.41)	.148	**1.52 (1.16 – 1.98)**	**.002**
Socioeconomic status	1.00 (0.99 – 1.01)	.924	1.00 (1.00 - 1.01)	.292
Inpatient admission ever (never)	**19.68 (14.06 –27.98)**	**<.001**	**15.64 (10.36 - 23.93)**	**<.001**
Age at index	**1.03 (1.01 – 1.06)**	**.016**	**0.94 (0.89 - 1.00)**	**.044**
Age group under 12 years (12+)	**0.85 (0.73 – 0.98)**	**.027**	**0.65 (0.47 - 0.91)**	**.011**
Ethnicity (White British)				
Asian ethnicity	**0.58 (0.39 – 0.84)**	**.005**	**0.63 (0.39 - 0.97)**	**.040**
White – Other	**0.61 (0.45 – 0.82)**	**.001**	0.77 (0.54 - 1.08)	.135
Mixed race – White and Black Caribbean	0.80 (0.60 – 1.06)	.131	**0.59 (0.42 - 0.83)**	**.002**
Mixed race – White and Black African	0.60 (0.31 – 1.06)	.096	0.49 (0.23 - 0.98)	.058
Mixed race – Other	0.76 (0.53 – 1.08)	.139	**0.64 (0.42 - 0.96)**	**.037**
Black/Black British – Caribbean	**0.48 (0.33 – 0.69)**	**<.001**	**0.50 (0.32 - 0.75)**	**.001**
Black/Black British – African	**0.33 (0.21 – 0.47)**	**<.001**	**0.34 (0.21 - 0.53)**	**<.001**
Black/Black British – Other	**0.62 (0.49 – 0.76)**	**<.001**	**0.53 (0.41 - 0.69)**	**<.001**
Other ethnicity	**0.47 (0.28 – 0.73)**	**.001**	**0.44 (0.24 - 0.74)**	**.004**
Missing ethnicity data	**0.25 (0.18 – 0.35)**	**<.001**	**0.64 (0.43 - 0.92)**	**.018**
Primary mental health diagnosis (ADHD/hyperkinetic)				
Anxiety, stress, and emotional	**0.24 (0.19 – 0.30)**	**<.001**	**0.40 (0.30 - 0.53)**	**<.001**
Autism or developmental	**0.68 (0.54 – 0.87)**	**.002**	0.94 (0.71 - 1.24)	.643
Depression	**0.39 (0.29 – 0.51)**	**<.001**	**0.58 (0.40 - 0.83)**	**.003**
OCD	**0.21 (0.11 – 0.37)**	**<.001**	**0.27 (0.13 - 0.52)**	**<.001**
Oppositional & conduct	**0.26 (0.18 – 0.38)**	**<.001**	**0.45 (0.29 - 0.68)**	**<.001**
Psychosis	**3.24 (1.89 – 6.26)**	**<.001**	**2.96 (1.40 - 6.27)**	**.005**
Other	**0.32 (0.24 – 0.41)**	**<.001**	**0.52 (0.37 - 0.71)**	**<.001**
None	**0.12 (0.09 – 0.16)**	**<.001**	**0.31 (0.22 - 0.43)**	**<.001**

Note: using index Revised Children's Anxiety and Depression Scale (RCADS) sleep issues. OR = odds ratios, aOR = adjusted odds ratio, CI = confidence interval. SLaM = South London and Maudsley. Relationships in bold were statistically significant.

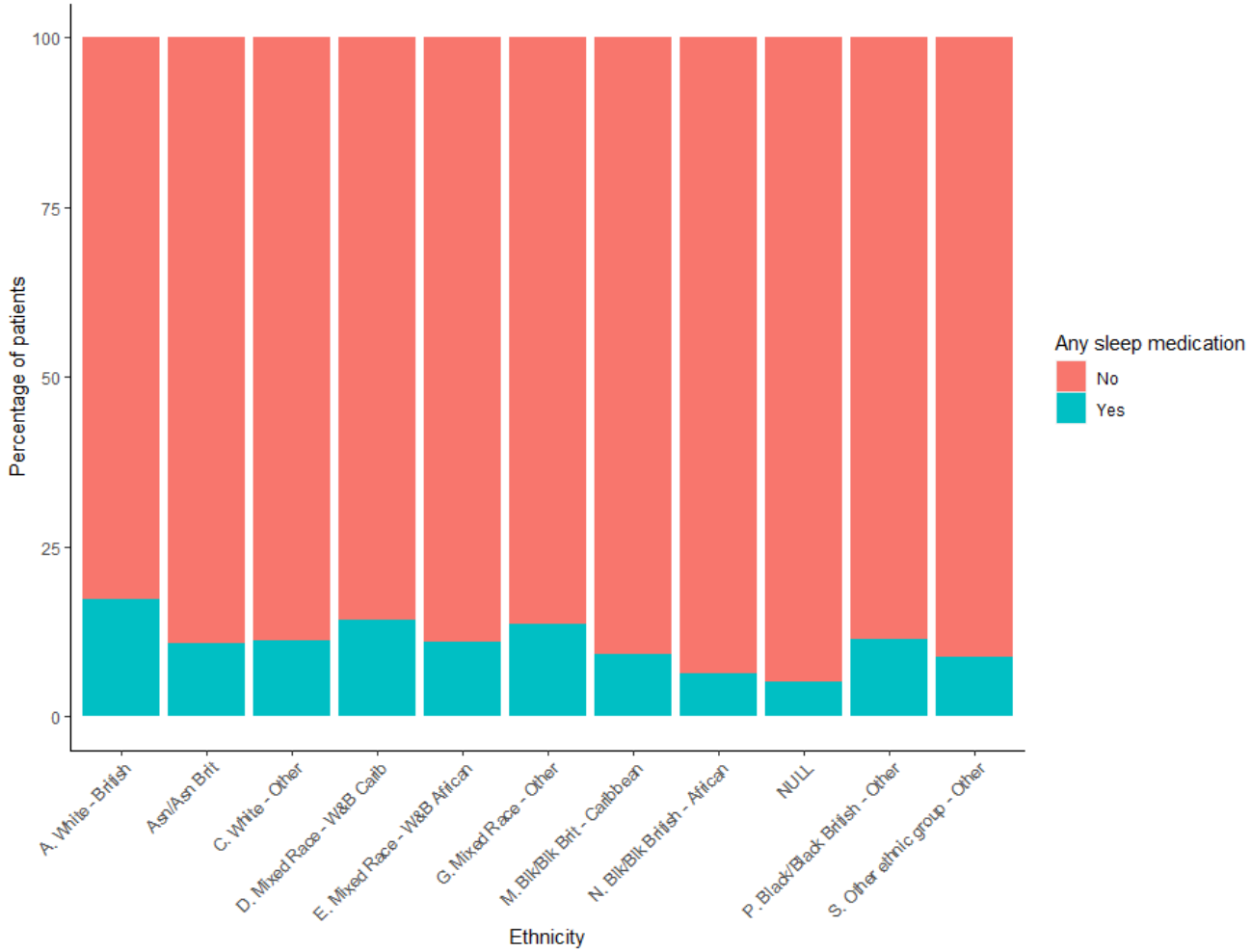
Reporting any sleep issues (compared to none) at index was significantly associated with sleep medication prescription, $B = 1.25$, $p < .001$. Patients who spent a longer time active in the South London and Maudsley CAMHS, $B = 1.06$, $p < .001$, and patients who had at least one inpatient stay during their time in the South London and Maudsley CAMHS, $B = 2.75$, $p < .001$, were significantly more likely to be prescribed sleep medication. Patients who were discharged were less likely to be prescribed sleep medication than patients who left the services at 18-years-old, $B = -0.40$, $p = .001$.

Sex was not significantly associated with sleep medication prescription, $B = 0.04$, $p = .658$ and neither was socioeconomic status $B = 0.004$, $p = .292$. Age at index, $B = -0.06$, $p = .044$, and being over 12 years old at index (compared to under 12), $B = -0.43$, $p = .011$, was significantly associated with sleep medication prescription, in that, younger patients and patients under 12-years-old at the time of completing the first ever Revised Children's Anxiety and Depression Scale (RCADS) were more likely to be prescribed sleep medication than older patients. It is important to note that this is not the age at prescription of sleep medication, but age at the time the first RCADS was completed. As we did not validate sleep medication prescription at the mention level, but rather at the patient level, we cannot determine the start date of medication or the age at medication prescription (see Appendix 6.2. for the amount of time between window start and first RCADS and first mention of medication).

Compared to White British patients, Asian, $B = -0.47$, $p = .040$, Mixed race White and Black Caribbean, $B = -0.52$, $p = .002$, Mixed race (other), $B = -0.44$, $p = .037$, Black/Black British Caribbean, $B = -0.70$, $p = .001$, Black British/Black British African, $B = -1.07$, $p < .001$, Black/Black British (other), $B = -0.63$, $p < .001$, other ethnicity, $B = -0.83$, $p = .004$, and no reported ethnicity B= -0.45, $p = .018$, patients were significantly less likely to be

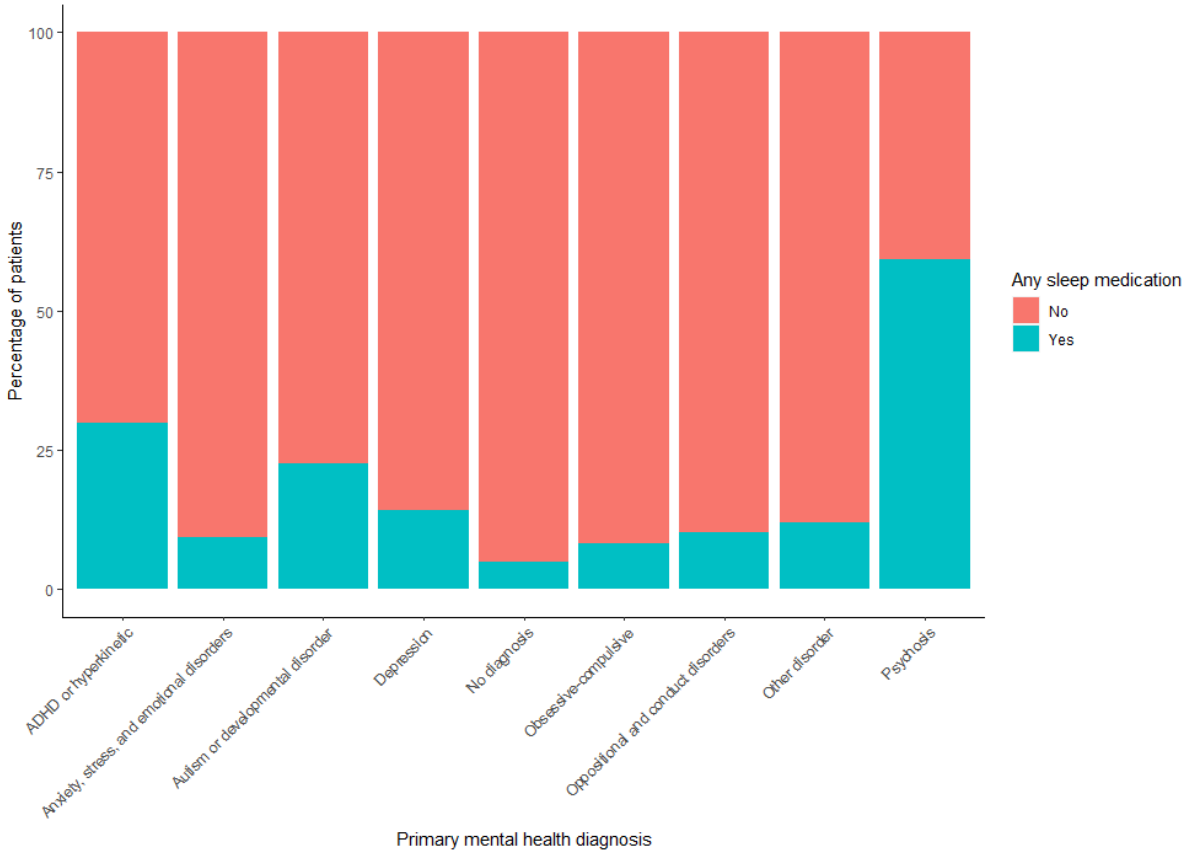
prescribed any sleep medication. White (other), B = -0.26, p = .135, and Mixed race White and Black African patients, B = -0.71, p = .058, were not prescribed sleep medication at significantly different rates than White British patients. Figure 6.3. below demonstrates the percentage of patients receiving sleep medication prescriptions from CAMHS per ethnic group (Appendix 6.4. provides a figure representing the number of patients for this relationship).

Figure 6.3. Percentage of patients receiving a validated sleep medication prescription from CAMHS per ethnicity group



Compared to patients with ADHD, patients with anxiety, stress, and emotional disorders, B = -0.92, p <.001, depression, B = -0.55, p = .003, no diagnosis, B = -1.18, p <.001, OCD, B = -1.31, p <.001, oppositional and conduct disorders, B = -0.80, p <.001, and other disorders, B = -0.66, p <.001, were significantly less likely to be prescribed sleep medication. Autistic patients and patients with other developmental disorders, B = -0.06, p = .643, were not significantly more or less likely to be prescribed any sleep medication than patients with ADHD. However, patients with psychosis were significantly more likely to be prescribed sleep medication, B = 1.08, p = .005, than patients with ADHD. Patients with psychosis were more likely to be prescribed promethazine and less likely to be prescribed melatonin (see Appendix 6.1.).

Figure 6.4. Percentage of patients receiving sleep medication prescription from CAMHS per primary mental health diagnosis



6.4. Discussion

The current study aimed to characterise South London and Maudsley CAMHS patients experiencing sleep issues and those receiving a sleep medication prescription. Over three-quarters of the sample (75 – 82%) reported trouble sleeping at least sometimes. Approximately one-fifth to one-quarter of patients reported always having trouble sleeping (21 – 25%), whilst one-in-eight (13%) patients were prescribed sleep medication. First, the characteristics of CAMHS patients reporting sleep issues are discussed in Section 6.4.1. Predictors of sleep medication prescription are then discussed in Section 6.4.2. Lastly, differences in the characteristics of patients that reported sleep issues and those who received sleep medication are discussed in Section 6.4.3. This exploratory study sets groundwork for future confirmatory analyses, including examining whether similar relationships are found across other UK CAMHS. Whilst the sharing of data is not possible due to confidentiality and data protection, data are available to researchers who apply to access CRIS and pass NHS ethics and information governance procedures (project number 20-033).

6.4.1. Sleep issues

Older CAMHS patients reported more sleep issues than younger patients, and girls reported greater levels of sleep issues than boys, as seen in other studies (de Zambotti et al., 2018). These relationships may be explained by differences in pubertal development across age and gender (Marshall & Tanner, 1969), and the impact of puberty on delaying circadian rhythms (Carskadon et al., 1993; Hummer & Lee, 2016). Unfortunately, it was not possible to examine the role of puberty or circadian rhythms within the current analyses and dataset, therefore, future research would be required to test this hypothesis. Patients who left CAMHS at 18-years-old reported sleep issues more often than patients who were discharged prior to the ‘adult’ threshold. Perhaps patients who entered CAMHS when they were older were more

likely to be discharged at 18-years-old due to less time between entering the services and turning 18. This explanation would suggest older patients would spend fewer active days in CAMHS. However, patients that were in CAMHS for longer reported sleep issues more often than patients who were in CAMHS for less time. Together, this suggests that sleep issues at index (perhaps at assessment upon entering CAMHS) may be associated with more complex mental health presentations, requiring longer treatment. This would make sense, given the complex relationships between poor sleep and mental health complexities such as self-harm (e.g., McGlinchey et al., 2017), suicidal ideation (e.g., Littlewood et al., 2019), emotion dysregulation (e.g., Baum et al., 2014) and intrusive thoughts (e.g., Harrington & Cairney, 2021; Chapter 5).

A previous study found that children from lower socioeconomic status backgrounds experience more pre-sleep worries (Bagley et al., 2015). Similarly, the present study found that lower socioeconomic status was associated with greater reporting of sleep issues. When sleep-deprived, the prefrontal cortex is less able to inhibit amygdala and hippocampal activity, resulting in more intrusive thoughts (e.g., pre-sleep worries) further exacerbating poor sleep (Harrington & Cairney, 2021). These relationships may be more prevalent during prefrontal cortex maturation and for adolescents from low socioeconomic status backgrounds. Future research could consider the role of pre-sleep worries or intrusive thoughts and socioeconomic status in the context of Harrington and Cairney's (2021) framework. Exploring these relationships may be of increasing importance, given the current cost of living crisis.

Another demographic variable that was significantly associated with reporting of trouble sleeping was ethnicity. White British patients were the most likely to report sleep issues (Figure 6.1.) and were also the largest group. Therefore, White British patients were used as the reference condition in all analyses. Compared to White British patients, Black/Black

British Caribbean, African or other Black/Black British ethnicity, mixed race White and Black Caribbean, and other White ethnicity patients were significantly less likely to report sleep issues. Black patients reported the lowest levels of sleep issues, particularly African, Caribbean, and then Other Black / Black British patients. We did not find a significant difference in reporting of sleep issues between White British patients and Asian, mixed race White and Black African, other mixed race, other ethnicities, or no reported ethnicity patients. Our interpretation of these results is speculative as these relationships typically lack empirical and theoretical investigation. Previous research shows that White participants report *fewer* sleep issues than participants of other ethnicities (Perlus et al., 2018; Zeiders et al., 2017), whereas our study finds the opposite. The previous studies were conducted in the USA. However, research in the UK with adults also suggests that participants from minority ethnicities experience poorer sleep than White participants (Dawkins et al., 2022). Compared to the average across England and Wales, the sample in this study included more Black people (29% compared to 4%) and fewer White people (38% compared to 81%; Office for National Statistics, 2022). However, when looking at the ethnicity of people living in the London boroughs of Lewisham, Croydon, Southwark, and Lambeth that South London and Maudsley CAMHS covers, more White British patients (42.57%) were included in the sample (excluding patients with recorded null ethnicity) than the average across the boroughs (36.93%) and more Black patients in total (28.69%) than the average in the boroughs (24.63%). However, the reporting of specific Black ethnicities (e.g., Caribbean, African) in the present study, under-represents Black African (6.93% compared to the borough average of 12.6%) and Black Caribbean (5.84% compared to 8.7%) but over-represents Black Other (15.92% compared to 3.33%). Nonetheless, the results across all Black ethnicities were in the same direction, suggesting that reporting of specific sub-categories of ethnicity would not influence the results. There were also fewer Asian (3.55% compared to 5.15%) and White

Other (7.65% compared to 14.68%) people included in the sample compared to the average across the South London and Maudsley boroughs.

Ethnicity itself may not be the driving factor for poor sleep, instead experiences of marginalisation and discrimination may contribute to sleep issues. Perhaps the young people that overcome the barriers and attend CAMHS are those that have experienced lower levels of discrimination or marginalisation than their peers who may face greater barriers. For example, experiences of racism and stigma around mental health have been identified as barriers to Black adolescents accessing mental health care in Canada, whilst supportive friends, family and service providers acted as facilitators (Fante-Coleman & Jackson-Best, 2020). Experiences of discrimination and marginalisation as barriers to accessing mental health services may be particularly the case when considering intersectionality between different combinations of systemic discrimination (e.g., socioeconomic status, ethnicity, gender, disability), which are not considered in the current study. The young people who do overcome the barriers and attend CAMHS may not be representative of other young people that are unable to access support from the services. Additionally, once young people reach CAMHS, they may still experience biases resulting in less representative data. For example, systematic biases in the demographics of CAMHS patients that complete Routine Outcome Measures (ROMs) has been identified (Morris et al., 2021).

Using the large NHS data set (CRIS) that was used to derive the sample for the current study, Morris et al. (2021) investigated the completion rates of caregiver reported Strengths and Difficulties Questionnaires (SDQ; Goodman, 2001), one of the commonly collected ROMs in CAMHS. Their results showed that boys were more likely to have a completed caregiver SDQ, whilst CAMHS patients that were older or from the lowest socioeconomic neighbourhoods were less likely. Further, Black, or Asian CAMHS patients were less likely than White CAMHS patients to have a completed caregiver SDQ. The representativeness of

the current study's sample is unknown, as Morris et al. (2021) was the first study to examine systematic biases in ROMs completion rates. Therefore, the CRIS data in the present study may not be fully representative of the CAMHS population. However, some of the biases identified by Morris et al. (2021) may be related to caregiver reports (e.g., older adolescents may involve their caregiver less), which would be less problematic in the current study that includes both caregiver- and self-reported measures. Further, CAMHS populations are not representative of the relationships between sleep and ethnicity in the general population.

Mental health diagnosis was significantly associated with reporting of sleep issues. ADHD was used as the reference condition due to high levels of sleep issues for children with ADHD (Cortese et al., 2009), high levels of sleep medication prescription for this population (Efron et al., 2014; Kimland et al., 2020), and previously available information regarding paediatric sleep medication prescription for children with ADHD (NICE, 2013). In the current study, patients with ADHD were also the second most likely to receive a sleep medication prescription (discussed in Section 6.4.2) apart from patients with a primary mental health diagnosis of psychosis, which had a much smaller sample size and therefore would not have been an appropriate reference condition. Patients with mood disorders such as depression, anxiety, stress, and emotional disorders were significantly more likely to report sleep issues than patients with ADHD. Theoretically, this could be explained by Goldstein and Walker's (2014) framework regarding how lack of REM sleep can contribute to mood dysregulation. The high rates of sleep issues associated with patients with depression may also be explained by negative, relative to positive, emotional memory biases (e.g., ATaC, Harrington et al., 2017; Chapter 4). This also supports the suggestions in Chapter 5 that positive and negative affect are important in the relationship between sleep and mental health. Further, patients with no mental health diagnosis were significantly less likely to report sleep issues than patients with ADHD, suggesting that patients with fewer sleep issues

perhaps have fewer symptoms of poor mental health. As this research is exploratory, correlational, and takes a broad-stroke approach, causation cannot be implied, and more research is needed to further explore these relationships and mechanisms.

6.4.2. Prescription of sleep medication

In the current study, we included the two main sleep medications that may be prescribed in CAMHS: melatonin (brand name Circadin) and promethazine (brand name Phenergan). We also included z-drugs that are less likely to be prescribed in CAMHS, but are still sleep specific (e.g., zopiclone, zolpidem; YoungMinds, 2021). Melatonin is a hormone involved in initiating sleep (Borbély, 1982; Borbély et al., 2016). Prescribing melatonin is not licensed for children and adolescents in the UK, due to lack of sufficient clinical trials. Whilst preliminary evidence suggests there are minimal side effects associated with melatonin prescription for children and teenagers with ADHD, autism, or other populations, these studies have been criticised for having small samples and typically investigating melatonin use only in the short-term (Cummings et al., 2012). Therefore, there may be potential unknown side effects which require further investigation. Promethazine is a drowsy antihistamine licensed for short-term sedation in all ages above 2-years-old; although prescription for sleep is less recommended than for its antihistamine/anti-emetic properties (Joint Formulary Committee, 2021). Therefore, promethazine may be used more often for short-term sedation following severe sleep difficulties. Prescription of sleep medication in the current study's CAMHS population (~13%) is much greater than the approximately 2% in the general population of children and young people in Norwegian and Swedish population studies (Kimland et al., 2020; Oerbeck et al., 2020), but lower than for children and adolescents in residential care (17%; Oerbeck et al., 2020). However, Oerbeck et al. (2020) included other sedative medications (e.g., alimemazine, benzodiazepines), which are not

included within our study due to their ability to treat a multitude of symptoms other than just sleep. As melatonin prescription in our study is similar to Oerbeck et al.'s (2020) study in youth residential care (~11%), it is possible that the exclusion of sedative medication with other functions (e.g., anti-depressants) in the current study may account for the approximately 4% difference between our studies. Future research may wish to examine the prescriptions of other psychotropic medications with sedating properties, as well as the use of ADHD stimulants or other medications that may impact sleep and increase the use of sleep medication.

In the present study, prescription of sleep medication was highest for patients with psychosis (see Figure 6.4.). However, this sample was much smaller. Therefore ADHD was used as the reference condition, as patients with ADHD also demonstrated high levels of sleep medication prescription. ADHD or autism and developmental disorders such as learning disabilities did not significantly differ in prescription of sleep medication, whilst other mental health disorders had significantly lower levels of sleep medication prescription. This has policy implications given the fight against over-medication of people with learning disabilities, autism, or both and appropriate prescriptions in paediatric services (e.g., STOMP-STAMP; NHS England & The RCPCH, 2018). These results are also in line with the NHS Digital data, which suggests that melatonin prescriptions were highest for patients with ADHD and a comorbid learning disability, when examining health care for people with learning disabilities (NHS Digital, 2021). However, our analyses only included primary diagnoses, which does not account for complexities associated with co-morbid presentations. Future research should consider the role of co-morbidities, as well as other psychotropic medication prescriptions to further address STOMP-STAMP aims. Such research in CAMHS could also have implications for NICE guidance, especially as the previous guidance on melatonin prescription for children with ADHD (NICE, 2013) has been withdrawn in light of

this emerging field of evidence. It may also be that different young people (e.g., with different diagnoses) receive various treatment modalities from CAMHS, which may include psychological sleep interventions. Therefore, there may be biases in the characteristics of CAMHS patients receiving pharmacological compared to psychological sleep interventions.

Considering the types of sleep medication prescribed in relation to primary mental health diagnoses and inpatient admission may suggest some potential mechanisms underlying these relationships, which could benefit from further exploration. Promethazine may be prescribed under conditions of short, but extreme periods of sleep issues (e.g., extended wakefulness or insomnia over a period of days). Such extreme sleep difficulties may be more commonly associated with acute mental distress, potentially more likely to result in inpatient admission. However, some patients (e.g., patients with depression) may have more chronic sleep issues and may benefit from a regular prescription of sleep medication (e.g., melatonin), to support their ongoing sleep issues. In general, the first mention of promethazine typically occurs later than the first mention of melatonin (see Appendix 6.2.). Therefore, it may be that promethazine is prescribed after a longer time in the services, perhaps when symptoms are more severe or escalating in comparison to earlier in treatment. Additionally, as an inpatient, prescriptions may be written pre-emptively, increasing the likelihood of their use. Part of the validation of sleep medication mentions included that it appeared young people took the medication, which should have reduced the likelihood of pre-emptive prescriptions without administration being validated.

It is worth noting that the current study does not necessarily capture the entire use of sleep medication in South London and Maudsley CAMHS. Some CAMHS patients may receive sleep medication without ever having completed an RCADS, and therefore would not be included in the current sample. Additionally, only the *prescription* of sleep medication was examined in this study, not the use or recommendation of over-the-counter sleep

medication. Melatonin can be purchased online or from abroad (although unlikely to be medical grade), and promethazine can be purchased over the counter in the UK. Therefore, just because a patient is not receiving a prescription for sleep medication, does not suggest they are not accessing pharmacological sleep aids through other methods. Additionally, just because a patient is not receiving prescribed pharmacological interventions for sleep does not mean that they are not receiving any sleep intervention. Non-pharmacological sleep interventions (e.g., sleep hygiene, cognitive behavioural therapy (CBT) for insomnia) may be used, but were not examined within the current study, due to lack of standardisation in delivery (e.g., Irish et al., 2015) and difficulty extracting information regarding these interventions from clinical records.

Non-pharmacological interventions may be offered alongside, or as an alternative to sleep medication. Mentions of sleep hygiene were anecdotally frequent within the free text in this study but would be difficult to extract due to lack of standardisation. The use of digital CBT for insomnia in CAMHS seems promising (Cliffe et al., 2020). CBT for insomnia is also found to reduce symptoms of major depression whilst supporting the efficacy of other mental health treatments (Manber et al., 2008). Conversely, psychological therapies for depression have been found to indirectly support sleep (Reynolds et al., 2020), and other therapies may involve aspects of sleep interventions not considered within the scope of the current study. This may result in a lower need for sleep medication prescription for these patients. Improving sleep via any method may have subsequent benefits for mental health. Future research could investigate the use of pharmacological and non-pharmacological sleep interventions within the NHS, including frequency of use, adherence, acceptability, and efficacy, likely using natural language processing applications. More research is required to support evidence-based practice for treating sleep issues in CAMHS.

6.4.3. Discrepancies between reported sleep issues and sleep medication prescription

As the models for the first and second research question are comparable, a narrative description of the discrepancies between reporting sleep issues and receiving sleep medication is possible. Sex was not significantly associated with sleep medication prescription, despite girls being more likely to report sleep issues than boys. When sex was the only predictor in the model, boys were significantly more likely than girls to be prescribed sleep medication. This supports Kimland et al.'s (2020) study that considers only prevalence and does not statistically adjust for other variables that may account for more variance than sex. Therefore, potential sex biases in sleep medication prescription could be associated with interactions with other variables (e.g., diagnostic biases). For example, diagnoses of ADHD and Autism are often earlier and at higher rates for boys than girls (e.g., Lai et al., 2022), and our results show that patients with these diagnoses have higher rates of sleep medication prescriptions. Socioeconomic status was also significantly associated with reporting of sleep issues, but not with sleep medication prescription. This is contrary to previous research that found higher rates of sleep medication prescription in more deprived localities (Soyombo et al., 2020). However, the current study includes individual SES, rather than SES of the services. The South London and Maudsley CAMHS operates cross the South London boroughs of Croydon, Lambeth, Lewisham, and Southwark. Relative to the rest of the UK, the average income in Croydon is £23,756 (1.13 times higher than the average), £25,046 in Lewisham and Southwark (1.19 times higher) and £29,003 in Lambeth (1.37 times higher), however this is comparatively low compared to some other London boroughs (Office for National Statistics, 2021). More research may be needed to examine the support that young people from lower socioeconomic backgrounds receive for sleep, given higher rates of sleep issues in this group.

There were also differences in the ethnicity of patients prescribed sleep medication and reporting sleep issues. White British patients were most likely to report sleep issues and receive sleep medication prescriptions. Compared to White British patients, patients of Asian, non-specified mixed ethnicity, other ethnicity, and no reported ethnicity were significantly less likely to be prescribed sleep medication, yet they did not significantly differ in reporting of sleep issues. In some cases, these relationships differed for melatonin and promethazine prescription (see Appendix 6.1.). Patients of other White ethnicity were significantly less likely to report sleep issues but were not significantly less likely to receive sleep medication, than White British patients. When controlling for all other variables in the study, including reported sleep issues, White British patients were significantly more likely to receive sleep medication prescriptions than Black patients and mixed race Black and White Caribbean patients. This supports McLay et al.'s (2021) research in New Zealand that found autistic children of European ethnicity were significantly more likely to receive melatonin prescriptions than autistic children of other ethnicities. McLay et al. (2021) suggested this could be related to cultural attitudes towards medication use and health services in general. Future research could examine the mechanisms behind the relationship between ethnicity, sleep, and sleep medication prescription.

Mental health diagnosis also demonstrated discrepancies between reported sleep issues and prescription of sleep medication. Compared to patients with ADHD, patients with depression, or anxiety, stress, and emotional disorders, were significantly more likely to report sleep issues, but less likely to receive sleep medication. Therefore, higher levels of sleep medication prescription for patients with ADHD may not be due to greater sleep issues alone. Patients with depression or anxiety, stress, and emotional disorders may be prescribed drowsy anti-depressant or anti-anxiety medications, reducing specific sleep medication prescriptions for these groups. Patients with ADHD may also be prescribed sleep medication

to counteract stimulant effects associated with ADHD medication (e.g., methylphenidate), thus increasing sleep medication prescription rates. Kimland et al.'s (2020) study found that approximately 80% of young people receiving melatonin prescriptions were also prescribed other psychotropic medications, suggesting the use of other medication may be a plausible explanation. Patients with psychosis were more likely to receive promethazine, but less likely to receive melatonin than patients with ADHD (see Appendix 6.1.), despite being significantly less likely to report sleep issues at index. This could reflect the use of promethazine for short-term sedation for more severe sleep issues (e.g., multiple days of complete sleeplessness or insomnia) that may be more common during psychotic episodes. For example, in an inpatient setting, people suffering during episodes of psychosis found sedative medications were effective in preventing suicide and overcoming the suffering caused by sleeplessness (Fredriksen et al., 2020).

Patients with at least one mental health inpatient admission were more likely to be prescribed sleep medication than patients who were never hospitalised. This was particularly the case for promethazine, but not melatonin (see Appendix 6.1.). However, inpatient admission was not significantly associated with index sleep issues. Nonetheless, sleep issues may have deteriorated from index RCADS (completed in community mental health settings), as previous research has found that *recent* sleep issues are a predictor of psychiatric inpatient hospitalisation following self-harm or suicide attempts (Borschmann et al., 2018). In future research natural language processing applications could be used to investigate the relationship between sleep issues, sleep medication, and self-harm in CAMHS. The current study was not time-bound, as mentions of sleep medication are validated at the patient, rather than mention level, due to ambiguity in free-text entries. Future research investigating sleep issues and sleep medication prescription more precisely over time may be able to address these questions further.

Service use was also associated with sleep issues and sleep medication prescriptions. Patients spending a greater number of days active in the South London and Maudsley CAMHS were significantly more likely to report sleep issues and be prescribed sleep medication. Prescription of sleep medication and reporting of sleep issues were also more likely for patients who left the services because they turned 18-years-old, compared to patients who were discharged. Together, this may suggest that sleep issues are a maintaining factor for poor mental health and require a greater length of ongoing mental health treatment, which is potentially costly to the NHS. This may also be related to more complex symptoms and presentations associated with poor sleep. Further investigation into evidence-based treatments of sleep issues may prove beneficial to reducing long lasting poor mental health and may be lifesaving and cost saving to the NHS in the long-term. However, causally sensitive research would be required in the future, as direction cannot be implied from associations identified in the current study. Limitations of the current study are discussed below.

6.4.4. General limitations

The reporting of sleep issues is limited by the use of a binary categorisation as 'any' or 'no' sleep issues. Other NHS data analysis has used 'any' or 'no' reported sleep issues defined as three or more days of various sleep issues reported in the past week (see Table 1.7a Lifestyles Team & NHS Digital, 2021). Unfortunately, the Revised Children's Anxiety and Depression Scale (RCADS) does not include any timeframe, as such there are subjective discrepancies in reporting due to greater ambiguity of response options (i.e., "always", "often", "sometimes", and "never") that are less objective than time frames. Nonetheless, binary categorisation allowed for a narrative comparison between the prescribed sleep medication and reported sleep issues models (Section 6.4.3. above), allowing a more coherent

interpretation of the results. The measurement of sleep issues in the current study is further limited by the use of a single item ("I have trouble sleeping"). From this single item it is not possible to examine the severity or types of sleep issues experienced by patients, which may require different treatments. It may also be that patients with long-standing and severe sleep issues are more likely to receive sleep medication prescriptions, but this cannot be examined through our limited sleep measure.

However, there are no routinely collected questionnaires used in the NHS to monitor sleep disturbances. Natural language processing applications have been developed to extract information from the free-text entries (e.g., clinical notes) and could be used to examine sleep disturbances (including nightmares, night-terrors etc.) in CRIS data. The natural language processing application which has been developed to identify sleep disturbances has only been validated on a small sample, with 68-75% agreement and issues surrounding false positives. Future research may wish to further validate and use a natural language processing application to examine sleep disturbances, rather than using the RCADS item, as this may allow finer-grained analyses. The natural language processing applications may also be able to provide information about sleep issues closer to the date of sleep medication prescription than the RCADS items used in the current study, therefore allowing time-bound analyses such as historical cohort studies. The current analyses are also limited by only considering independent effects, rather than interactions between variables, even in the adjusted models. Future research could consider interactions between demographic variables such as via latent class analyses, in which emerging groups of patients with sleep issues and/or receiving sleep medication may be observed.

Unfortunately, due to the risk of identification associated with smaller samples, some variables were collapsed. For example, sample sizes were too small for 'other gender' or 'gender not specified' to conduct meaningful analyses. Therefore, gender outside of the

binary categories assigned to sex (e.g., male and female) were assigned NA in the current study. Diagnoses were grouped, in line with methodologies in previous research (Downs et al., 2016). Additionally, ethnic groups with less than 100 patients were collapsed into larger groups (e.g., "White – Irish" was collapsed into "White – Other"). Future research with even larger samples (e.g., population level data) may be able to maintain the original groupings and examine these relationships in a more fine-grained manner. Nonetheless, CRIS data provides a large sample of patients in South London and Maudsley CAMHS with a wide range of available variables and possibility to link to other data such as the national pupil's database, social care records, and hospital admission data. The current study extracted data for 7,196 patients, although other extraction criteria may yield larger samples. Additionally, there may be biases in the extraction of data, or in the presence of missing data. For example, looked after children may lack continuity and may experience greater amounts of missing data. This may mean that some biases in sleep medication prescription are missed within the current study. For example, previous research has found that young people in residential care who are asylum seekers without accompanying guardians are less likely, than other children in residential care, to receive sleep medication (Oerbeck et al., 2020). It is also important to note that I was able to access the data up until January 2022, when the information governance requirements expired. After this, the data were only accessible via a collaborator and King's College London, limiting the opportunity for further analyses.

6.4.5. Conclusion

In conclusion, the prescription of sleep medication in the South London and Maudsley CAMHS is similar to prescription patterns in youth residential care (Oerbeck et al., 2020), with melatonin being the most prescribed. Narrative comparison of the two models analysed in the present study identified some potential discrepancies between reporting of sleep issues

and prescription of sleep medication, which are associated with demographic variables such as ethnicity, sex, and mental health diagnosis. This raises some questions regarding potential biases in prescription of sleep medication that may be of interest in future research that is causally sensitive. Further research regarding sleep issues and sleep treatment within CAMHS populations would be beneficial for informing care plans for treating sleep issues. This is particularly important, as transdiagnostic issues, such as sleep, are often not prioritised within mental health care (Freeman et al., 2020). Treating sleep issues may also have implications for service-use such as active days in CAMHS, inpatient admissions, and achieving discharge rather than transitions to adult mental health services. More in depth exploration of the associations between sleep issues and CAMHS service-use may be beneficial, especially when adult mental health services have fewer resources than CAMHS (Hill et al., 2019).

Chapter 7: General Discussion

7.1. Thesis summary

This thesis includes five empirical chapters, which explore the impact of adolescent sleep on learning and memory (Chapters 2 – 4), and affective wellbeing and mental health (Chapters 4 – 6). The thesis takes a broad view of adolescence, analysing data from a total of 557 participants aged 13-24-years-old (Chapters 2 – 5) and a further sample of 7,167 patient records from the child and adolescent mental health services (CAMHS, Chapter 6). Chapters 2 and 3 used the "cathedruke" non-word learning paradigm to investigate the role of sleep for memory consolidation (Chapter 2) and the restoration of encoding capacity (Chapter 3). Chapter 4 considered how our understanding of the relationship between sleep, learning and memory (e.g., Chapters 2 and 3) may also be related to positive and negative affect as well as emotional memory biases. As such, Chapter 4 links the two halves of the thesis and has implications for understanding the potential mechanisms between sleep and mental health. Chapter 5 explored further mechanisms behind the relationship between sleep and mental health, such as the ability to inhibit thoughts and responses to emotional stimuli. Given the relationships between sleep and mental health (e.g., seen in Chapter 5), Chapter 6 explored the demographic and service use predictors of sleep issues and sleep medication prescriptions in the Child and Adolescent Mental Health Services (CAMHS). Below is an overview of each empirical chapter included in the thesis.

7.1.1. Chapter 2: Adolescent sleep, learning, and memory during the school term and holiday

The first study of the thesis is a pilot study exploring the relationship between adolescent sleep restriction (13-16-year-olds) during the school term compared to the

holiday. This study used polysomnography in participants' own homes in an attempt to capture naturalistic sleep. Whilst sleep restriction was not observed in the polysomnographic data, due to methodological issues such as the small sample size and lack of counterbalancing caused by COVID-19 lockdowns, adolescents appeared to get less sleep on school nights compared to term-time weekends and school holidays according to actigraphic and diary data collected over multiple nights. The impacts of potential sleep restriction on overnight memory consolidation were examined using the "cathedruke" non-word learning paradigm. Performance on this paradigm has been associated with sleep in adult (e.g., Tamminen et al., 2010) and child (e.g., Henderson et al., 2012) participants. The present study is the first time, to our knowledge, that this paradigm has been used with adolescent participants. Thus, the study in Chapter 2 fills a current gap in the word-learning literature. In the non-word learning paradigm lexical integration was greater in the holiday than the term-time immediate test. However, in the delayed test, lexical integration remained the same for the holiday condition, but showed overnight increases in the term-time, suggesting consolidation occurred in the term but not the holiday.

This pilot study provides evidence that these methodologies (e.g., "cathedruke" paradigm and at-home polysomnography) would be appropriate for future research aiming to examine the impact of school on adolescent sleep restriction and the consequences for learning and memory. This research design would be appropriate for addressing two major gaps in the literature. Firstly, the impact of current UK school start times on adolescent sleep architecture without experimental manipulation of sleep. Secondly, understanding word-learning and its associations with sleep during adolescent development, explored using the "cathedruke" paradigm. Potential amendments to the design to ensure greater feasibility in the future were suggested earlier, in Section 2.5.3.

7.1.2. Chapter 3: Is word learning capacity restored after a daytime nap?

The same "cathedruke" non-word learning paradigm used in Chapter 2 was also used in the second study of the thesis, which was a laboratory-based study published as a registered report (March et al., 2023). In this chapter, instead of investigating memory consolidation after sleep (as in Chapter 2), Chapter 3 used the "cathedruke" non-word learning paradigm to explore how a daytime nap *before* learning prepares the brain for subsequent encoding and memory. Using a sensitive measure of cued recall accuracy (Levenshtein distance), the current study found that cued recall was significantly more accurate following a daytime nap than an equivalent period of wake. The study presented in Chapter 3 did not support the predictions made by the synaptic homeostasis hypothesis, that slow-wave sleep would be fundamental for restoring encoding capacity (Tononi & Cirelli, 2003, 2006, 2014). Instead, frontal sleep spindle density prior to encoding was significantly associated with greater overnight lexical integration. Therefore, sleep spindles may have been involved in preparing the brain for subsequent encoding. Another study, published after our stage one registered report was accepted, found that spindle density during a daytime nap was associated with restoring the hippocampus for subsequent encoding (Ong et al., 2020). Perhaps the active systems consolidation account could explain the findings in Chapter 3 and Ong et al. (2020), in that hippocampal-to-neocortical dialogue may occur via sleep spindles. Such hippocampal-to-neocortical dialogue may reorganise information into the neocortex, allowing the hippocampus to be restored for subsequent encoding. The findings in Chapter 3 further add to this, by suggesting that sleep spindles prior to encoding may have implications for subsequent consolidation. The impact of sleep both prior to and following encoding will need to be examined in future research to further explore this question.

7.1.3. Chapter 4: Sleep, affect, and emotional memory biases

Similarly, to Chapter 3, Chapter 4 also examined encoding after sleep. This chapter focused on encoding and memory for emotionally valenced stimuli, creating links between learning, memory, and affective processes. Data were collected online from 18-24-years-olds during the June COVID-19 lockdown. Repeated data were collected one-year later (outside of lockdown conditions) for approximately half of the sample. Contrary to predictions, participants higher in negative affect recalled *fewer* negative word-pairs than participants lower in negative affect. This may have been an inhibitory mechanism to regulate negative affect. The role of sleep for inhibition and affective wellbeing in relation to mental health is explored further in Chapter 5. At Timepoint 2 (one-year follow-up), but not Timepoint 1 (during COVID-19 lockdowns), poor habitual sleepers were found to have higher memory for positive than neutral word-pairs, whilst the opposite was found for participants that slept better. Negative memory was not significantly associated with habitual sleep, suggesting negative memory was preserved despite poor sleep, as demonstrated by previous research (e.g., Vargas et al., 2019, Tempesta et al., 2016, Walker et al., 2009). However, whilst the relationship was not significant ($p = .057$), last night's sleep at Timepoint 2 showed some association with positive and negative memory biases, in that, participants with poorer sleep on the night preceding the study demonstrated a larger gap between negative (higher) and positive (lower) memories. Using more emotionally salient stimuli, for example under laboratory conditions, this relationship may be significant. Given the different relationships between Timepoints 1 and 2, and between habitual sleep and last night's sleep, selective consolidation may explain the results (e.g., ATaC model; Harrington et al., 2017). Poor habitual sleepers may have selectively consolidated positive, relative to neutral memories, as a way to counteract potential negative memory biases. As such, higher positive memory may act as an emotion regulation strategy, similar to the relationship between higher negative

affect and lower negative memory biases observed at Timepoint 1. Whereas last night's sleep may influence recall to a greater extent, in that poor sleep on the night preceding the study may increase the recollection of negative memories (e.g., Harrington & Cairney, 2021), whilst good sleepers on the preceding night may have been able to inhibit negative memory recall to a greater extent.

7.1.4. Chapter 5: Adolescent sleep, mental health, and mediators

In addition to the role of sleep for inhibiting negative thoughts and the recollection of negative memories (e.g., Harrington & Cairney, 2021; Chapter 4), sleep may also impact on the ability to up-regulate positive affect and down-regulate negative affect. Chapter 5 investigated the role of emotion recognition, inhibition of responses to affective stimuli, and the ability to control unwanted thoughts as potential mechanisms of the relationship between sleep, depression, anxiety, and stress. These relationships investigate predictions made by models proposed by Harrington and Cairney (2021) and Goldstein and Walker (2014). Both models suggest a role of the prefrontal cortex, which matures throughout adolescence. Yet neither model directly considers how adolescent development may influence the relationships proposed in their models. The study in Chapter 5 explores collected data with 13-16-year-olds and 18-24-year-olds to examine if there are any developmental differences in the mediators between sleep and mental health that may be associated with brain maturation. Positive and negative affect were also considered, given the tripartite theory that suggests that positive affect is a unique characteristic of depression that distinguishes it from anxiety, both of which are characterised by high negative affect (Clark & Watson, 1991). Furthermore, adolescents are found to have a greater relationship between sleep and positive affect than negative affect or mental health (e.g., Short et al., 2020).

The study described in Chapter 5 found that lower thought control ability, as a serial mediator with negative affect and independently, significantly mediated the relationship between poor habitual sleep and higher levels of mental health symptoms. For the 13-16-year-old group only, high positive affect significantly mediated the relationship between good habitual sleep and lower levels of depression. Positive affect is widely understudied and is not considered within Harrington and Cairney's (2021) or Goldstein and Walker's (2014) model in detail. The findings in Chapter 5 suggest positive affect may be a useful target in the relationship between sleep and mental health for teenagers. This may have implications for interventions within Child and Adolescent Mental Health Services (CAMHS), which could improve sleep and increase positive affect to reduce depression. For example, interventions such as behavioural activation that aim to increase positive affect for people with depression, should be considered alongside sleep, at least for teenagers. Additionally, Cognitive Behavioural Therapy (CBT) emotion regulation strategies that rely on thought control abilities (e.g., cognitive reappraisal; Engen & Anderson, 2018) may also be more effective when sleep is supported.

7.1.5. Chapter 6: Sleep problems and sleep medication prescription in CAMHS

Given the relationship between sleep and mental health during adolescent development (e.g., Chapter 5), sleep may be particularly important in CAMHS. Chapter 6 used secondary NHS data from South London and Maudsley CAMHS to explore the characteristics of patients reporting of sleep issues and those receiving sleep medication prescriptions. This study found that most CAMHS patients reported at least some sleep issues. Specifically, girls were more likely to report sleep issues than boys. Patients reporting sleep issues were also more likely, than patients reporting no sleep issues, to spend longer in CAMHS and leave the services when they turned 18-years-old rather than being discharged sooner. Melatonin was

the most prescribed sleep medication; fewer patients were prescribed promethazine. Patients with any psychiatric inpatient admission were significantly more likely to receive a prescription for sleep medication than patients who never had an inpatient admission. Further, patients with a primary mental health diagnosis of psychosis, ADHD, Autism or learning disability were the most likely to be prescribed sleep medication. Despite patients with depression, anxiety and emotional disorders reporting higher levels of sleep issues, they were significantly less likely to receive a sleep medication prescription than patients with ADHD. The present study focused only on sleep medication, however, other psychotropic medications such as anti-depressants and anti-anxiety medication may also have a sedative effect, which may reduce the need for sleep medication in patients with a primary diagnosis of depression and anxiety. Regarding sociodemographic information, Black patients were significantly less likely than White British patients to receive a sleep medication prescription even after controlling for reporting of sleep issues. Together, this suggests that the prescription of pharmacological sleep intervention may demonstrate biases compared to the characteristics of CAMHS patients reporting sleep issues. Furthermore, sleep issues may be associated with longer and more complex service-use. Therefore, the accessibility and allocation of sleep interventions in CAMHS should be explored further in future research.

7.2. Sleep, learning, and memory

The present thesis investigated the role of adolescent sleep for learning, memory, affect and mental wellbeing. In this section, the chapters examining sleep, learning and memory are explored. The following section (Section 7.3.) discusses the thesis research that investigated sleep, affect, and mental health. Chapters 2, 3, and 4 examined the impact of sleep on learning and memory. Chapter 2 examined how school may impact adolescent sleep restriction and the consequences of this for subsequent consolidation. Chapter 3 investigated

the function of sleep for restoring encoding capacity. Chapter 4 also explored sleep *prior* to encoding in relation to emotional memory biases, as well as memory one-year later. The present section focuses on Chapters 2 and 3, as they explore sleep architectural correlates of sleep consolidation and the restoration of encoding capacity. Chapter 4 is discussed in Section 7.3. given the implications of this research for understanding affective wellbeing and mental health. Chapters 2 and 3 both used the well-established "cathedruke" non-word learning paradigm in novel ways. Chapter 2 was the first to use this paradigm with teenagers, whilst Chapter 3 used this paradigm for the first time to examine sleep *prior* to encoding. In the present thesis, neither Chapter 2 nor Chapter 3 observed a significant association between slow-wave sleep and performance on behavioural tasks, despite clear predictions made by the synaptic homeostasis hypothesis.

Chapter 2 was interrupted by COVID-19 lockdowns, and therefore is only useable as a pilot study. However, the multi-night data (diary and actigraphy) demonstrate sleep restriction during school nights with longer sleep on term-time weekends (compensatory sleep) and school holidays (free sleep; sleep need). The sleep architectural correlates of sleep restriction are supportive of laboratory studies, in that slow-wave sleep is relatively preserved whilst other sleep stages are restricted. This suggests that slow-wave sleep may not be an explanatory factor for the relationship between sleep restriction and reduced consolidation, during adolescence at least. The synaptic homeostasis hypothesis suggests that depotentiation of synapses occurs during slow-wave sleep, which increases signal-to-noise ratio, allowing easier retrieval of learned information (i.e., memory consolidation effects, Chapter 2) and prepares the brain for subsequent learning by reducing saturation (e.g., Chapter 3; Tononi & Cirelli, 2003, 2014). Whilst the sample in Chapter 2 is small (underpowered) and the data are limited (e.g., lack of counterbalancing), hence this study is only useable as pilot data, questions regarding the robustness of the relationship between slow-wave sleep and

consolidation have been raised in a review by Cordi and Rasch (2021) and require further study. In Chapter 3, there was also not a significant relationship between cognitive performance and slow-wave sleep *prior* to encoding in an older adolescent sample (18-24-year-olds).

The null results regarding slow-wave sleep, particularly in Chapter 2, may be due to the developmental reduction in slow-wave sleep duration occurring during adolescent development when maturing from childhood to adulthood (e.g., Jenni & Carskadon, 2004; Buchmann et al., 2011). Adults may benefit from existing knowledge allowing greater acquisition of further knowledge, whilst children may compensate for lack of prior knowledge during sleep consolidation, specifically capitalising on greater levels of slow-wave sleep (e.g., James et al., 2019, James et al., 2017). Lexical integration involves integrating new and old information, and therefore requires a level of existing knowledge of the 'old' information. However, adolescents may have less prior knowledge than adults and less slow-wave sleep than children, so it is not clear what mechanisms they would benefit from. Studies that carefully control for prior knowledge (e.g., James et al., 2021) may allow investigation of the relative contributions of slow-wave sleep and prior knowledge throughout development and how this may be associated with integrative consolidation processes. Nonetheless, the lack of relationship between slow-wave sleep and behavioural performance in Chapter 2 and 3, despite predictions made by the synaptic homeostasis hypothesis (based primarily on the adult literature) suggests that the role of slow-wave sleep may not be as robust as previously thought. Cordi and Rasch (2021) have reviewed the literature and found recent replication failures regarding the role of slow-wave sleep for cognition. As such, a movement toward greater open science practices would allow greater scrutiny of theoretical models by reducing publication biases against null results.

Instead, Chapter 3 found that greater Stage 2 frontal sleep spindle density, during a daytime nap preceding learning, was associated with later overnight improvements in lexical integration. Therefore, sleep spindles may function to prepare the brain for later integration of new and old information. Other research has also implicated sleep spindles, rather than slow-wave sleep, in restoring encoding capacity (e.g., Mander et al., 201; Ong et al., 2020). Chapter 3 demonstrates that sleep spindles during prior to encoding may play a role in later consolidation and integration of new and old information. However, Chapter 2, which investigated consolidation, did not analyse sleep spindles due to the small sample size. Another study has found that sleep spindles prior to encoding were associated with restoration of the hippocampus for subsequent learning (Ong et al., 2020). Additionally, sleep spindles have been implicated in the integration of new and old information (Tamminen et al., 2010), similarly to the findings in Chapter 3. Theories such as the active systems consolidation model, suggest that sleep spindles may function to communicate between the neocortex and the hippocampus (Rasch & Born, 2013; Klinzing et al., 2019), allowing interleaving or new and old information during the reorganization and storage of learned information, as proposed by complementary learning systems accounts (e.g., McClelland et al., 1995; Kumaran et al., 2016). Such reorganization via sleep spindles would allow the hippocampus to become restored for subsequent learning, as demonstrated in Ong et al.'s (2020) research. Therefore, the findings in Chapter 3 could be explained by sleep spindles restoring the hippocampus prior to encoding, supporting later integration of new and old information during overnight sleep consolidation.

The maturation of sleep spindles during adolescence may also influence this relationship, and requires further investigation. Sleep spindles appear to mature throughout adolescence (Hahn, 2018; Hoedlmoser, 2020). Therefore, the role of sleep spindles in active systems consolidation accounts may only apply once sleep spindles are matured. As such, the

older adolescent sample (18-24-years-old) in Chapter 3 may have mature sleep spindles, whilst younger samples (e.g., 13-16-years-old in Chapter 2) may show different relationships. Sleep spindles may also be a marker of individual difference. Some research has associated sleep spindle density with 'fluid intelligence' (i.e., cognitive reasoning) but not verbal abilities or short-term memory abilities (Fang et al., 2017). These individual differences in 'fluid intelligence', associated with sleep spindles, may facilitate the integration of new and old information due to greater pattern recognition ability. Sleep spindles may also have individual differences in density (e.g., de Gennaro et al., 2005), which may be seen both in nocturnal and daytime sleep episodes (e.g., Mylonas et al., 2020). This suggests that greater sleep spindle density in daytime naps may also be seen again in overnight sleep, which may benefit learning and memory. However, the within-subjects designs used in Chapter 2 and 3 should have reduced the confounding effects of individual differences.

A large difference between Chapter 2 and Chapter 3 is that they investigated sleep consolidation and restoration of encoding capacity respectively. As such, these processes may have different sleep architectural correlates. Chapter 3 is one of very few to consider the role of sleep *prior* to encoding. The impact of sleep on subsequent learning may be of particular interest during adolescence. For example, if adolescents are receiving insufficient sleep on weekdays, as suggested in Chapter 2 and theories such as the Perfect Storm Model (Crowley et al., 2018), then their brains may not be suitably prepared for encoding during a school day. Interventions such as a daytime nap (Cousins, Wong et al., 2019; Cousins, van Rijn et al., 2019) or delaying school start times (Kelley et al., 2017; Marx et al., 2017) may provide additional or adequate sleep, allowing adolescents to be in an optimal position for learning during the school day. Future theories of sleep and cognition should consider not only the impact of adolescent development, but also the role of sleep for both the consolidation of memories *and* preparation for learning. Despite some limitations of the synaptic homeostasis

hypothesis (as described earlier), its strength is in considering both sleep prior to and after encoding.

The methodology in Chapter 2 could be used as a feasible quasi-experimental alternative to a randomised controlled trial delaying schools start times, as randomised controlled trials are currently not feasible in the UK (Illingworth et al., 2018). Suggestions are made in Section 2.5.3. to increase the likelihood of successful future research using a similar design and methodology to Chapter 2. The research questions in Chapter 3 and Chapter 2 could be combined into an intervention study to investigate whether daytime naps during the school day compensate for sleep restriction and provide benefits for subsequent encoding. Similar research questions have been investigated in experimental designs, which included encoding different factual stimuli before and after the daytime nap (e.g., Cousins, Wong, et al., 2019). A similar training and testing paradigm used in a naturalistic context (e.g., school, polysomnography at home) could be used to investigate the role of sleep for consolidation and restoring encoding capacity. Additionally, neuroimaging (e.g., fMRI) could provide further evidence to inform theories regarding the benefit of sleep for restoring encoding capacity, in a developmental context. For example, Ong et al. (2020) found sleep spindles during a daytime nap restored the hippocampus for subsequent learning, in a sample of young adults (mean age 23.3 years). However, these relationships may be different for younger adolescents who have maturing sleep spindles (Hahn et al., 2018; Hoedlmoser, 2020). Therefore, investigation of these research questions in teenagers may further inform theories such as the complementary learning systems accounts and active systems consolidation models.

Multi-night or multi-episode (e.g., daytime nap and overnight) sleep recordings could also be beneficial for examining the effects of sleep before and after learning, as well as investigating how adolescent sleep restriction may change over the course of a school week.

Chapter 3 demonstrated that sleep spindles prior to encoding were associated with later consolidation benefits for lexical integration, suggesting that recordings preceding (e.g., during the nap) and following encoding (i.e., overnight consolidation) could inform better about the neural processes underlying this relationship. Further, Chapter 2 identified adolescent sleep restriction during school nights on the multi-night recordings (e.g., actigraphy, daily diary) but not on the single night recordings (polysomnography). Therefore, taking multiple consecutive recordings of sleep may be able to inform about cumulative effects of sleep on cognition. Multi-night recordings could also answer further research questions regarding sleep variability or regularity, which have been considered theoretically important during adolescence (e.g., Becker et al., 2017). Additionally, multi-night recordings (e.g., using Dreem headbands, discussed further in Section 7.6.1.) could examine the sleep architectural correlates of adolescent sleep restriction over multiple nights, to further investigate whether findings in laboratory studies (Ong et al., 2016) replicate in naturalistic settings.

7.3. Sleep, affect, and mental health

The latter half of the thesis (Chapters 4 – 6) explored the relationship between sleep, affect, and mental health. Chapters 4 and 5 were conducted online during the COVID-19 lockdowns, and therefore measured sleep subjectively through self-report. Chapter 4 sought to investigate the role of sleep, as well as positive and negative affect, for emotional memory biases. Contrary to the mood-congruent hypothesis (Bower, 1981), higher negative affect was significantly associated with *lower* negative memory. Therefore, it is plausible that inhibitory affective regulation processes may been employed. At Timepoint 2 (one-year follow-up, outside of lockdown conditions), different relationships emerged for habitual sleep and last night's sleep. Poorer habitual sleepers may have selectively consolidated positive, relative to

neutral memories, whilst the opposite was seen for good sleepers. However, whilst not statistically significant, investigating sleep on the night preceding the study demonstrated that good sleepers may have inhibited the recollection of negative memories, whilst poor sleepers may not have been able to use such mechanisms. However, there are methodological limitations associated with conducting this research online (e.g., low-level emotional valence manipulation) that suggests these relationships may be significant with more salient stimuli (e.g., in laboratory studies). The role of inhibition for affective wellbeing and mental health was explored further in Chapter 5. Due to lack of established methodology for collecting data quickly with participants under 18-years-old at the beginning of the pandemic, data for Chapter 4 were collected with 18-24-year-olds only. Later, in Chapter 5, we established a recruitment method that allowed quick data collection with younger participants. Chapter 5 found that participants' ability to inhibit intrusive thoughts (e.g., Harrington & Cairney, 2021) but not reactivity to nor recognition of emotional stimuli (e.g., Goldstein & Walker, 2014) mediated the relationship between poor sleep and poor mental health for both 13-16-year-old and 18-24-year-old adolescents. Given the relationships between poor sleep and mental health (e.g., Chapter 5), Chapter 6 aimed to investigate pharmacological treatment of sleep issues, using a large secondary dataset of medical records from the National Health Service (NHS) CAMHS in South London and Maudsley. Whilst Chapter 6 study took a broad approach to sleep issues, due to limitations in the data available, the results highlighted potential biases in the prescription of sleep medication in CAMHS that may require further investigation.

Similarly to the criticisms of theories of learning and memory (discussed above in Section 7.2.), the theories regarding the role of sleep and mental health also mention mechanisms that may change developmentally. In both Harrington and Cairney's (2021) and Goldstein and Walker's (2014) theories, the role of the prefrontal cortex is emphasised.

However, the prefrontal cortex matures up until the end of adolescence at approximately 24-years-old (Arain et al., 2013; Sawyer et al., 2018). As such, the ability of the prefrontal cortex to inhibit unwanted thoughts and emotional reactivity may be particularly challenging for adolescents. This makes sense, given that research demonstrates adolescents have lower emotion regulation abilities than adults, and that these emotion regulation skills mature developmentally (e.g., Schweizer et al., 2020). Chapter 5 found that both younger and older adolescents demonstrated a significant relationship between poor sleep, lower ability to control thoughts, and greater levels of mental health issues. However, older adolescents (18-24-years-old) reported poorer sleep, greater mental health issues, lower thought control ability, lower positive and higher negative affect than younger adolescents (13-16-year-olds). It is not clear whether this is associated with reporting patterns, social factors at the time of the study (e.g., during COVID-19, older adolescents may have been more isolated if they lived away from family), or brain maturation (e.g., the prefrontal cortex). Future research could consider neuroimaging (e.g., fMRI) to examine the role of brain maturation in the relationships between sleep and mental health proposed by Harrington and Cairney (2021) and Goldstein and Walker (2014).

As well as considering adolescent development, combining our understanding of the relationship between sleep, learning, memory (e.g., Section 7.2., Chapter 2 – 4), affect, and mental health (e.g., Chapters 4 – 6) could also support in informing these theories. Studies investigating the effects of total sleep deprivation have found that negative memories are equally high, but positive and neutral memories are significantly lower when sleep deprived rather than well-rested (Tempesta et al., 2018). This creates a greater ratio of negative to positive or neutral memories, resulting in an overall negative memory bias. Interestingly, current theories may suggest that negative memories would be higher in the presence of poor sleep, for example due to affective tagging in the Affective Tagging and Consolidation model

(Harrington et al., 2017) or amygdala over-reactivity to affective stimuli (Goldstein & Walker, 2014). These theories may explain the findings in Chapter 4. Together, this suggests the importance of considering a range of sleep paradigms, for example total sleep deprivation (e.g., Tempesta et al., 2016, 2018), more naturalistic sleep restriction (e.g., Chapter 2), and subjective sleep quality or perceptions of sleep (e.g., Chapters 4 – 6). It would also be useful to consider the role of sleep preceding (e.g., Chapters 3 and 4) and following (e.g., Chapter 2) encoding or emotional distress. Such research could be causally sensitive and imply the directionality of relationships between sleep, learning, memory, affect, and mental wellbeing.

A meta-analysis of longitudinal research demonstrated that participants with insomnia at baseline were significantly more likely to develop depression symptoms at follow-up (Baglioni et al., 2011). The longitudinal nature of the data suggests that insomnia may precede the onset of depression. However, a systematic review of longitudinal data suggests relationships between poor sleep and depression or anxiety are bidirectional (Alvaro et al., 2013). Goldstein and Walker's (2014) theory does not discern between the role of sleep preceding and following the exposure to emotional stimuli, but does propose that REM sleep functions to recalculate the emotional systems for next day functioning. Harrington and Cairney's (2021) framework suggest the relationships between poor sleep and intrusive memories may accumulate and gradually worsen over time, suggesting the importance of habitual sleep difficulties. Whilst correlation does not equate to causation, and the methodologies of the online studies conducted in this thesis are limited, significant relationships were observed between habitual sleep and general tendencies towards lower thought control abilities and greater symptoms of mental health issues such as depression, anxiety, and stress (Chapter 5). Whereas, whilst not significant, negative emotional memory biases were more likely when participants had poor sleep on *the preceding* night (Chapter 4, Timepoint 2). Potentially this suggests that inhibitory processes are important for emotion

regulation and may be influenced by sleep. Future research investigating sleep preceding and following encoding could further inform theories of the relationship between sleep and mental health via different theoretical mechanisms such as inhibition, consolidation and cognitive biases (e.g., Goldstein & Walker, 2014; Harrington et al., 2017; Harrington & Cairney, 2021). A model is proposed in later in Section 7.5. that may help direct future research questions in this field.

The research questions in Chapters 4 and 5 are important for informing theories and understanding the mechanisms behind the relationship between poor sleep and poor mental health. Such research has implications for designing targeted interventions to support sleep in the prevention or treatment mental health difficulties. Chapter 6 specifically investigated pharmacological sleep intervention in CAMHS. However, due to the secondary nature of the data, and limitations of the existing natural language processing applications, a broad approach was used. Mixed quantitative and qualitative research could investigate clinicians, CAMHS patients, and families' attitudes towards and understanding of sleep medication prescriptions. Furthermore, the present thesis considers pharmacological interventions (i.e., sleep medication) but not non-pharmacological interventions (e.g., CBT for insomnia, sleep hygiene) for poor adolescent sleep. Future research could consider how different sleep interventions may affect learning, memory, mental health, and wellbeing.

Chapter 5 suggests that intrusive thoughts mediate the relationship between poor sleep and mental health symptoms. Harrington and Cairney (2021) suggest that this is a cyclical relationship that continues to exacerbate the difficulties. Identifying the mechanisms for intervention within the relationship between poor sleep and thought control abilities would be important for future research. For example, Harrington and Cairney (2021) suggest that thought substitution may be less impacted by poor sleep, and as such may be an effective intervention for managing intrusive thoughts to break the cycle. It would be of interest to

examine whether treating sleep issues (pharmacologically or non-pharmacologically) would reduce intrusive thoughts, or whether interventions would be required to directly target intrusive thoughts, which may further function to reduce sleep issues. Understanding the directionality of these relationships could support the efficacy of interventions. For example, treating depression has been associated with a reduction in sleep issues for some but not all adolescents (Reynolds et al., 2020). Applying a theoretical stance, and considering the impact of adolescent development, whilst investigating the efficacy of different interventions could also provide more information about the mechanisms of change. Often, potential mechanisms of change in the relationship between adolescent sleep and mental health are considered outside of theoretical frameworks (e.g., Gradisar et al., 2022). Exploring these relationships theoretically could also inform about complex mental health presentations.

The research presented in Chapter 6 originally intended to examine the relationship between sleep issues and self-harm. However, the natural language processing application to extract mentions of self-harm is currently being validated in CAMHS populations. Once this natural language processing application is validated, associations between sleep, medication and self-harm may be examined. Sleep issues and self-harm are transdiagnostic symptoms and as such are sometimes neglected within mental health literature and practice (Ennis et al., 2017; Freeman et al., 2020). Previous research has found that poor sleep is predictive of mental health hospitalisation (Borschmann et al., 2018) and self-harm (Hysing et al., 2015). This has cost implications for the NHS, for example, each episode of self-harm requiring Accident and Emergency treatment cost the NHS an average of £809 in 2017, which is likely much higher now (Tsiachristas et al., 2017). It is also important to continue investigating the associations between sleep and complex mental health presentations, given the risk for life. The study in Chapter 6 also planned to link CRIS data with the National Pupil Database (NPD). Poor sleep is associated with greater levels of school absenteeism (Matos et al., 2016;

Murray et al., 2019) and reduced academic performance (e.g., Enright & Refinetti, 2017; Owens & Weiss, 2017; Phillips et al., 2017). This may be exacerbated in CAMHS samples, as depression is associated with declining grades (Wickersham et al., 2020) and school absenteeism (Finning et al., 2019). Similar relationships are seen for self-harm and absenteeism (Epstein et al., 2019). Thus, the relationship between sleep and mental health may have wider implications for education, which is particularly important for CAMHS patients. Together, this suggests that the relationships between sleep and mental health may also have implications for pedagogical practice, which implores the integration of research investigating sleep, learning, memory, affect, and mental-health.

7.4. Implications for policy and practice

Recently, there has been a movement toward greater integration of mental health and educational services (e.g., Department of Health & Department for Education, 2017). The research presented in this thesis has also been topical with regards to the recent school start time debates (e.g., Chapter 2; Petitions Committee, 2019; Zeichner, 2019) and STOMP-STAMP movements toward the appropriate prescription of medication in paediatric samples (Chapter 6; NHS England & RCPCH, 2018). As such, the research questions in these chapters have implications for policy and practice in mental health and education. The present research has also occurred alongside a recent resurgence in adolescent sleep literature (e.g., Lo & Chee, 2020), which has allowed the development of different theories underpinning adolescent sleep (e.g., Crowley, 2018, Becker, 2017). A strength of the present thesis is that it applies theories of sleep and cognition, that are often explored primarily in adult literature, to development during adolescence. Our current theoretical models for understanding the role of sleep may require further scrutiny to identify their limitations and robustness, with a particular focus on open science. More research could help to create developmentally relevant theoretical models, which could inform our understanding of the

mechanisms behind the relationships between sleep and learning, memory, and mental wellbeing could inform the development of effective interventions.

Despite the importance of sleep, sleep may be insufficient for many adolescents particularly during the school term (e.g., Gradisar et al., 2011, findings from the limited pilot study in Chapter 2). Poor or insufficient sleep may be supported by pharmacological (e.g., sleep medication, Chapter 6) and non-pharmacological (e.g., daytime nap, discussed in Chapter 3; delaying school start times, discussed in Chapter 2) interventions. These interventions require investigation to ensure they are beneficial and accessible to adolescents that need support with sleep. Whilst the present thesis does not specifically research interventions for poor sleep during adolescence, this thesis does identify potential consequences of poor sleep that could inform interventions. For example, the research questions in Chapter 2 are associated with the debate around secondary school start times (Petitions Committee, 2019; Zeichner, 2019), including the extent to which adolescents are sleep restricted in the UK during school nights (e.g., Kelley et al., 2017) and the impact of sleep restriction on learning and memory. However, more high-quality research is required to allow secondary schools to confidently decide to delay school start times, especially given the wide range of barriers that may be experienced with such a drastic change. One study identified that issues associated with the working hours of caregiver(s) and teachers, the school start times of other siblings, and transportation issues may be significant barriers to changes in school start times (Illingworth et al., 2017). This highlights the need for translational research, discussed in the next section (Section 7.4.1.).

Instead, other researchers have suggested that incorporating daytime napping within educational settings may compensate for the sleep restriction adolescents experience due to early school start times (e.g., Cousins, Wong et al., 2019). Daytime naps may have the additional advantage of both consolidating information learned in the morning as well as

preparing the brain for subsequent encoding in the afternoon. For example, Chapter 3 found that immediate recall was significantly more accurate following a daytime nap than an equivalent period of wake. Sleep architectural correlates were not found for this behavioural advantage, and this benefit was not maintained in the next-day follow-up test. However, sleep spindles during the nap were associated with subsequent sleep consolidation effects in the lexical competition task. As such, there may be interactions between pre-encoding sleep and consolidation that require further investigation. Additionally, participants in Chapter 3 woke up at 6 a.m. and remained awake until the nap, and therefore may have been sleep restricted in comparison to the normal sleep patterns of university students. As such, Chapter 3 may also have demonstrated that remedial effects of a daytime nap in compensating for nocturnal sleep restriction, similarly to Cousins, van Rijn et al. (2019). However, the role of a daytime nap in restoring encoding capacity, in the absence of sleep restriction, may show different results. Therefore, daytime naps may only benefit adolescents who receive an insufficient nocturnal sleep duration. Other students may use the time to continue studying, as one study has shown that the short-term memory benefits of a daytime nap compared to continual studying are comparable (Cousins, Wong et al., 2019).

Considering potential interventions to prevent or compensate for insufficient sleep may improve academic attainment and mental health. A meta-analysis conducted by Marx et al. (2017) showed several benefits of delaying school start time, such as longer sleep duration and improved academic attainment. Marx et al. (2017) also found that delayed school start times were associated with fewer symptoms of depression. Chapters 4 and 5 indicate that poor sleep may have implications for worse affective wellbeing and mental health, suggesting delayed school start times may improve these relationships. This may also extend to colleges, universities and the workplace, given that participants in Chapters 4 and 5 included 18-24-year-olds. Sleep may be a useful target given its relationship with both education and mental

health. Poor mental health also has potential consequences for education. For example, higher levels of depression have been associated with declining grades (Wickersham et al., 2020) and lower school attendance has been associated with increased risk of self-harm (Epstein et al., 2019). Together, this suggests that sleep should be prioritised in CAMHS and schools.

Within a CAMHS dataset, Chapter 6 identified that some patients may be more likely to receive sleep medication prescriptions than others. White British patients were significantly more likely than Black patients to receive a sleep medication prescription even after controlling for service use and demographic factors, including reporting of trouble sleeping. This may demonstrate cultural differences in attitudes towards psychotropic medication, alternatively, it may demonstrate racial barriers in receiving sleep medication prescription. Whilst Chapter 6 is unable to inform the reasons behind this relationship, further investigation of this may highlight the need for understanding cultural barriers and training on potential cultural barriers within CAMHS. Patients with a diagnosis of autism or learning/developmental disability, psychosis, or ADHD were prescribed sleep medications at the highest rates in Chapter 6. The STOMP/STAMP initiatives highlight that psychotropic medications in general may be over prescribed to patients with these diagnoses (NHS England & RCPCH, 2018).

Understanding the relationships between poor sleep and mental health may inform about complex presentations and service use. Chapters 4 and 5 explored the mechanisms behind how poor sleep may contribute to poor mental health. Chapter 6 identified that reporting of any trouble sleeping was associated with greater time in the services and lower likelihood of being successfully discharged before turning 18-years-old. Therefore, this suggests that poor sleep may be associated with more complex or enduring mental health presentations, requiring longer lengths of treatment (as discussed in Section 7.3. above). Given the current pressures on the National Health Service (NHS), interventions targeting

sleep issues may prevent the development of mental illness or support faster treatment and improve prognoses. However, these relationships require further investigation in more fine-grained research that targets the specific mechanisms identified in Chapters 4 and 5, and considers the implications for services (as in Chapter 6)

7.4.1. Translational research

Given the potential implications of the research questions posed in this thesis (discussed above in Section 7.4.), future research should conduct translational research. Solari et al. (2020) outline contributing factors that support translational research in the field of reading, which may be applied more widely to the field of educational and mental health practice. Such recommendations include interdisciplinary teams with complementary skills (e.g., methodology, research communication). Throughout my PhD, I have been fortunate to develop skills in many of these areas and to collaborate with practitioners (e.g., psychiatrists, teachers). However, future research may be more successful in wider teams comprised of a range of expertise and experiences. These teams should include co-production with service-users, such as adolescents, families, teachers, and clinicians, which is a key part of the NHS Long-Term Plan (Jeremy, 2019). Chapter 2 was originally intended to collaborate with interested parties such as schools, families, and policy makers, regarding the interpretation and dissemination of the study's findings. Unfortunately, this was not possible due to the COVID-19 pandemic interrupting the study. To produce meaningful and translational research, future research should consider co-production in all stages of the research, which may support with barriers experienced in the current research such as recruitment challenges in Chapter 2. This translational research would have implications for improving our theoretical understanding of adolescent sleep and the implications this may have for policy and practice.

7.5. Proposed model

Together, this thesis has investigated the associations between adolescent sleep and learning, memory, wellbeing, and mental health. The model proposed in Figure 7.1. below aims to combine our current understanding of the cognitive and affective processes between the role of sleep and mental health, whilst considering the impact of adolescent development. The model proposes that poor sleep before a stressor increases negative affect in relation to the stressor (e.g., Minkel et al., 2012). In the presence of good sleep, inhibitory control over recollection of unwanted (e.g., negative) thoughts (e.g., Chapter 5, Harrington & Cairney, 2021) and memories (e.g., Chapter 4) can break the cycle (see Figure 7.2.). When sleep is poor however, inhibitory control may be less possible (e.g., Goldstein & Walker, 2014; Harrington & Cairney, 2021). This later results in greater levels of ruminating (e.g., Harrington & Cairney, 2021) and potentially nightmares, further exacerbating sleep issues the next night.

These sleep issues may reduce consolidation capacity, increasing the need for selective consolidation. Chapter 4 shows that poor habitual sleepers may have selectively consolidated positive, relative to neutral memories, whilst negative memories did not vary with habitual sleep. This may have been an affective regulation mechanism, given our sample tended to report good sleep and low negative affect. In samples with much poorer sleep or higher negative affect, negative information may be selectively consolidated at the expense of both positive and neutral information, resulting in negative emotional memory biases (e.g., Affective Tagging and Consolidation model, Harrington et al., 2017). Lower positive memories could lead to lower positive affect. This lower positive affect is uniquely associated with higher levels of depression symptoms (e.g., tripartite theory, Clark & Watson, 1991). Additionally, this relationship between sleep and positive affect may be more

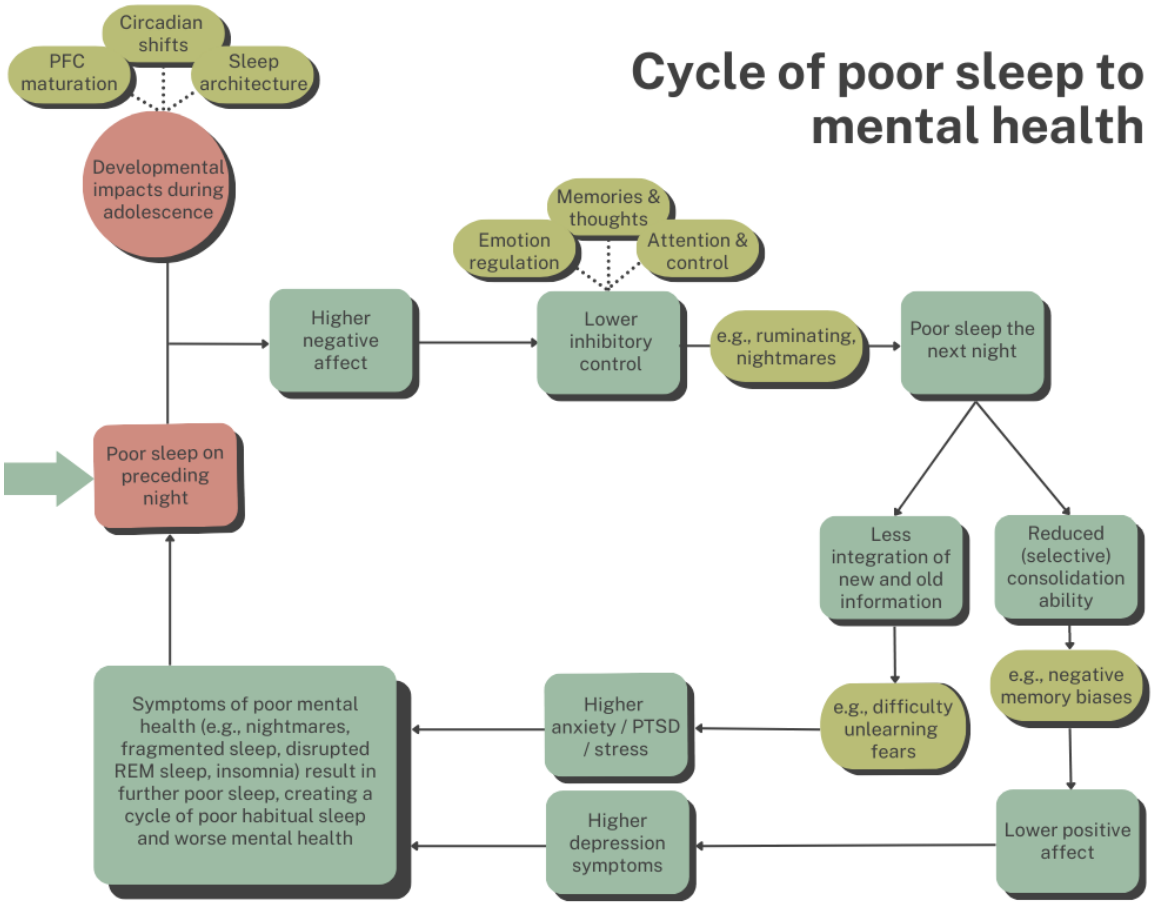
prevalent for younger adolescents (e.g., Chapter 5) and may be stronger than for negative affect during adolescence (e.g., Short et al., 2020).

Poorer sleep may also result in lower sleep spindle density. Fewer sleep spindles may be associated with less integration of new and old information (e.g., Chapter 3, Tamminen et al., 2010). The research presented in Chapter 3 and the research conducted by Tamminen et al. (2010) specifically focuses on *lexical* integration and declarative memory. However, the ability to integrate new and old memories more broadly during consolidation could allow fear extinction. One suggestion is that reconsolidation allows integration of the old fearful memory with new information learning safety, which can result in long term fear extinction and prevent recurrence of the fear in humans (Schiller et al., 2010). As fearful (old) and safety (new) memories could also be considered declarative memories, they may also be supported by integrative consolidation processes facilitated by sleep spindles. For example, Wilhelm et al. (2017) found that adolescents with social anxiety had lower spindle density, resulting in greater reactivity to negative stimuli the next day. However, emotional memories may be stored in the amygdala rather than distributed broadly across the neocortex (as in standard declarative memory); nonetheless, emotional memories interact with the hippocampus as in Harrington and Cairney's (2021) model. The difficulty unlearning fears through integrative processes during consolidation may result in higher symptoms of anxiety, post-traumatic stress disorder, and stress (e.g., increased physiological stress responses). These relationships may further be impacted by adolescent development. A recent review has found that adolescents are more sensitive to associative learning for social and emotional information but show less specificity so are more likely to generalise this information (Towner et al., in press). As such, fears may be generalised to a greater extent in adolescents with poor sleep, and fewer sleep spindles may reduce opportunities to unlearn these fears.

The relationship between sleep and mental health continues in a cycle, further worsening sleep, affective regulation processes, and mental health. Poor sleep should be considered, not just with regards to total sleep time, but also to sleep quality more broadly. For example, lower sleep efficiency has been associated with worse mental health and wellbeing, particularly for autistic adults (Henderson et al., 2023). This suggests that neurodevelopmental conditions could also be considered in the proposed relationships in the future. Furthermore, longer sleep onset latency was associated with higher levels of anxiety at bedtime for children (approximately 8-years-old on average) in the UK during the COVID-19 pandemic (Knowland et al., 2022). However, Knowland et al. (2022) did not find a relationship between anxiety and total sleep time, which they suggest may be due to lack of early school start times during the lockdowns.

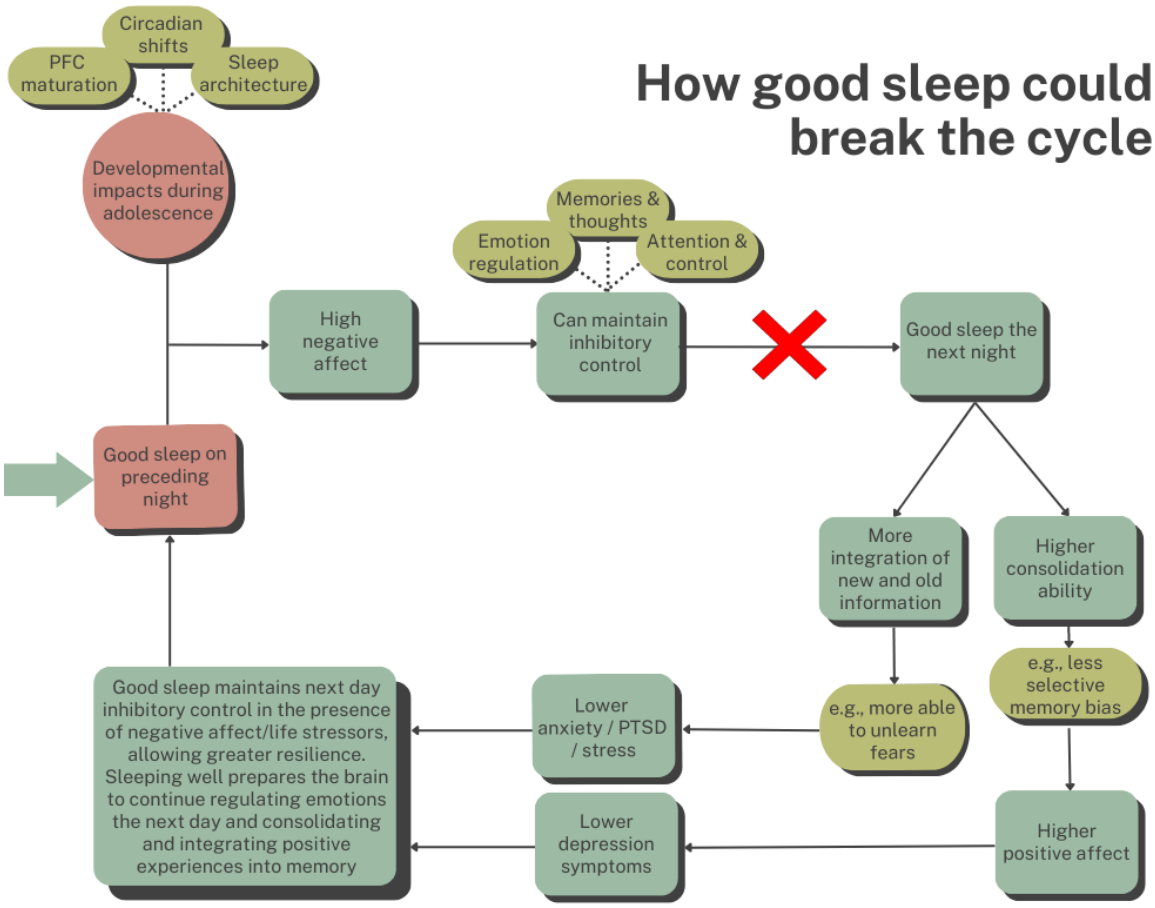
Adolescents may be particularly vulnerable in the proposed cycle, given changes in sleep architecture such as the maturation of sleep spindles and slow-wave sleep (Section 1.4.). Adolescent changes in circadian rhythms shifting towards evening chronotypes and resulting in less sleep (e.g., Section 1.4.1.) particularly on school nights (e.g., Chapter 2) may also increase vulnerability to these relationships, with consequences for education too. Additionally, prefrontal cortex maturation may influence inhibitory control mechanisms proposed in the model. This preliminary model considers both positive and negative affect, as well as developmental impacts on the relationships, and the role of sleep before and after potential stressors. Future research examining the sleep architectural correlates of different pathways in the model could provide further information to update our theoretical knowledge of these mechanisms.

Figure 7.1. Proposed model of how poor sleep, negative affect, inhibitory control, overnight consolidation and integration, positive affect, and mental health may interact in a cycle



Note: In Figure 7.1. and 7.2., pink indicates the possible point at which sleep issues may onset. Blue indicates the factors that may contribute toward and maintain difficulties within the cycle. Green indicates examples of the way the factors may manifest.

Figure 7.2. Diagram of how good sleep can break the cycle described in the proposed model in Figure 7.1.



7.6. Strengths and limitations of the studies

The research presented in this thesis benefits from using open-science practices. Chapter 3 is in the format of a registered report, which reduces the likelihood of publication biases against null hypotheses, allowing greater scrutiny of theoretical concepts. The online studies (Chapters 4 and 5) were also pre-registered on the Open Science Framework. This allows other researchers to access the data, analyses scripts, and materials used for these studies. These open science practices help to combat the replication crisis, which is particularly important in the sleep and cognition literature. A recent review by Cordi and Rasch (2021) outlined replication failures and null hypotheses regarding the role of (slow wave) sleep for

active memory consolidation. This highlights the importance of more open science practices being used within the field of sleep and cognition and shows the strengths of the present research in applying such open science practices. The majority of the studies included in this thesis benefit from large, well-powered sample sizes (discussed further in Section 7.6.3.). The methodologies used in the current study have different strengths and applicability to future research. Specifically, the different measures of sleep are discussed in Section 7.6.1. below.

7.6.1. Measuring sleep

Throughout the thesis, a mixture of subjective and objective sleep measurement has been used. This is considered a strength, as objective and subjective measures may inform about different phenomena (Dewald et al., 2010). Polysomnography is considered the gold-standard for sleep measurement, as it is objective and can validly measure sleep architecture. Polysomnography was used in participants' own homes in Chapter 2, as this method is found to be measure sleep more naturalistically than polysomnography conducted in less familiar environments such as hospitals (Bruyneel et al., 2011) or the laboratory. However, the requirement for a researcher to attend participants' homes added recruitment challenges, particularly following the COVID-19 pandemic. There was also a large amount of data lost due to lack of researcher observation, for example, early removal or electrodes or equipment failures during the sleep period. In Chapter 3, laboratory-based polysomnography was used, which had the advantage of greater experimental control and less data attrition but was perhaps less naturalistic than sleep at home. The collection of polysomnographic data is typically time intensive, requiring a trained researcher to apply the equipment and analyse the data. The equipment itself is also expensive, and subscriptions are often required to score the data (e.g., REMLogic). With recent open science movements, SpiSOP (Weber, 2013) is a freely available open science resource that can be used to score polysomnographic data.

However, as our data were recorded on REMLogic, they were also scored in REMLogic with the aid of algorithms to examine inter-rater reliability. Due to the time required and financial expense associated with polysomnography data collection and analyses, polysomnographic data is often limited to a single session or a small number of recordings, as was the case in Chapters 2 and 3.

However, collecting data over multiple time points has highlighted that multiple night sleep recording may be better positioned to investigate the mechanisms of interest. For example, Chapter 2 observed sleep restriction in multi-night recordings (actigraphy and daily diary), but not in single nights of polysomnography. Additionally, Chapter 3 found that sleep spindles prior to encoding were associated with later overnight consolidation in the lexical competition task. Therefore, taking recordings over multiple periods of sleep may have been better placed to inform about the sleep architectural correlates associated with the interaction between encoding and consolidation (Chapter 3) or sleep restriction during a school week (Chapter 2), which may provide further information for theoretical development. Recent technological advances, such as the Dreem headband (e.g., Guillot et al., 2019; Thorey et al., 2019), allow for the collection of multiple nights of sleep architectural data with greater ease. Dreem headbands are currently available for commercial and research purposes. These headbands do not require a trained researcher to apply or monitor the headband, as this can be done remotely, by participants themselves. Dreem headbands were considered as an alternative for polysomnography in Chapter 2, before the study was cancelled due to COVID-19. However, Dreem headbands are limited as they do not capture posterior sleep spindles, and therefore may not be appropriate for all research questions. Future research would need to validate the efficacy of the Dreem headbands, or other dry electrode polysomnography devices for use with adolescents. Some research already demonstrates that Dreem headband recordings may be comparable to polysomnography (Thorey et al., 2019). The use of Dreem

headbands could also allow for larger sample sizes and multi-night recordings in future research.

Alternatively, actigraphy devices can be used to objectively record multiple nights of sleep. Actigraphs are cheaper than polysomnography and can be easily delivered to participants, allowing reduced researcher contact. These devices are also available commercially or for research and can collect data on heart rate, light, movement, and sleep. Actigraphy devices are unable to inform about sleep architectural correlates, which were an important part of the research questions for Chapter 2 and 3. The algorithms used to score actigraphy data can also show some issues in discerning sleep and wakeful inactivity, often over- and under-estimating different variables, and performing differently for adolescents than adults (Quante et al., 2018). Therefore, the present research benefitted from including a subjective sleep diary alongside the objective actigraphy data. The combination of subjective and objective sleep measurement is a strength of the present research, and is recommended for future research, as they capture different phenomena (e.g., Dewald et al., 2010).

If no contact with participants is possible, then self-report measures of sleep may be most appropriate. The Pittsburgh Sleep Quality Index (PSQI, Buysse et al., 1989) was used across multiple chapters (e.g., Chapters 4 and 5). The PSQI can be laborious to score, due to calculating component scores and equations used to determine some variables (e.g., sleep efficiency), which can make it difficult to use for larger samples. The R code used in the present thesis, may be of use to future researchers to allow them to quickly and efficiently gather and score PSQI data from larger samples (e.g., online research). The code is openly available on the Open Science Framework along with the data for Chapters 4 and 5. Coding in R also supports open science practices, as the code provides evidence of data cleaning and analysis decisions and processes. Habitual sleep measures such as the Pittsburgh Sleep Quality Index cannot assess nightly changes in sleep. Instead, sleep diaries may be used.

However, multi-night reporting on sleep diaries could result in greater attrition of data or recruitment challenges as repeated measures can be more time-intensive for participants. In the current thesis, an amended version of the Pittsburgh Sleep Quality Index questionnaire was created to measure last night's sleep, similarly to a sleep diary (Appendix 4.1.). Future research would be needed to validate the psychometric properties of this questionnaire and its potential use as a low intensity daily diary. Using a consistent daily diary questionnaire in wider research could support meta-analyses and integration of data from multiple studies. However, there are limitations to self-report measures, in that good sleepers have been found to over-report sleep duration, whilst participants with insomnia under-reported sleep duration, in comparison to polysomnographically measured sleep (Benz et al., 2023). Therefore, this new questionnaire would also need to be validated in relation to objective sleep measures. Nonetheless, self-report can allow for online research, which has many benefits such as quick recruitment and large sample-sizes (e.g., Chapters 4 and 5). The present studies were conducted only with samples currently residing in the UK, due to varying lockdown conditions internationally. However, large-scale online research can also collect data internationally, which can inform our understanding of the relationship between sleep and cognition across the world (Chee & Willoughby, 2023).

7.6.2. The impact of COVID-19

The COVID-19 pandemic disrupted the present research (e.g., Chapter 2) and prevented further face-to-face research that had been planned, such as a study investigating sleep variability using actigraphy devices, and research investigating the impact of chronotype and time of day on cognitive performance. Instead, research was conducted online, using self-reported sleep and shorter experimental paradigms to maintain participants attention. Online studies are not recommended to be longer than 20-30 minutes (Revilla & Höhne, 2020).

Therefore, the experimental paradigms used in Chapters 4 and 5 were required to be shorter, with fewer items and repetitions, than may be expected in laboratory studies. For example, Antonenko et al.'s (2016) study used hundreds of items with multiple repetitions in a cognitive battery of tests. In the present thesis, the non-word learning paradigm used in Chapter 2 and Chapter 3 had 22 items with multiple repetition blocks (e.g., five blocks with six repetitions per item in the phoneme monitoring training task). The "cathedruke" non-word learning tasks took an average of 36.84 minutes per training session (of which there were two) and 21.81 minutes per testing session (of which there were four) when delivered face-to-face. Instead, shorter experimental paradigms were used in Chapter 4 (30 word-pairs over two encoding and recall blocks) and Chapter 5 (30 trials over eight blocks). The paradigms used in Chapters 4 and 5 intended to elicit emotional reactions from participants in an ethical manner. However, they may not have been salient or arousing enough, for example, the manipulation check was not significant in Chapter 4. More emotionally arousing stimuli could be used in laboratory research (e.g., The International Affective Picture System; Lang et al., 1997). A new open access affective image system has been developed for online and laboratory-based studies (Kurdi et al., 2017). However, due to the additional stressors and potential isolation caused by the COVID-19 pandemic, it would not have been ethical to use such strongly valenced stimuli online at the time the present research was conducted. Online studies with emotionally distressing content are ethnically challenging, due to lack of researchers' ability to monitor participants wellbeing, and participants right to withdraw at any point, which may involve them skipping the research debrief or ignoring supportive resources. Therefore, the emotional valence of the paradigms used in the current study was not particularly strong. As such, this will have limited the ability of these studies to draw conclusions regarding the impact of sleep on interpreting emotionally valent information.

7.6.3. The samples in the current thesis

Whilst Chapter 2's sample (*N* = 11) was limited by COVID-19 lockdowns preventing further data collection, the other studies presented in this thesis benefit from large sample sizes. Chapter 3 (*N* = 45) was adequately powered, based on the conducted power analyses. We were able to collect and analyse data with large samples in Chapters 4 (*N* = 247) and 5 (*N* = 118, 136), due to the online nature of the studies. The largest sample in this thesis was in Chapter 6 (*N* = 7167), which used a secondary NHS dataset. These large-scale studies are useful in supporting our understanding of sleep, as the laboratory-based studies in previous research often use smaller sample sizes. For example, in a meta-analysis of research examining sleep consolidation of emotional memories, the sample size of the included studies was an average of 41 (*SD* = 14, range = 23 – 86; Lipinska et al., 2019), whilst our sample in Chapter 4 is approximately six times larger (at Timepoint 1, at least). Additionally, as described in Section 7.6.1., future research could maximise sample sizes and still investigate sleep architectural correlates, using new sleep methodologies.

At the beginning of the COVID-19 lockdown, there was not an established methodology to collect data with participants aged under 18-years-old, other than via social media or word-of-mouth, which can be time-consuming to recruit. There were also challenges in recruiting adolescents in the first empirical study of this thesis (Chapter 2) prior to COVID-19. A pragmatic decision was made to ease recruitment difficulties by collecting data with samples aged 18-24-years. The present thesis takes a broad definition of adolescence (e.g., 10-24-years-old; Sawyer et al., 2018), therefore, 18-24-year-olds are within the latter stage of adolescent development. A large part of the thesis (e.g., Chapter 3, 4, and 5) focuses on older adolescents (18-24-years-old), who are often living as young adults, with greater independence from parent(s)/caregiver(s) and outside of mandatory education. As such, the

experience of older adolescents examined in Chapters 3 – 5 may be very different from younger participants (e.g., teenagers) examined in Chapters 2, 5 and 6.

The samples recruited for the experimental research in this thesis (Chapters 2 – 5) were aged 13-24-years-old. These chapters did not collect data with participants aged 10-12-years-old, partially due to the nature of online research (e.g., Chapters 4 and 5), but also as younger participants may experience fewer difficulties with sleep. Typically, online platforms have a minimum age of 13-years-old, with parental consent; however more than half of young people have a social media profile before this age (Ofcom, 2020). Therefore, whilst there may be regulatory restrictions on online activity for adolescents under 13-years-old, they may still have online literacy skills. Younger adolescents (10-12-years-old) may be pre-pubertal or in earlier stages of pubertal development and therefore, may not experience sleep restriction to the same degree as older adolescents whose circadian rhythms may be more disrupted by pubertal development. However, some research has demonstrated that weekend compensatory sleep is evident as early as 9-years-old (Thorleifsdottir et al., 2002). Therefore, future research with 10-12-year-olds may be of interest to investigate how relationships between sleep and cognition develop throughout the wide span of adolescence (10-24-years-old).

The age range of 13-16-years-old was selected as these students are typically in Key Stage 2, making the research relevant to education (e.g., Chapter 2), whilst 18-24-year-olds were sampled for practicalities of recruitment (e.g., Chapter 4). It is important to note that these samples excluded participants aged 17-years-old from the majority of the research conducted within the current thesis. Given the importance of the research questions explored in this thesis, for CAMHS and schools, it is important to acknowledge that adolescents in CAMHS and mandatory schooling are under 18-years-old. Therefore, future research with 17-year-olds would also be relevant. The sample in Chapter 6, included data from patients

aged from 3.46 – 18-years-old, with a median age of 13.28 years old. Therefore, whilst the sample does include adolescents aged 10-18-years-old, it also includes children under the age of 10-years-old.

7.7. Conclusion

The studies presented in this thesis contribute to the existing literature in three main ways:

1) Sleep spindles appear to be important for preparing the brain to integrate new and old information in subsequent periods of overnight consolidation (Chapter 3). Whilst slow-wave sleep may be preserved, even in the presence of restricted sleep duration (Chapter 2).
2) The ability to inhibit or control thoughts and memories may be facilitated by good sleep (Chapters 4 & 5). These inhibitory processes may be a mechanism explaining the relationship between poor sleep and mental health, which would have implications for our understanding of the development and treatment of mental health disorders.
3) The relationship between sleep and positive affect appears to effect teenagers more than older adolescents (Chapter 5), which has implications for understanding adolescent depression.

A model is proposed in Section 7.5., which combines these conclusions and integrates them with existing theory and previous research to explain the relationship between poor sleep and mental health. This model considers both the function of sleep in restoring the brain and consolidating memories, and how this can impact learning, memory, and affect, with consequences for mental health. Future research would benefit from taking a theoretically driven approach focusing on producing translational research to investigate the potential

mechanisms behind these relationships. Such research could inform the design and implementation of interventions to support adolescent sleep.

References

Adan, A., & Admirall, H. (1991). Horne & Östberg morningness-eveningness questionnaire: a reduced scale. *Personality and Individual Differences*, *12*(3), 241–253.

Ali Memon, M., Ting, H., Ramayah, T., Chuah, F., & Cheah, J.-H. (2017). A review of the methodological misconceptions and guidelines related to the application of structural equation modeling: A malaysian scenario. *Journal of Applied Structural Equation Modeling*, *1*(1), 1–8.

Alimoradi, Z., Broström, A., Tsang, H. W. H., Griffiths, M. D., Haghayegh, S., Ohayon, M. M., Lin, C. Y., & Pakpour, A. H. (2021). Sleep problems during COVID-19 pandemic and its' association to psychological distress: A systematic review and meta-analysis. *EClinicalMedicine*, *36*. https://doi.org/10.1016/j.eclinm.2021.100916

Alimoradi, Z., Gozal, D., Tsang, H. W. H., Lin, C. Y., Broström, A., Ohayon, M. M., & Pakpour, A. H. (2022). Gender-specific estimates of sleep problems during the COVID-19 pandemic: Systematic review and meta-analysis. *Journal of Sleep Research*, *31*(1), 1–16. https://doi.org/10.1111/jsr.13432

Alvaro, P. K., Roberts, R. M., & Harris, J. K. (2013). A systematic review assessing bidirectionality between sleep disturbances, anxiety, and depression. *Sleep*, *36*(7), 1059–1068. https://doi.org/10.5665/sleep.2810

Analysis Inn (2020). How to calculate average variance extracted and composite reliability. Retrieved from: https://www.analysisinn.com/post/how-to-calculate-average-variance-extracted-and-composite-reliability/

Antonenko, D., Diekelmann, S., Olsen, C., Born, J., & Mölle, M. (2013). Napping to renew learning capacity: Enhanced encoding after stimulation of sleep slow oscillations. *European Journal of Neuroscience*, *37*(7), 1142–1151. https://doi.org/10.1111/ejn.12118

Anwyl-Irvine, A. L., Massonié, J., Flitton, A., Kirkham, N. Z., & Evershed, J. K. (2019). Gorilla in our midst: an online beahvioural experiment builder. *Behavior Research Methods*. https://doi.org/https://doi.org/10.3758/s13428-019-01237-x

Arain, M., Lina, M. H., Mathur, J. P., Afsha, W. N., Ranbir, R., Sushil, S., & Sharma. (2013). Maturation of the adolescent brain. *Neuropsychiatric Disease and Treatment*, *9*, 449–461. https://doi.org/http://dx.doi.org/10.2147/NDT.S39776

Arrona-Palacios, A., García, A., & Valdez, P. (2015). Sleep-wake habits and circadian preference in Mexican secondary school. *Sleep Medicine*, *16*(10), 1259–1264. https://doi.org/10.1016/j.sleep.2015.05.026

Asarnow, J. R., Bai, S., Babeva, K. N., Adrian, M., Berk, M. S., Asarnow, L. D., Senturk, D., Linehan, M. M., & McCauley, E. (2020). Sleep in youth with repeated self-harm and high suicidality: Does sleep predict self-harm risk? *Suicide and Life-Threatening Behavior*, *50*(6), 1189–1197. https://doi.org/10.1111/sltb.12658

Bagley, E. J., Kelly, R. J., Buckhalt, J. A., & El-Sheikh, M. (2015). What keeps low-SES children from sleeping well: The role of presleep worries and sleep environment. *Sleep Medicine*, *16*(4), 496–502. https://doi.org/10.1016/j.sleep.2014.10.008

Baglioni, C., Battagliese, G., Feige, B., Spiegelhalder, K., Nissen, C., Voderholzer, U., Lombardo, C., & Riemann, D. (2011). Insomnia as a predictor of depression: A meta-analytic evaluation of longitudinal epidemiological studies. In *Journal of Affective Disorders*. https://doi.org/10.1016/j.jad.2011.01.011

Baglioni, C., Nanovska, S., Regen, W., Spiegelhalder, K., Feige, B., Nissen, C., Reynolds, C. F., & Riemann, D. (2016). Sleep and mental disorders: A meta-analysis of polysomnographic research. *Psychological Bulletin*, *142*(9), 969–990. https://doi.org/10.1037/bul0000053

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Barrett, J. R., Tracy, D. K., & Giaroli, G. (2013). To sleep or not to sleep: A systematic review of the literature of pharmacological treatments of insomnia in children and adolescents with attention-deficit/hyperactivity disorder. *Journal of Child and Adolescent Psychopharmacology*, *23*(10), 640–647. https://doi.org/10.1089/cap.2013.0059

Baum, K. T., Desai, A., Field, J., Miller, L. E., Rausch, J., & Beebe, D. W. (2014). Sleep restriction worsens mood and emotion regulation in adolescents. *The Journal of Child Psychology and Psychiatry*, *2*(55), 180–190. https://doi.org/10.1111/jcpp.12125

Becker, S. P., Langberg, J. M., & Byars, K. C. (2015). Advancing a biopsychosocial and contextual model of sleep in adolescence: A review and introduction to the special issue. *Journal of Youth and Adolescence* (Vol. 44, Issue 2, pp. 239–270). https://doi.org/10.1007/s10964-014-0248-y

Becker, S. P., Schindler, D. N., Holdaway, A. S., Tamm, L., Epstein, J. N., & Luebbe, A. M. (2019). The Revised Child Anxiety and Depression Scales (RCADS): Psychometric evaluation in children evaluated for ADHD. *Journal of Psychopathology and Behavioral Assessment*, *41*(1), 93–106. https://doi.org/10.1007/s10862-018-9702-6

Becker, S. P., Sidol, C. A., Van Dyk, T. R., Epstein, J. N., & Beebe, D. W. (2017). Intraindividual variability of sleep/wake patterns in relation to child and adolescent functioning: A systematic review. *Sleep Medicine Reviews*, *34*, 94–121. https://doi.org/10.1016/j.smrv.2016.07.004

Belenky, G., Wesensten, N. J., Thorne, D. R., Thomas, M. L., Sing, H. C., Redmond, D. P., Russo, M. B., & Balkin, T. J. (2003). Patterns of performance degradation and

restoration during sleep restriction and subsequent recovery: A sleep dose-response study. *Journal of Sleep Research*, *12*(1), 1–12. https://doi.org/10.1046/j.1365-2869.2003.00337.x

Ben Simon, E., Vallat, R., Barnes, C. M., & Walker, M. P. (2020). Sleep loss and the socio-emotional brain. *Trends in Cognitive Sciences*, *24*(6), 435–450. https://doi.org/10.1016/j.tics.2020.02.003

Benz, F., Riemann, D. Domschke, K., Spiegelhalder, K., Johann, A. F., Marshall, N. S., & Feige, B. (2023). How many hours do you sleep? A comparison of subjective and objective sleep duration measures in a sample of insonia patients and good sleepers. *Journal of Sleep Research*, 1-13. DOI: 10.111/jsr.13802

Berger, R. H., Miller, A. L., Seifer, R., Cares, S. R., & LeBourgeois, M. K. (2012). Acute sleep restriction effects on emotion responses in 30- to 36-month-old children. *Journal of Sleep Research, 21*(3), 235-246. DOI:10.1111/j.1365-2869.2011.00962.x

Berres, S., & Erdfelder, E. (2021). The sleep benefit in episodic memory: An integrative review and a meta-analysis. *Psychological Bulletin, 147*(12), 1309–1353. https://doi.org/10.1037/bul0000350

Blume, C., Schmidt, M. H., & Cajochen, C. (2020). Effects of the COVID-19 lockdown on human sleep and rest-activity rhythms, *Current Biology, 30*(14). https://doi.org/10.1016/j.cub.2020.06.021

Borbély, A. A. (1982). Two-process model of sleep regulation. *Human Neurobiology*, *1*, 195–204. https://doi.org/10.1007/978-3-540-29678-2_6166

Borbély, A. A., Daan, S., Wirz-Justice, A., & Deboer, T. (2016). The two-process model of sleep regulation: A reappraisal. *Journal of Sleep Research*, *25*(2), 131–143. https://doi.org/10.1111/jsr.12371

Born, J., & Wilhelm, I. (2012). System consolidation of memory during sleep. *Psychological Research*, *76*(2), 192–203. https://doi.org/10.1007/s00426-011-0335-6

Borschmann, R., Stark, P., Prakash, C., & Sawyer, S. M. (2018). Risk profile of young people admitted to hospital for suicidal behaviour in Melbourne, Australia. *Journal of Paediatrics and Child Health*, *54*(11), 1213–1220. https://doi.org/10.1111/jpc.13938

Bower, G. H. (1981). Mood and memory. *Behaviour Research and Therapy*, *36*(2), 129–148. https://doi.org/10.1016/0005-7967(87)90052-0

Bradley, M. M., & Lang, P. L. (1999). *Affective Norms for English Words (ANEW): Instruction Manual and Affective Ratings* (pp. 1–48). The Center for Research in Psychophysiology. https://e-lub.net/media/anew.pdf

Brand, S., Schilling, R., Ludyga, S., Colledge, F., Bahmani, D. S., Holsboer-Trachsler, E., Pühse, U., & Gerber, M. (2019). Further evidence of the zero-association between symptoms of insomnia and facial emotion recognition—results from a sample of adults in their late 30s. *Frontiers in Psychiatry*, *10*(JAN), 1–9. https://doi.org/10.3389/fpsyt.2018.00754

Brauer, M., & Curtin, J. J. (2018). Linear mixed-effects models and the analysis of nonindependent data: A unified framework to analyze categorical and continuous independent variables that vary within-subjects and/or within-items. *Psychological Methods*, *23*(3), 389–411. https://doi.org/10.1037/met0000159

Brown, V. A. (2020). An introduction to linear mixed effects modeling in R. *PsyArXiv Preprints*, 1–46. https://doi.org/10.31234/osf.io/9vghm

Bruyneel, M., Sanida, C., Art, G., Libert, W., Cuvelier, L., Paesmans, M., Sergysels, R., & Ninane, V. (2011). Sleep efficiency during sleep studies: Results of a prospective study comparing home-based and in-hospital polysomnography. *Journal of Sleep*

Research, *20*(1 PART II), 201–206. https://doi.org/10.1111/j.1365-2869.2010.00859.x

Buchmann, A., Ringli, M., Kurth, S., Schaerer, M., Geiger, A., Jenni, O. G., & Huber, R. (2011). EEG sleep slow-wave activity as a mirror of cortical maturation. *Cerebral Cortex*, *21*(3), 607–615. https://doi.org/10.1093/cercor/bhq129

Buckhalt, J. A., El-Sheikh, M., & Keller, P. (2007). Children's sleep and cognitive functioning: Race and socioeconomic status as moderators of effects. *Child Development*, *78*(1), 213–231.

Buysse, D. J., Reynolds, C. F., Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Research, 28,* 193–213.

Cairney, S. A., Durrant, S. J., Power, R., & Lewis, P. A. (2015). Complementary roles of slow-wave sleep and rapid eye movement sleep in emotional memory consolidation. *Cerebral Cortex, 25,* 1565 – 1575.

Carskadon, M. A., Vieira, C., & Acebo, C. (1993). Association between puberty and delayed phase preference. *Sleep* (Vol. 16, Issue 3). https://academic.oup.com/sleep/article-abstract/16/3/258/2749376

Chatburn, A., Lushington, K., & Kohler, M. J. (2014). Complex associative memory processing and sleep: A systematic review and meta-analysis of behavioural evidence and underlying EEG mechanisms. *Neuroscience and Biobehavioral Reviews*, *47*, 646–655. https://doi.org/10.1016/j.neubiorev.2014.10.018

Chee, M. W. L., & Willoughby, A. R. (2023). The value of large-scale studies of sleep and cognition. *Trends in Neurosciences*, *46*(4), 255-256.

Chorpita, B. F., Yim, L. M., Moffit, C. E., Umemoto, L. A., & Francis, S. E. (2000). Assessment of symptoms of DSM-IV anxiety and depression in children: A Revised

Child Anxiety and Depression Scale. *Behaviour Research and Therapy*, 38, 835-855.

Clark, L. A., & Watson, D. (1991). Tripartite model of anxiety and depression: Psychometric evidence and taxonomic implications. *Journal of Abnormal Psychology*, *100*(3), 316–336. https://doi.org/10.1037//0021-843x.100.3.316

Cliffe, B., Croker, A., Denne, M., Smith, J., & Stallard, P. (2020). Digital cognitive behavioral therapy for insomnia for adolescents with mental health problems: Feasibility open trial. *JMIR Mental Health*, *7*(3). https://doi.org/10.2196/14842

Committee on Toxicity. (2001). *COT statement on the reproductive effects of caffeine. London: Committee on Toxicity of Chemicals in Food, Consumer Products and the Environment.* https://cot.food.gov.uk/cotstatements/cotstatementsyrs/cotstatements2001/caffeine

Cordi, M. J., & Rasch, B. (2020). How robust are sleep-mediated memory benefits? *Current Opinion in Neurobiology*, *67*, 1–7. https://doi.org/10.1016/j.conb.2020.06.002

Cortese, S., Faraone, S. V., Konofal, E., & Lecendreux, M. (2009). Sleep in children with attention-deficit/hyperactivity disorder: meta-analysis of subjective and objective studies. *Journal of the American Academy of Child and Adolescent Psychiatry*. https://doi.org/10.1097/CHI.0b013e3181ae09c9

Cousins, J. N., Sasmita, K., & Chee, M. W. L. (2018). Memory encoding is impaired after multiple nights of partial sleep restriction. *Journal of Sleep Research*, *27*, 138–145. https://doi.org/10.1111/jsr.12578

Cousins, J. N., van Rijn, E., Ong, J. L., Wong, K. F., & Chee, M. W. L. (2019). Does splitting sleep improve long-term memory in chronically sleep deprived adolescents? *Npj Science of Learning*, *4*(1), 8. https://doi.org/10.1038/s41539-019-0047-z

Cousins, J. N., Wong, K. F., & Chee, M. W. L. (2019). Multi-night sleep restriction impairs long-term retention of factual knowledge in adolescents. *Journal of Adolescent Health*, *65*(4), 549–557. https://doi.org/10.1016/j.jadohealth.2019.04.030

Cousins, J. N., Wong, K. F., Raghunath, B. L., Look, C., & Chee, M. W. L. (2019). The long-term memory benefits of a daytime nap compared with cramming. *Sleep*, *42*(1), 1–7. https://doi.org/10.1093/sleep/zsy207

Crawford, J. R., & Henry, J. D. (2004). The Positive and Negative Affect Scale (PANAS): Construct validity, measurement properties and normative data in a non-clinical sample. *British Journal of Clinical Psychology, 43*, 245–265.

Crowley, S. J., Cain, S. W., Burns, A. C., Acebo, C., & Carskadon, M. A. (2015). Increased sensitivity of the circadian system to light in early/mid-puberty. *Journal of Clinical Endocrinology and Metabolism*. https://doi.org/10.1210/jc.2015-2775

Crowley, S. J., Wolfson, A. R., Tarokh, L., & Carskadon, M. A. (2018). An update on adolescent sleep: New evidence informing the perfect storm model. *Journal of Adolescence*, *67*, 55–65. https://doi.org/10.1016/j.adolescence.2018.06.001

Cummings, C., Canadian Paediatric Society, Community Paediatrics Committee (2012). Melatonin for the management of sleep disorders in children and adolescents. *Paediatrics & Child Health,* 17(6), 331 – 333. https://doi.org/10.1093/pch/17.6.331

Curcio, G., Ferrara, M., & De Gennaro, L. (2006). Sleep loss, learning capacity and academic performance. *Sleep Medicine Reviews*, *10*(5), 323–337. https://doi.org/10.1016/j.smrv.2005.11.001

Darden, M., Espie, C. A., Carl, J. R., Henry, A. L., Kanady, J. C., Krystal, A. D., & Miller, C. B. (2021). Cost-effectiveness of digital cognitive behavioral therapy (Sleepio) for insomnia: A Markov simulation model in the United States. *Sleep*, 1-12, doi.10.1093/sleep/zsaa223

Davis, M. H., Di Betta, A. M., Macdonald, M. J. E., Gaskell, M. G., (2009). Learning and consolidation of novel spoken words. *Journal of Cognitive Neuroscience, 21*(4), 803-820. https://doi.org/10.1016/j.sleep.2019.05.015

Davis, C. J. (2005). N-watch: A program for deriving neighborhood size and other psycholinguistic statistics. *Behavior Research Methods*, *37*(1), 65–70. https://doi.org/10.3758/BF03206399

Diagnostic and Statistical Manual of Mental Disorders: DSM-5. 5th ed., American Psychiatric Association, 2013. DSM-V, doi-org.db29.linccweb.org/10.1176/appi.books.99780890425596.dsm02

De la Vega, R., Tomé-Pires, C., Solé, E., Racine, M., Castarlenas, E., Jensen, M. P., & Miró, J. (2015). The Pittsburgh Sleep Quality Index: Validity and factor structure in young people. *Psychological Assessment*, *27*(4), e22–e27. https://doi.org/10.1037/pas0000128

De Souza, J. C., De Sousa, I. C., Maia, A. P. L., & De Azevedo, C. V. M. (2012). Sleep patterns of teachers and adolescents who attend school in the morning. *Biological Rhythm Research*, *43*(1), 65–72. https://doi.org/10.1080/09291016.2011.638156

De Zambotti, M., Goldstone, A., Colrain, I. M., & Baker, F. C. (2018). Insomnia disorder in adolescence: Diagnosis, impact, and treatment. *Sleep Medicine Reviews*, *39*, 12–24. https://doi.org/10.1016/j.smrv.2017.06.009

Deng, J., Zhou, F., Hou, W., Silver, Z., Wong, C. Y., Chang, O., Huang, E., & Zuo, Q. K. (2021). The prevalence of depression, anxiety, and sleep disturbances in COVID-19 patients: a meta-analysis. *Annals of the New York Academy of Sciences*, *1486*(1), 90–111. https://doi.org/10.1111/nyas.14506

Department of Health, & Department for Education. (2017). *Transforming Children and Young People's Mental Health Provision: A Green Paper*.

Dewald, J. F., Meijer, A. M., Oort, F. J., Kerkhof, G. A., & Bögels, S. M. (2010). The influence of sleep quality, sleep duration and sleepiness on school performance in children and adolescents: A meta-analytic review. *Sleep Medicine Reviews*. https://doi.org/10.1016/j.smrv.2009.10.004

Díaz-Morales, J. F., & Escribano, C. (2014). Consequences of adolescent's evening preference on psychological functioning: a review. *Anales de Psicología*, *30*(3), 1096–1104. https://doi.org/10.6018/analesps.30.3.167941

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

Dietch, J. R., Sethi, K., Slavish, D. C., & Taylor, D. J. (2019). Validity of two retrospective questionnaire versions of the Consensus Sleep Diary: the whole week and split week Self-Assessment of Sleep Surveys. *Sleep Medicine*, *63*, 127–136. https://doi.org/10.1016/j.sleep.2019.05.015

Downs, J., Ford, T., Stewart, R., Epstein, S., Shetty, H., Little, R., Jewell, A., Broadbent, M., Deighton, J., Mostafa, T., Gilbert, R., Hotopf, M., & Hayes, R. (2019). An approach to linking education, social care and electronic health records for children and young people in South London: A linkage study of child and adolescent mental health service data. *BMJ Open*, *9*(1), 1–12. https://doi.org/10.1136/bmjopen-2018-024355

Downs, J., Hotopf, M., Ford, T., Simonoff, E., Jackson, R. G., Shetty, H., Stewart, R., & Hayes, R. D. (2016). Clinical predictors of antipsychotic use in children and adolescents with autism spectrum disorders: a historical open cohort study using electronic health records. *European Child and Adolescent Psychiatry*, *25*(6), 649–658. https://doi.org/10.1007/s00787-015-0780-7

Du Prel, J. B., Hommel, G., Röhrig, B., & Blettner, M. (2009). Confidence interval or P-value? Part 4 of a series on evaluation of scientific publications. *Deutsches Arzteblatt*, *106*(19), 335–339. https://doi.org/10.3238/arztebl.2009.0335

Dumay, N., & Gaskell, M. G. (2007). Sleep-associated changes in the mental representation of spoken words. *Psychological Science*, *18*(1), 35–39. https://doi.org/10.1111/j.1467-9280.2007.01845.x

Earis, F. J., Brooks-Gunn, J., Raudenbush, S. W., & Sampson, R. J. (2000). Project on human development in Chicago neighborhoods (PHDCN): Physical Development Scale, wave 3, 2000-2002. *ICPSR*, *13730*, 120–125.

Ebesutani, C., Reise, S. P., Chorpita, B. F., Ale, C., Regan, J., Young, J., Charmaine, H. M. M., & Weisz, J. R. (2012). The Revised Child Anxiety and Depression Scale-Short Version: Scale reduction via exploratory bifactor modeling of the broad anxiety factor. *Psychological Assessment*, *24*(4), 833–845. https://doi.org/10.1037/a0027283

Edens, K. M. (2006). The relationship of university students' sleep habits and academic motivation. *NASPA Journal*, *43*(3), 432–445.

Efron, D., Lycett, K., & Sciberras, E. (2014). Use of sleep medication in children with ADHD. *Sleep Medicine*, *15*(4), 472–475. https://doi.org/10.1016/j.sleep.2013.10.018

Ellenbogen, J. M., Hulbert, J. C., Jiang, Y., & Stickgold, R. (2009). The sleeping brain's influence on verbal memory: Boosting resistance to interference. *PLoS ONE*, *4*(1), 2–5. https://doi.org/10.1371/journal.pone.0004117

Engen, H. G., & Anderson, M. C. (2018). Memory control: A fundamental mechanism of emotion regulation. *Trends in Cognitive Sciences*, *22*(11), 982–995. https://doi.org/10.1016/j.tics.2018.07.015

Ennis, C. R., Short, N. A., Moltisanti, A. J., Smith, C. E., Joiner, T. E., & Taylor, J. (2017). Nightmares and nonsuicidal self-injury : The mediating role of emotional

dysregulation. *Comprehensive Psychiatry*, *76*, 104–112. https://doi.org/10.1016/j.comppsych.2017.04.003

Enright, T., & Refinetti, R. (2017). Chronotype, class times, and academic achievement of university students. *Chronobiology International*, *34*(4), 445–450. https://doi.org/10.1080/07420528.2017.1281287

Epstein, S., Roberts, E., Sedgwick, R., Polling, C., Finning, K., Ford, T., Dutta, R., & Downs, J. (2019). School absenteeism as a risk factor for self-harm and suicidal ideation in children and adolescents: a systematic review and meta-analysis. *European Child and Adolescent Psychiatry*. https://doi.org/10.1007/s00787-019-01327-3

Erdfelder, E., Faul, F., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Estevan, I., Silva, A., & Tassino, B. (2018). School start times matter, eveningness does not. *Chronobiology International*, *35*(12), 1753–1757. https://doi.org/10.1080/07420528.2018.1504785

Facer-Childs, E., Middleton, B., Skene, D. J., & Bagshaw, A. P. (2019). Resetting the late timing of 'night owls' has a positive impact on mental health and performance. *Sleep Medicine, 60,* 236 – 247. https://doi.org/10.1016/j.sleep.2019.05.001

Fante-Coleman, T., & Jackson-Best, F. (2020). Barriers and facilitators to accessing mental healthcare in canada for black youth: A scoping review. *Adolescent Research Review*, *5*(2), 115–136. https://doi.org/10.1007/s40894-020-00133-2

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Feliu-Soler, A., Pérez-Aranda, A., Montero-Marín, J., Herrera-Mercadal, P., Andrés-Rodríguez, L., Angarita-Osorio, N., Williams, A. D., & Luciano, J. V. (2019). Fifteen years controlling unwanted thoughts: A systematic review of the thought control ability questionnaire (TCAQ). *Frontiers in Psychology*, *10*(JUN). https://doi.org/10.3389/fpsyg.2019.01446

Fenn, K. M., & Hambrick, D. Z. (2012). Individual differences in working memory capacity predict sleep-dependent memory consolidation. *Journal of Experimental Psychology: General*, *41*(3), 404–410. https://doi.org/10.1037/a0025268.supp

Fenn, K. M., & Hambrick, D. Z. (2015). General intelligence predicts memory change across sleep. *Psychonomic Bulletin and Review*, *22*, 791–799. https://doi.org/10.3758/s13423-014-0731-1

Ferrarelli, F., Huber, R., Peterson, M. J., Massimini, M., Murphy, M., Reidner, B. A., Watson, A., Bria, P., Tononi, G. (2007). Reduted sleep spindle activity in schizophrenia patients. *American Journal of Psychiatry,* 164, 483-492. http://doi.org/10.1176/appi.ajp.164.3.483

Frances, C., de Bruin, A., & Duñabeitia, J. A. (2020). The effects of language and emotionality of stimuli on vocabulary learning. *PLoS ONE*, *15*(10), e0240252. https://doi.org/10.1371/journal.pone.0240252

Freeman, D., Sheaves, B., Waite, F., Harvey, A. G., & Harrison, P. J. (2020). Sleep disturbance and psychiatric disorders. *The Lancet Psychiatry*, *7*(7), 628–637. https://doi.org/10.1016/S2215-0366(20)30136-X

Fusar-Poli, P. (2019). Integrated mental health services for the developmental period (0 to 25 years): A critical review of the evidence. *Frontiers in Psychiatry*, *10*(JUN), 1–17. https://doi.org/10.3389/fpsyt.2019.00355

García-Escalera, J., Chorot, P., Valiente, R. M., Reales, J. M., & Sandín, B. (2016). Efficacy of transdiagnostic cognitive-behavioral therapy for anxiety and depression in adults, children and adolescents: A meta-analysis. *Revista de Psicopatologia y Psicologia Clinica*, *21*(3), 147–175. https://doi.org/10.5944/rppc.vol.21.num.3.2016.17811

Garnefski, N., & Kraaij, V. (2006). Cognitive emotion regulation questionnaire - development of a short 18-item version (CERQ-short). *Personality and Individual Differences*, *41*(6), 1045–1053. https://doi.org/10.1016/j.paid.2006.04.010

Gaskell, M. G., & Dumay, N. (2003). Lexical competition and the acquisition of novel words. *Cognition*, *89*, 105–132. https://doi.org/10.1016/S0010-0277(03)00070-2

Genzel, L., & Wixted, J. T. Cellular and systems consolidation of declarative memory. *Cognitive Neuroscience of Memory Consolidation*, 3-16, DOI: 10.1007/978-3-319-45066-7_1

Gobin, C. M., Banks, J. B., Fins, A. I., & Tartar, J. L. (2015). Poor sleep quality is associated with a negative cognitive bias and decreased sustained attention. *Journal of Sleep Research*, *24*, 535–542. https://doi.org/10.1111/jsr.12302

Goldstein, A. N., & Walker, M. P. (2014). The role of sleep in emotional brain function. *Annual Review of Clinical Psychology*, *10*(1), 679–708. https://doi.org/10.1146/annurev-clinpsy-032813-153716

Gomez, F. (2016). A guide to the Depression, Anxiety and Stress Scale (DASS 21)-Consultant Clinical Psychologist. *Black Dog Institute*, 1–4. https://www.cesphn.org.au/images/mental_health/Frequently_Used/Outcome_Tools/Dass21.pdf

Gomez Fonseca, A., & Genzel, L. (2020). Sleep and academic performance: considering amount, quality and timing. *Current Opinion in Behavioral Sciences*, *33*, 65–71. https://doi.org/10.1016/j.cobeha.2019.12.008

Goodman, R. (2001). Psychometric properties of the strengths and difficulties questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry,* 40 (11), 1337-1345.

Gradisar, M., Gardner, G., & Dohnt, H. (2011). Recent worldwide sleep patterns and problems during adolescence: A review and meta-analysis of age, region, and sleep. *Sleep Medicine, 12*, 110–118. https://doi.org/10.1016/j.sleep.2010.11.008

Gradisar, M., Kahn, M., Micic, G., Short, M., Reynolds, C., Orchard, F., Bauducco, S., Bartel, K., & Richardson, C. (2022). Sleep's role in the development and resolution of adolescent depression. *Nature Reviews in Psychology,* 1, 512 – 523, https://doi.org/10.1038/s44159-022-00074-8

Gunn, H. E., O'Rourke, F., Dahl, R. E., Goldstein, T. R., Rofey, D. L., Forbes, E. E., & Shaw, D. S. (2019). Young adolescent sleep is associated with parental monitoring. *Sleep Health, 5*(1), 58–63. https://doi.org/10.1016/j.sleh.2018.09.001

Hagenauer, M. H., Perryman, J. I., Lee, T. M., & Carskadon, M. A. (2009). Adolescent changes in the homeostatic and circadian regulation of sleep. *Developmental Neuroscience*, *31*(4), 276–284. https://doi.org/10.1159/000216538

Hahn, M. A., Heib, D., Schabus, M., Hoedlmoser, K., & Helfrich, R. F. (2020). Slow oscillation-spindle coupling predicts enhanced memory formation from childhood to adolescence. *ELife*, 1–21. https://doi.org/10.7554/eLife.53730

Hahn, M. A., Joechner, A. K., Roell, J., Schabus, M., Heib, D. P. J., Gruber, G., Peigneux, P., & Hoedlmoser, K. (2018). Developmental changes of sleep spindles and their impact on sleep-dependent memory consolidation and general cognitive abilities: A longitudinal approach. *Developmental Science*, *22*(1), 1–16. https://doi.org/10.1111/desc.12706

Hair, J. F., Howard, M. C., & Nitzl, C. (2020). Assessing measurement model quality in PLS-SEM using confirmatory composite analysis. *Journal of Business Research*, *109*(August 2019), 101–110. https://doi.org/10.1016/j.jbusres.2019.11.069

Hair, J. F., Sarstedt, M., Hopkins, L., & Kuppelwieser, V. G. (2014). Partial least squares structural equation modeling (PLS-SEM): An emerging tool in business research. *European Business Review*, *26*(2), 106–121. https://doi.org/10.1108/EBR-10-2013-0128

Hair, J. F., Sarstedt, M., Pieper, T. M., & Ringle, C. M. (2012). The use of partial least squares structural equation modeling in strategic management research: A review of past practices and recommendations for future applications. *Long Range Planning*, *45*(5–6), 320–340. https://doi.org/10.1016/j.lrp.2012.09.008

Hair, J. F., Sarstedt, M., Ringle, C. M., & Mena, J. A. (2012). An assessment of the use of partial least squares structural equation modeling in marketing research. *Journal of the Academy of Marketing Science*, *40*(3), 414–433. https://doi.org/10.1007/s11747-011-0261-6

Hairston, I. S., Talbot, L. S., Eidelman, P., Gruber, J., & Harvey, A. G. (2010). Sensory gating in primary insomnia. *European Journal of Neuroscience*, *31*(11), 2112–2121. https://doi.org/10.1111/j.1460-9568.2010.07237.x

Harrington, M. O., & Cairney, S. A. (2021). Sleep loss gives rise to intrusive thoughts. *Trends in Cognitive Sciences*, 1–3. https://doi.org/10.1016/j.tics.2021.03.001

Harrington, M. O., Pennington, K., & Durrant, S. J. (2017). The 'Affect Tagging and Consolidation' (ATaC) model of depression vulnerability. *Neurobiology of Learning and Memory*, *140*, 43–51. https://doi.org/10.1016/j.nlm.2017.02.003

Hawton, K., Witt, K. G., Taylor Salisbury, T. L., Arensman, E., Gunnell, D., Townsend, E., van Heeringen, K., & Hazell, P. (2015). Interventions for self-harm in children and

adolescents. *Cochrane Database of Systematic Reviews*, *2015*(12). https://doi.org/10.1002/14651858.CD012013

Heaton, J. A., Ogle, R. L., & Noel, N. E. (2012). *Modifying the caffeine consumption questionnaire: impulsivity and expectancies as predictors of caffeine consumption*.

Henderson, L. M., St Clair, M., Knowland, V., van Rijn, E., Walker, S., & Gaskell, M. G. (2023). Stronger associations between sleep and mental health in adults with autism: A biobank study. *Journal of Autism and Developmental Disordesr, 53*(4), 1543-1559. https://doi.org/10.1007/s10803-021-05382-1

Henderson, L. M., Weighall, A. R., Brown, H., & Gaskell, M. G. (2012). Consolidation of vocabulary is associated with sleep in children. *Developmental Science, 15*(5), 674–687. https://doi.org/10.1111/j.1467-7687.2012.01172.x

Henry, J. D., & Crawford, J. R. (2005). The short-form version of the Depression Anxiety Stress scales (DASS-21): Construct validity and normative data in a large non-clinical sample. *British Journal of Clinical Psychology, 44*(2), 227–239. https://doi.org/10.1348/014466505X29657

Henseler, J., Ringle, C. M., & Sarstedt, M. (2014). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science, 43*(1), 115–135. https://doi.org/10.1007/s11747-014-0403-8

Hill, A., Wilde, S., & Tickle, A. (2019). Review: Transition from Child and Adolescent Mental Health Services (CAMHS) to Adult Mental Health Services (AMHS): a meta-synthesis of parental and professional perspectives. *Child and Adolescent Mental Health, 24*(4), 295–306. https://doi.org/10.1111/camh.12339

Hirshkowitz, M., Whiton, K., Albert, S. M., Alessi, C., Bruni, O., DonCarlos, L., Hazen, N., Herman, J., Adams Hillard, P. J., Katz, E. S., Kheirandish-Gozal, L., Neubauer, D. N., O'Donnell, A. E., Ohayon, M., Peever, J., Rawding, R., Sachdeva, R. C., Setters, B.,

Vitiello, M. V., & Ware, J. C. (2015). National Sleep Foundation's updated sleep duration recommendations: Final report. *Sleep Health*, *1*(4), 233–243. https://doi.org/10.1016/j.sleh.2015.10.004

Hoddes, E., Zarcone, V., & Dement, W. (1972). Development and use of the Stanford sleepiness scale (SSS). *Psychophysiology*, *9*(1), 150-.

Hoedlmoser, K. (2020). Co-evolution of sleep spindles, learning and memory in children. *Current Opinion in Behavioral Sciences*, *33*, 138–143. https://doi.org/10.1016/j.cobeha.2020.03.005

Hoedlmoser, K., Peigneux, P., & Rauchs, G. (2022). Recent advances in memory consolidation and information processing during sleep. *Journal of Sleep Research, 31*(4), 1 – 12. https://doi.org/10.1111/jsr.13607

Hsu, C. Y., Gau, S. S. F., Shang, C. Y., Chiu, Y. N., & Lee, M. B. (2012). Associations between chronotypes, psychopathology, and personality among incoming college students. *Chronobiology International*, *29*(4), 491–501. https://doi.org/10.3109/07420528.2012.668995

Hu, X., Cheng, L., Chiu, M., & Paller, K. A. (2021). Promoting memory consolidation during sleep: A meta-analysis of targeted memory reactivation. *Psychological Bulleting, 146*(3), 218-244. https://doi.org/10.1037/bul0000223

Hummer, D. L., & Lee, T. M. (2016). Daily timing of the adolescent sleep phase: Insights from a cross-species comparison. *Neuroscience and Biobehavioral Reviews*, *70*, 171–181. https://doi.org/10.1016/j.neubiorev.2016.07.023

Hysing, M., Harvey, A. G., Linton, S. J., Askeland, K. G., & Sivertsen, B. (2016). Sleep and academic performance in later adolescence: Results from a large population-based study. *Journal of Sleep Research*, *25*(3), 318–324. https://doi.org/10.1111/jsr.12373

Ilie, S., Sutherland, A., & Vignoles, A. (2017). Revisiting free school meal eligibility as a proxy for pupil socio-economic deprivation. *British Educational Research Journal*. https://doi.org/10.1002/berj.3260

Illingworth, G., Sharman, R., Jowett, A., Harvey, C., Foster, R. G., & Espie, C. A. (2018). Challenges in implementing and assessing outcomes of school start time change in the UK : experience of the Oxford Teensleep study. *Sleep Medicine*, 1–7. https://doi.org/10.1016/j.sleep.2018.10.021

Irish, L. A., Kline, C. E., Gunn, H. E., Buysee, D. J., & Hall, M. H. (2015). The role of sleep hygiene in promoting public health: A review of empirical evidence. *Sleep Medicine Reviews*, 22, 23-36, https://doi.org/10.1016/j.smrv.2014.10.001

James, E., Gaskell, M. G., Pearce, R., Korell, C., Dean, C., & Henderson, L. M. (2021). The role of prior lexical knowledge in children's and adults' incidental word learning from illustrated stories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *47*(11), 1856–1869. https://doi.org/10.1037/xlm0001080

James, E., Gaskell, M. G., Weighall, A., & Henderson, L. (2017). Consolidation of vocabulary during sleep: The rich get richer? *Neuroscience and Biobehavioral Reviews*, *77*, 1–13. https://doi.org/10.1016/j.neubiorev.2017.01.054

JASP Team (2022). JASP (Version 0.16.4) [Computer software].

Jenni, O. G., & Carskadon, M. A. (2004). Spectral analysis of the sleep electroencephalogram during adolescence. *Sleep*, *27*(4), 774–783. https://doi.org/10.1093/sleep/27.4.774

Jeremy, G. (2019, February 6). We can talk – how co-produced mental health training for hospital staff improves outcomes for young people and professionals. *NHS Long Term Plan Blog.* www.longtermplan.nhs.uk/blog/we-can-talk/

Joint Formulary Committee (2021). *Promethazine Hydrochloride.* National Institute for Health and Care Excellence British National Formulary.

https://bnf.nice.org.uk/drug/promethazine-hydrochloride.html

Kang, S. G., Lee, Y. J., Kim, S. J., Lim, W., Lee, H. J., Park, Y. M., Cho, I. H., Cho, S. J., & Hong, J. P. (2014). Weekend catch-up sleep is independently associated with suicide attempts and self-injury in Korean adolescents. *Comprehensive Psychiatry*, *55*, 319–325. https://doi.org/10.1016/j.comppsych.2013.08.023

Keating, P. (2005). *Dprime analysis.* UCLA Phonetics Lab. http://phonetics.linguistics.ucla.edu/facilities/statistics/dprime.htm

Kelley, P., Lockley, S. W., Foster, R. G., & Kelley, J. (2015). Synchronizing education to adolescent biology: 'let teens sleep, start school later.' *Learning, Media and Technology*, *40*(2), 210–226. https://doi.org/10.1080/17439884.2014.942666

Kelley, P., Lockley, S. W., Kelley, J., & Evans, M. D. R. (2017). Is 8:30 a.m. still too early to start school? A 10:00 a.m. school start time improves health and performance of students aged 13-16. *Frontiers in Human Neuroscience*, *11*(December). https://doi.org/10.3389/fnhum.2017.00588

Kim, S. Y., & Payne, J. D. (2020). Neural correlates of sleep, stress, and selective memory consolidation. *Current Opinion in Behavioral Sciences*, *33*, 57–64. https://doi.org/10.1016/j.cobeha.2019.12.009

Kimland, E. E., Bardage, C., Collin, J., Järleborg, A., Ljung, R., & Iliadou, A. N. (2020). Pediatric use of prescribed melatonin in Sweden 2006–2017: a register based study. *European Child and Adolescent Psychiatry*, *0123456789*. https://doi.org/10.1007/s00787-020-01598-1

Kiyonaga, A., & Scimeca, J. M. (2019). Practical considerations for navigating registered reports. *Trends in Neurosciences, 42,* 568-571.

Klein, R. A., Vianello, M., Hasselman, F., Adams, B. G., Adams, Jr, R. B., Alper, S., Aveyard, M, et al. (2018). Many labs 2: Investigating variation in replicability across

samples and settings. *Advances in Methods and Practices in Psychological Science, 1*(4), 443-490.

Klinzing, J., Niethard, N., & Born, J. Mechanisms of systems memory consolidation during sleep. *Nature Neuroscience,* 22, 1598-1610. https://doi.org/10.1038/s41593-019-0467-3

Knowland, V. C. P., van Rijn, E., Gaskell, M. G., & Henderson, L. (2022). UK children's sleep and anxiety during the COVID-19 pandemic. *BMC Psychology, 10*(76). 1-13. https://doi.org/10.1186/s40359-022-00729-4

Küpper, C. S., Benoit, R. G., Dalgleish, T., & Anderson, M. C. (2014). Direct suppression as a mechanism for controlling unpleasant memories in daily life. *Journal of Experimental Psychology: General*, *143*(4), 1443–1449. https://doi.org/10.1037/a0036518

Kurdziel, L. B. F., & Spencer, R. M. C. (2016). Consolidation of novel word learning in native English-speaking adults. *Memory, 24*, 471-481.

Kuula, L., Tamminen, J., Makkonen, T., Merikanto, I., Räikkönen, K., & Pesonen, A. (2019). Higher sleep spindle activity is associated with fewer false memories in adolescent girls. *Neurobiology of Learning and Memory*, *157*, 96–105. https://doi.org/10.1016/j.nlm.2018.12.005

Lai, M. C., Lin, H. Y. & Ameis, S. H. (2022). Towards equitable diagnoses for autism and attention-deficit/hyperactivity disorder across sexes and genders. *Current Opinion in Psychiatry,* 35(2), 90-100. https://doi.org/10.1097/YCO.0000000000000770

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). International affective picture system (IAPS): Technical manual and affective ratings. *NIMH Center for the Study of Emotion and Attention*, 39–58.

Lau, E. Y. Y., Wong. M. L., Rusak, B., Lam, Y. C., Wing, Y. K., Tseng, C., Lee, T. M. C.

(2019). The coupling of short sleep duration and high sleep need predicts riskier decision making. *Psychology & Health,* 1-18. https://doi.org/10.1080/08870446.2019.1594807

Leong, R. L. F., Lau, T. Y., Dicom, A. E., Teo, T. B., Ong, J. L., Chee, M. W. L. (2023). Infleunce of mid-afternoon nap duration and sleep parameters on memory encoding, mood, processing speed, and vigilance. *Sleep,* 1-11. https://doi.org/10.1093/sleep/zsad025

Leong, R. L. F., Yu, N., Ong, J. L., Ng, A. S. C., Jamaluddin, S. A., Cousins, J. N., Chee, N. I. Y. N., Chee, M. W. L. (2021). Memory performance following napping in habitual and non-habitual nappers. *Sleep*, *44*(7), 1-11. https://doi.org/10.1093/sleep/zsab038

Leung, S. O. (2011). A comparison of psychometric properties and normality in 4-, 5-, 6-, and 11-point likert scales. *Journal of Social Service Research*, *37*(4), 412–421. https://doi.org/10.1080/01488376.2011.580697

Lev Ari, L., & Shulman, S. (2012). Pathways of sleep, affect, and stress constellations during the first year of college: Transition difficulties of emerging adults. *Journal of Youth Studies*, *15*(3), 273–292. https://doi.org/10.1080/13676261.2011.635196

Lifestyles Team, NHS Digital. (2021). Mental health of children and young people in England 2021 – wave 2 follow up to the 2017 survey. *Government Statistical Service: NHS Digital,* 1-48.

Lin, H. (n.d.). *Effect size converter*. Retrieved October 27, 2020, from https://www.escal.site/#footnote-2

Lindsay, S., & Gaskell, G. M. (2013). Lexical integration of novel words without sleep. *Journal of Experimental Psychology: Learning Memory and Cognition*, *39*(2), 608–622. https://doi.org/10.1037/a0029243

Lipinska, G., Stuart, B., Thomas, K. G. F., Baldwin, D. S., & Bolinger, E. (2019). Preferential consolidation of emotional memory during sleep: A meta-analysis. *Frontiers in Psychology*, *10*(MAY). https://doi.org/10.3389/fpsyg.2019.01014

Littlewood, D. L., Kyle, S. D., Carter, L., Peters, S., Pratt, D., & Gooding, P. (2019). Short sleep duration and poor sleep quality predict next-day suicidal ideation : an ecological momentary assessment study. *Psychological Medicine*, *49*, 403–411.

Liu, Z., Zhao, Z., Jia, C., & Buysse D. J. (2008). Sleep patterns and problems among Chinese adolescents. *Pediatrics*, *121,* 1165-1173. https://doi.org/10.1542/peds.2007-1464

Lo, J. C., & Chee, M. W. (2020). Cognitive effects of multi-night adolescent sleep restriction: current data and future possibilities. *Current Opinion in Behavioral Sciences*, *33*, 34–41. https://doi.org/10.1016/j.cobeha.2019.12.005

Lo, J. C., Ong, J. L., Leong, R. L. F., Gooley, J. J., & Chee, M. W. L. (2016). Cognitive performance, sleepiness, and mood in partially sleep deprived adolescents: The need for sleep study. *Sleep*, *39*(3), 687–698. https://doi.org/10.5665/sleep.5552

Logan, R. W., Hasler, B. P., Forbes, E. E., Franzen, P. L., Torregrossa, M. M., Huang, Y. H., Buysse, D. J., Clark, D. B., & McClung, C. A. (2018). Impact of sleep and circadian rhythms on addiction vulnerability in adolescents. *Biological Psychiatry*, *83*(12), 987-996. https://doi.org.10.1016/j.biopsych.2017.11.035

Lowe, C. J., Safati, A., & Hall, P. A. The neurocgnitive consequences of sleep restriction: A meta-analytic review. *Neuroscience and Biobehavioral Reviews*, *80*, 586-604, https://dx.doi.org/10.1016/j.neubiorev.2017.07.010

Manber, R., Edinger, J. D., Gress, J. L., San Pedro-Salcedo, M. G., Kuo, T. F., & Kalista, T. (2008). Cognitive behavioral therapy for insomnia enhances depression outcome in patients with comorbid major depressive disorder and insomnia. *Sleep*. https://doi.org/10.1093/sleep/31.4.489

Mander, B. A., Santhanam, S., Saletin, J. M., & Walker, M. P. (2011). Wake deterioration and sleep restoration of human learning. *Current Biology*, *21*(5), R183–R184. https://doi.org/10.1016/j.cub.2011.01.019

Mandera, P., Keuleers, E., & Brysbaert, M. (2017). Explaining human performance in psycholinguistic tasks with models of semantic similarity based on prediction and counting: A review and empirical validation. *Journal of Memory and Language, 92*, 57–78. https://doi.org/10.1016/j.jml.2016.04.001

March, J. A., Ricketts, J., Tamminen, J. (2023). Is word learning capacity restored after a daytime nap? *Cortex, 159,* 142-166. https://doi.org/10.1016/j.cortex.2022.10.013

Marco, C. A., Wolfson, A. R., Sparling, M., & Azuaje, A. (2011). Family socioeconomic status and sleep patterns of young adolescents. *Behavioral Sleep Medicin, 10*(1), 70–80. https://doi.org/10.1080/15402002.2012.636298

Marx, R., Tanner-Smith, E. E., Davison, C. M., Ufholz, L. A., Freeman, J., Shankar, R., Newton, L., Brown, R. S., Parpia, A. S., Cozma, I., & Hendrikx, S. (2017). Later school start times for supporting the education, health, and well-being of high school students. In *Cochrane Database of Systematic Reviews*. https://doi.org/10.1002/14651858.CD009467.pub2

Mattys, S. L., & Clark, J. H. (2002). Lexical activity in speech processing: Evidence from pause detection. *Journal of Memory and Language, 47*(3), 343-359. https://doi.org/10.1016/S0749-596X(02)00037-2

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, *102*(3), 419–457. https://doi.org/10.1037/0033-295X.102.3.419

McDevitt, E. A., Alaynick, W. A., & Mednick, S. C. (2012). The effect of nap frequency on daytime sleep architecture. *Physiology & Behaviour*, *107*(1), 40-44. https://doi.org/10.1016/j.physbeh.2012.05.021

McGlinchey, E. L., Courtney-Seidler, E. A., German, M., & Miller, A. L. (2017). The role of sleep disturbance in suicidal and nonsuicidal self-injurious behavior among adolescents. *Suicide and Life-Threatening Behavior*, *47*(1), 103–111. https://doi.org/10.1111/sltb.12268

McLay, L. K., Schluter, P. J., Eggleston, M. J. F., Woodford, E. C., & Bowden, N. (2021). Melatonin dispensing among New Zealand children aged 0–18 years with autism: a nationwide cross-sectional study. *Sleep Medicine, 80*, 184–192. https://doi.org/10.1016/j.sleep.2021.01.028

Merikanto, I., Lahti, J., Kuula, L., Heinonen, K., Räikkönen, K., Andersson, S., Strandberg, T., & Pesonen, A. K. (2018). Circadian preference and sleep timing from childhood to adolescence in relation to genetic variants from a genome-wide association study. *Sleep Medicine, 50*, 36–41. https://doi.org/10.1016/j.sleep.2018.04.015

Meteyard, L., & Davies, R. A. I. (2020). Best practice guidance for linear mixed-effecs models in psychological science. *Journal of Memory and Language, 112*(104092), 1–22. https://doi.org/https://doi.org/10.1016/j.jml.2020.104092

Minges, K. E., & Redeker, N. S. (2016). Delayed school start times and adolescent sleep: A systematic review of the experimental evidence. *Sleep Medicine Reviews, 28*, 86–95. https://doi.org/10.1016/j.smrv.2015.06.002

Ministry of Housing, Communities & Local Government (2019). English indices of deprivation 2019. Retrieved February 1, 2022, from http://imd-by-postcode.opendatacommunities.org/imd/2019

Minkel, J. D., Banks, S., Htaik, O., Moreta, M. C., Jones, C. W., McGlinchey, E. L., Simpson, N. S., & Dinges, D. F. (2012). Sleep deprivation and stressors: Evidence for elevated negative affect in response to mild stressors when sleep deprived. *Emotion*, *12*(5), 1015–1020. https://doi.org/10.1037/a0026871

Mölle, M., Bergmann, T., 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

Mondiana, Y. Q., Pramoedyo, H., & Sumarminingsih, E. (2018). Structural equation modeling on likert scale data with transformation by successive interval method and with no transformation. *International Journal of Scientific and Research Publications (IJSRP)*, *8*(5), 398–405. https://doi.org/10.29322/ijsrp.8.5.2018.p7751

Monecke, A. (2015). *Package "semPLS"* (pp. 1–26). *CRAN Repository*

Monecke, A., & Leisch, F. (2012). SemPLS: Structural equation modeling using partial least squares. *Journal of Statistical Software*, *48*(3). https://doi.org/10.18637/jss.v048.i03

Muller, D., Judd, C. M., & Yzerbyt, V. Y. (2005). When moderation is mediated and mediation is moderated. *Journal of Personality and Social Psychology*. https://doi.org/10.1037/0022-3514.89.6.852

Natale, V., & Cicogna, P. C. (2002). Morningness-eveningness dimension: Is it really a continuum? *Personality and Individual Differences*, *32*(5), 809–816. https://doi.org/10.1016/S0191-8869(01)00085-X

Nelson, D. L., Evoy, C. L. M. C., & Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computer.University of Kansas, Lawrence, Kansas Copyright 2004 Psychonomic Society, Inc*, *36*(3), 402–407. https://doi.org/10.3758/bf03195588

NHS Digital (2021). Health and care of people with learning disabilities experimental

statistics 2019 to 2020. Retrieved from: https://digital.nhs.uk/data-and-information/publications/statistical/health-and-care-of-people-with-learning-disabilities/experimental-statistics-2019-to-2020/adhd-autism-and-epilepsy Accessed 10th February 2023.

NHS England & The Royal College of Paediatrics and Child Health (2018). Supporting treatment and appropriate medication in paediatrics (STAMP). Retrieved from: https://www.england.nhs.uk/learning-disabilities/improving-health/stamp/ Accessed 10th February 2023.

NICE. (2013). Using melatonin to treat sleep disorders in children and young people with ADHD. *Information for the Public*, 1–6. Retrieved from: https://www.nice.org.uk/advice/esuom2/resources/using-melatonin-to-treat-sleep-disorders-in-children-and-young-people-with-adhd-pdf-17486436805

Nitzl, C., Roldan, J. L., & Cepeda, G. (2016). Mediation analysis in partial least squares path modelling, Helping researchers discuss more sophisticated models. *Industrial Management and Data Systems*, *116*(9), 1849–1864. https://doi.org/10.1108/IMDS-07-2015-0302

Nunes, M. L., & Bruni, O. (2015). Insônia na infância e adolescência: aspectos clínicos, diagnóstico e abordagem terapêutica. *Jornal de Pediatria*, *91*(6), S26–S35. https://doi.org/10.1016/j.jped.2015.08.006

Office for National Statistics (2022, November 29). Ethnic group, England and Wales: Census 2021. *Census 2021*. Retrieved from: www.ons.gov.uk/peoplepopulationandcommunity/culturalidentity/ethnicity/bulletins/ethnicgroupenglandandwales/census2021

Office for National Statistics (2021, May 17). What are the regional differences in income and productivity? Explore ecnomic inequality in the UK with our map. Retrieved from: https://www.ons.gov.uk/visualisations/dvc1370/

O'Leary, K., Bylsma, L. M., & Rottenberg, J. (2017). Why might poor sleep quality lead to depression? A role for emotion regulation. *Cognition and Emotion*, *31*(8), 1698–1706. https://doi.org/10.1080/02699931.2016.1247035

Oerbeck, B., Overgaard, K. R., Hjellvik, V., Bramness, J. G., Hansen, B. H., & Lien, L. (2020). The use of sleep medication in youth residential care. *Journal of Child and Adolescent Psychopharmacology*, *30*(5), 335–341. https://doi.org/10.1089/cap.2019.0172

Olcese, U., Esser, S. K., & Tononi, G. (2010). Sleep and synaptic renormalization: A computational study. *Journal of Neurophysiology*, *104*(6), 3476–3493. https://doi.org/10.1152/jn.00593.2010

Olya, H. (2017). Partial least squares based structural equation modeling (PLS-SEM). *Global Conference on Services Management*, 1–29.

Ong, A. D., Kim, S., Young, S., & Steptoe, A. (2017). Positive affect and sleep: A systematic review. *Sleep Medicine Reviews*, *35*, 21–32. https://doi.org/10.1016/j.smrv.2016.07.006

Ong, J. L., Lau, T. Y., Lee, X. K., van Rijn, E., & Chee, M. W. L. (2020). A daytime nap restores hippocampal function and improves declarative learning. *Sleep*, 1–9. https://doi.org/10.1093/sleep/zsaa058

Ong, J. L., Lo, J. C., Gooley, J. J., & Chee, M. W. L. (2016). EEG changes across multiple nights of sleep restriction and recovery in adolescents: The need for sleep study. *Sleep*, *39*(6), 1233–1240. https://doi.org/10.5665/sleep.5840

Ong, J. L., Lo, J. C., Gooley, J. J., & Chee, M. W. L. (2017). EEG changes accompanying successive cycles of sleep restriction with and without naps in adolescents. *Sleep, 40*(4). https://doi.org/10.1093/sleep/zsx030

Ong, J. L., Patanaik, A., Chee, N. I. Y. N., Lee, X. K., Poh, J. H., Chee, M. W. L. (2018). Auditory stimulation of sleep slow oscillations modulates subsequent memory encoding through altered hippocampal function. *Sleep, 41*(5) 1-11. https://doi.org/10/1093/sleep/zsy031

Owens, J. A., Rosen, C. L., & Mindell, J. A. (2003). Medication use in the treatment of pediatric insomnia: results of a survey of community-based pediatricians. *Pediatrics, 111*.

Paz-Alonso, P. M., Ghetti, S., Matlen, B. J., Anderson, M. C., & Bunge, S. A. (2009). Memory suppression is an active process that improves over childhood. *Frontiers in Human Neuroscience, 3*(SEP), 1–6. https://doi.org/10.3389/neuro.09.024.2009

Perlus, J. G., Brien, F. O., Haynie, D. L., & Simons-morton, B. G. (2018). Adolescent sleep insufficiency one year after high school. *Journal of Adolescence*, *68*(January), 165–170. https://doi.org/10.1016/j.adolescence.2018.07.016

Petersen, A. C., Crockett, L., Richards, M., & Boxer, A. (1987). A self-report measure of pubertal status: Reliability, validity, and initial norms. *Journal of Youth and Adolescence*, *17*(2), 117–133. https://doi.org/10.1103/PhysRevLett.78.1396

Petitions Committee. (2019). *School should start at 10am as teenagers are too tired*. UK Government and Parliament. https://petition.parliament.uk/petitions/229178

Pfost, M., Hattie, J., Dörfler, T., & Artelt, C. (2014). Individual differences in reading development: A review of 25 years of empirical research on Matthew Effects in reading. *Review of Educational Research, 84*(2), 203-244. doi:10.3102/0034654313509492

Phillips, A. J. K., Clerx, W. M., O'Brien, C. S., Sano, A., Barger, L. K., Picard, R. W., Lockley, S. W., Klerman, E. B., & Czeisler, C. A. (2017). Irregular sleep/wake patterns are associated with poorer academic performance and delayed circadian and sleep/wake timing. *Scientific Reports*, *7*(1), 1–13. https://doi.org/10.1038/s41598-017-03171-4

Pöhlchen, D., Pawlizki, A., Gais, S., & Schönauer, M. (2020). Evidence against a large effect of sleep in protecting verbal memories from interference. *Journal of Sleep Research*, *November 2019*, 1–7. https://doi.org/10.1111/jsr.13042

Pöhlchen, D., & Schönauer, M. (2020). Sleep-dependent memory consolidation in the light of rapid neocortical plasticity. *Current Opinion in Behavioral Sciences*, *33*, 118–125. https://doi.org/10.1016/j.cobeha.2020.02.001

Polos, P. G., Bhat, S., Gupta, D., O'Malley, R. J., DeBari, V. A., Upadhyay, H., Chaudhry, S., Nimma, A., Pinto-Zipp, G., & Chokroverty, S. (2015). The impact of Sleep Time-Related Information and Communication Technology (STRICT) on sleep patterns and daytime functioning in American adolescents. *Journal of Adolescence*. https://doi.org/10.1016/j.adolescence.2015.08.002

Quante, M., Kaplan, E. R., Cailler, M., Rueschman, M., Wang, R., Weng, J., Taveras, E. M., & Redline, S. (2018). Actigraphy-based sleep estimation in adolescents and adults: a comparison with polysomnography using two scoring algorithms. *Nature and Science of Sleep*, *10*, 13–20. https://doi.org/10.2147/NSS.S151085

Quartel, L. (2014). The effect of the circadian rhythm of body temperature on A-level exam performance. *Undergraduate Journal of Psychology*, *27*(1).

Ramtekkar, U., & Ivanenko, A. (2015). Sleep in children with psychiatric disorders. *Seminars in Pediatric Neurology*, *22*(2), 148–155. https://doi.org/10.1016/j.spen.2015.04.004

Randler, C., Vollmer, C., Kalb, N., & Itzek-Greulich, H. (2019). Breakpoints of time in bed, midpoint of sleep, and social jetlag from infancy to early adulthood. *Sleep Medicine*, *57*, 80–86. https://doi.org/10.1016/j.sleep.2019.01.023

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

Raven, F., Van der Zee, E. A., Meerlo, P., & Havekes, R. (2018). The role of sleep in regulating structural plasticity and synaptic strength: Implications for memory and cognitive function. *Sleep Medicine Reviews*, *39*, 3–11. https://doi.org/10.1016/j.smrv.2017.05.002

Revilla, M., & Hohne, J. K. (2020). How long do respondents think online surveys should be? New evidence from two online panels in Germany. *International Journal of Market Research*, *62*(5), https://doi.org/10.1177/1470785320943049

Reynolds, C. M., Short, M. A., & Gradisar, M. (2018). Sleep spindles and cognitive performance across adolescence: A meta-analytic review. *Journal of Adolescence*, *66*(January), 55–70. https://doi.org/10.1016/j.adolescence.2018.04.003

Reynolds, S., Orchard, F., Midgley, N., Kelvin, R., & Goodyer, I. (2020). Do sleep disturbances in depressed adolescents improve following psychological treatment for depression? *Journal of Affective Disorders*, *262*(April 2019), 205–210. https://doi.org/10.1016/j.jad.2019.10.029

Ricketts, J., Bishop, D. V. M., Pimperton, H., & Nation, K. (2011). The role of self-teaching in learning orthographic and semantic aspects of new words. *Scientific Studies of Reading*, *15*(1), 47 - 70. doi:10.1080/10888438.2011.536129

Ricketts, J., Lervåg, A., Dawson, N., Taylor, L. A., & Hulme, C. (2020). Reading and oral vocabulary development in early adolescence. *Scientific Studies of Reading*, *24*(5), 380-396. DOI: 10.1080/10888438.2019.1689244

Ricketts, J., Dawson, N., & Davies, R. (2021). The hidden depths of new word knowledge: Using graded measures of orthographic and semantic learning to measure vocabulary acquisition. *Learning and Instruction, 74*, 101468. doi:10.1016/j.learninstruc.2021.101468

Rosanova, M., & Ulrich, D. (2005). Pattern-specific associative long-term potentiation induced by a sleep spindle-related spike train. *The Journal of Neuroscience,* 25 (41), 9398-9405. https://doi.org/10.1523/JNEUROSCI.2149-05.2005

Rusterholz, T., Hamann, C., Markovic, A., Schmidt, S. J., Achermann, P., & Tarokh, L. (2018). Nature and nurture: Brain region-specific inheritance of sleep neurophysiology in adolescence. *Journal of Neuroscience, 38*(43), 9275–9285. https://doi.org/10.1523/JNEUROSCI.0945-18.2018

Rzepka-Migut, B., & Paprocka, J. (2020). Efficacy and safety of melatonin treatment in children with autism spectrum disorder and attention-deficit/hyperactivity disorder-A review of the literature. *Brain Sciences*, *10*(4). https://doi.org/10.3390/brainsci10040219

Saletin, J. M., & Walker, M. P. (2012). Nocturnal mnemonics: sleep and hippocampal memory processing. *Frontiers in Neurology,* 3 (59), 1-12. https://doi.org/10.3389/fneur.2012.00059

Sarstedt, M. (n.d.). *What is PLS-SEM?* PLS-SEM Academy. https://www.pls-sem-academy.com/courses/588139/lectures/10566745

Sarstedt, M., Hair, J. F., Nitzl, C., Ringle, C. M., & Howard, M. C. (2020). Beyond a tandem analysis of SEM and PROCESS: Use of PLS-SEM for mediation analyses! *International Journal of Market Research*, *62*(3), 288–299. https://doi.org/10.1177/1470785320915686

Sawyer, S. M., Azzopardi, P. S., Wickremarathne, D., & Patton, G. C. (2018). Viewpoint The age of adolescence. *The Lancet Child and Adolescent Health*, *2*(18), 223–228.

Schneider B. A., Avivi-Reich, M., & Mozuraitis, M. (2015). A cautionary note on the use of the Analysis of Covariance (ANCOVA) in classification designs with and without within-subject factors. *Frontiers in Psychology, 6,* 474.

Schneider, I. K., Veenstra, L., van Harreveld, F., Schwarz, N., & Koole, S. L. (2016). Let's not be indifferent about neutrality: Neutral ratings in the International Affective Picture System(IAPS) mask mixed affective responses. *Emotion, 16*(4), 426-430. https://doi.org/10.1037/emo0000164

Schreiner, T., & Rasch, B. (2017). The beneficial role of memory reactivation for language learning during sleep: A review. *Brain and Language,* 167, 94-105. https://dx.doi.org/10.1016/j.bandl.2016.02.005

Schweizer, S., Gotlib, I. H., & Blakemore, S. J. (2020). The role of affective control in emotion regulation during adolescence. *Emotion (Washington, D.C.), 20*(1), 80–86. https://doi.org/10.1037/emo0000695

Shekleton, J. A., Rajaratnam, S. M. W., Gooley, J. J., Van Reen, E., Czeisler, C. A., & Lockley, S. W. (2013). Improved neurobehavioral performance during the wake maintenance zone. *Journal of Clinical Sleep Medicine*, *9*(4), 353–362. https://doi.org/10.5664/jcsm.2588

Siapas, A. G., & Wilson, M. A. (1998). Coordinated interactions between hippocampal riples and cortical spindles during slow-wave sleep. *Neuron, 21*(5), 1123-1128. https://doi.org/10.1016/S0896-6273(00)80629-7

Stanovich, K. E. (1986). Mathew effects in reading: Some consequences of individual differences in the acquisition of litearcy. *Reading Research Quarterly,* 21*(4),* 360-407. DOI: 10.1598/RRQ.21.4.1

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,* 1679-1686.

Ray, S., Danks, N. P., Valdez, A. C., Estradam J. M. V., Uanhoro, J., Nakayama, J., Koya, L., Burbach, L., Bejar, A. H. C., & Adler, S. (2022). Package 'SeminR'. *CRAN Repository,* version 2.3.2.

Ringle, C. M., Wende, S., & Becker, J. M. (2022). SmartPLS-4. Oststeinbek: SmartPLS. Retrrieved from: https://www.smartpls.com

Shen, L., Schie, J. Van, Ditchburn, G., Brook, L., & Bei, B. (2018). Positive and negative emotions: Differential associations with sleep duration and quality in adolescents. *Journal of Youth and Adolescence*, 2584–2595. https://doi.org/10.1007/s10964-018-0899-1

Short, M. A., Booth, S. A., Omar, O., Ostlundh, L., & Arora, T. (2020). The relationship between sleep duration and mood in adolescents: A systematic review and meta-analysis. In *Sleep Medicine Reviews* (Vol. 52). W.B. Saunders Ltd. https://doi.org/10.1016/j.smrv.2020.101311

Soetewey, A. (2020). *Outliers detection in R*. Stats and R. Retrived from: https://www.statsandr.com/blog/outliers-detection-in-r/#fn2

Soyombo, S., Stanbrook, R., Aujla, H., Capewell, D., Shantikumar, M., Kidy, F., Todkill, D., & Shantikumar, S. (2020). Socioeconomic status and benzodiazepine and Z-drug prescribing: a cross-sectional study of practice-level data in England. *Family Practice*, *37*(2), 194–199. https://doi.org/10.1093/fampra/cmz054

Speedy, K., Nukalapati, L., Speedy, K., & Davies-Kabir, M. (2021). Melatonin prescribing practices and the provision of sleep hygiene/parent-led sleep behavioural

Interventions in S-CAMHS, Aneurin Bevan University Health Board (ABUHB)-Service evaluation as part of quality improvement project. *BJPsych Open*, *7*(S1), S351–S352. https://doi.org/10.1192/bjo.2021.920

Streukens, S., & Leroi-Werelds, S. (2016). Bootstrapping and PLS-SEM: A step-by-step guide to get more out of your bootstrap results. *European Management Journal*, *34*(6), 618–632. https://doi.org/10.1016/j.emj.2016.06.003

Sun, W., Ling, J., Zhu, X., Lee, T. M. C., & Li, S. X. (2019). Associations of weekday-to-weekend sleep differences with academic performance and health-related outcomes in school-age children and youths. *Sleep Medicine Reviews*, *46*, 27–53. https://doi.org/10.1016/j.smrv.2019.04.003

Talbot, L. S., McGlinchey, E. L., Kaplan, K. A., Dahl, R. E., & Harvey, A. G. (2010). Sleep deprivation in adolescents and adults: Changes in affect. *Emotion*, *10*(6), 831–841. https://doi.org/10.1037/a0020138

Tamminen, J., & Gaskell, M. G. (2008). Newly learned spoken words show long-term lexical competition effects. *Quarterly Journal of Experimental Psychology*, *61*(3), 361–371. https://doi.org/10.1080/17470210701634545

Tamminen, J., Lambon Ralph, M. A., & Lewis, P. A. (2013). The role of sleep spindles and slow-wave activity in integrating new information in semantic memory. *Journal of Neuroscience*, *33*(39), 15376–15381. https://doi.org/10.1523/JNEUROSCI.5093-12.2013

Tamminen, J., Lambon Ralph, M. A., & Lewis, P. A. (2017). Targeted memory reactivation of newly learned words during sleep triggers REM-mediated integration of new memories and existing knowledge. *Neurobiology of Learning and Memory,* 137, 88-82. https://doi.org/10.1016/j.nlm.2016.11.012

Tamminen, J., Payne, J. D., Stickgold, R., Wamsley, E. J., & Gaskell, M. G. (2010). Sleep spindle activity is associated with the integration of new memories and existing knowledge. *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.3028-10.2010

Tashjian, S. M., Mullins, J. L., & Galván, A. (2019). Bedtime autonomy and cellphone use influence sleep duration in adolescents. *Journal of Adolescent Health*, *64*(1), 124–130. https://doi.org/10.1016/j.jadohealth.2018.07.018

Tempesta, D., Socci, V., Coppo, M., Dello Ioio, G., Nepa, V., De Gennaro, L., & Ferrara, M. (2016). The effect of sleep deprivation on the encoding of contextual and non-contextual aspects of emotional memory. *Neurobiology of Learning and Memory*, *131*, 9–17. https://doi.org/10.1016/j.nlm.2016.03.007

Tempesta, D., Socci, V., De Gennaro, L., & Ferrara, M. (2018). Sleep and emotional processing. In *Sleep Medicine Reviews*. https://doi.org/10.1016/j.smrv.2017.12.005

Tham, E. K. H., Lindsay, S., & Gaskell, M. G. (2015). Markers of automaticity in sleep-associated consolidation of novel words. *Neuropsychologia, 71*, 146-157. https://doi.org/10.1016/j.neuropsychologia.2015.03.025

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. (2014). Sleep and the Price of Plasticity: From Synaptic and Cellular Homeostasis to Memory Consolidation and Integration. *Neuron*, *81*(1), 12–34. https://doi.org/10.1016/j.neuron.2013.12.025

Tottenham, N., Hare, T. A., & Casey, B. J. (2011). Behavioral assessment of emotion discrimination, emotion regulation, and cognitive control in childhood, adolescence, and adulthood. *Frontiers in Psychology*, *2*(MAR), 1–9. https://doi.org/10.3389/fpsyg.2011.00039

Towner, E., Chierchia, G., Blakemore, S. J. (in press). Sensitivity and specificity in affective and social learning in adolescence. *Trends in Cognitive Science.*

Tsiachristas, A., McDaid, D., Casey, D., Brand, F., Leal, J., Park, A. La, Geulayov, G., & Hawton, K. (2017). General hospital costs in England of medical and psychiatric care for patients who self-harm: a retrospective analysis. *The Lancet Psychiatry*, *4*(10), 759–767. https://doi.org/10.1016/S2215-0366(17)30367-X

Van Den Bergh, D., van Doorn, J., Marsman, M., Draws, T., van Kesteren, E., Derks, K. & Wagenmakers, E. (2020). A tutorial on conducting and interpreting a Bayesian ANOVA in JASP. L'Année psychologique, 120, 73-96.

Van Der Vinne, V., Zerbini, G., Siersema, A., Pieper, A., Merrow, M., Hut, R. A., Roenneberg, T., & Kantermann, T. (2015). Timing of examinations affects school performance differently in early and late chronotypes. *Journal of Biological Rhythms*, *30*(1), 53–60. https://doi.org/10.1177/0748730414564786

Van Der Werf, Y. D., Altena, E., Schoonheim, M. M., Sanz-Arigita, E. J., Vis, J. C., De Rijke, W., & Van Someren, E. J. W. (2009). Sleep benefits subsequent hippocampal functioning. *Nature Neuroscience*, *12*(2), 122–123. https://doi.org/10.1038/nn.2253

Van Doorn, J., van den Bergh, D., Böhm, U. et al. (2021). The JASP guidelines for conducting and reporting a Bayesian analysis. *Psychonomic Bulletin & Review*, 28, 813–826.

Van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology, 67,* 1176-1190.

Van Rijn, E., Gouws, A., Walker, S. A., Knowland, V. C. P., Cairney, S. A., Gaskell, M. G., & Henderson, L. M. (2023). Do naps benefit novel word learning? Developmental

differences nad white matter correlates. *Cortex,* 37-60. https://doi.org/10.1016.j.cortex.2022.09.016

Vargas, I., Payne, J. D., Muench, A., Kuhlman, K. R., & Lopez-Duran, N. L. (2019). Acute sleep deprivation and the selective consolidation of emotional memories. *Learning and Memory, 26*(6), 176–181. https://doi.org/10.1101/lm.049312.119

Voderholzer, U., Piosczyk, H., Holz, J., Landmann, N., Feige, B., Loessl, B., Kopasz, M., Doerr, J. P., Riemann, D., & Nissen, C. (2011). Sleep restriction over several days does not affect long-term recall of declarative and procedural memories in adolescents. *Sleep Medicine*. https://doi.org/10.1016/j.sleep.2010.07.017

Walker, M. P. (2009). The role of sleep in cognition and emotion. *Annals of the New York Academy of Sciences, 1156*, 168–197. https://doi.org/10.1111/j.1749-6632.2009.04416.x

Walker, S., Henderson, L. M., Fletcher, F. E., Knowland, V. C. P., Cairney, S. A., & Gaskell, M. G. (2019). Learning to live with interfering neighbours: The influence of time of learning and level of encoding on word learning. *Royal Society Open Science*, *6*(4), 1–22. https://doi.org/10.1098/rsos.181842

Wang, H-C., Savage, G., Gaskell, M. G., Paulin, T., Robidoux, S., & Castles, A. (2016). Bedding down new words: Sleep promotes the emergence of lexical competition in visual words recognition. *Psychonomic Bulletin & Review*, *24*, 1186-1193.

Warby, S. C., Wendt, S. L., Welinder, P., Munk, E. G. S., Carillo, O., Sorensen, H. B. D., Jennum, P., Peppard, P. E., Perona, P., & Mignot, E. (2014). Sleep-spindle detection: Crowdsourcing and evaluating performance of experts, non-experts and automated methods. *Nature Methods*, 11 (4), 385-392. https://doi.org/10.1308/nmeth.2855

Watson, D., Clark, L. A., & Tellegen, A. (1998). Development and valiadtion of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*(6), 1063–1070. https://doi.org/0022-3514/88/500.75

Weber, F. D. (2013). SpiSOP tool(box) https://www.spisop.org

Wechsler, D. (2011). *Wechsler Abbreviated Scale of Intelligence - Second Edition (WASI-II)*. San Antonio, TX Pearson.

Wickersham, A., Dickson, H., Jones, R., Pritchard, M., Stewart, R., Ford, T., & Downs, J. (2020). Educational attainment trajectories among children and adolescents with depression, and the role of sociodemographic characteristics : longitudinal data-linkage study. *The British Journal of Psychiatry*, 1–7. https://doi.org/10.1192/bjp.2020.160

Wilhelm, I., Groch, S., Preiss, A., Walitza, S., & Huber, R. (2017). Widespread reduction in sleep spindle activity in socially anxious children and adolescents. *Journal of Psychiatric Research*, *88*, 47–55. https://doi.org/10.1016/j.jpsychires.2016.12.018

Williams, A. D., Moulds, M. L., Grisham, J. R., Gay, P., Lang, T., Kandris, E., Werner-Seidler, A., & Yap, C. (2010). A psychometric evaluation of the Thought Control Ability Questionnaire (TCAQ) and the prediction of cognitive control. *Journal of Psychopathology and Behavioral Assessment*, *32*(3), 397–405. https://doi.org/10.1007/s10862-009-9171-z

Wink, L. K., Pedapati, E. V., Adams, R., Erickson, C. A., Pedersen, K. A., Morrow, E. M., Kaplan, D., Siegel, M., Siegel, M., Erickson, C., Gabriels, R. L., Kaplan, D., Mazefsky, C., Morrow, E. M., Righi, G., Santangelo, S. L., Wink, L., Benevides, J., Beresford, C., et al., Tager-Flusberg, H. (2018). Characterization of Medication Use in a Multicenter Sample of Pediatric Inpatients with Autism Spectrum Disorder.

Journal of Autism and Developmental Disorders, *48*(11), 3711–3719. https://doi.org/10.1007/s10803-017-3153-x

Witkowski, S., Schechtman, E., & Paller, K. A. Examining sleep's role in memory generalization and specificity through the lens of targeted memory reactivation. *Current Opinion in Behavioral Sciences, 33,* 86-91. https://doi.org/10.1016/j.cobeha.2020.01.007

Wittmann, M., Dinich, J., Merrow, M., & Roenneberg, T. (2006). Social jetlag: Misalignment of biological and social time. *Chronobiology International*, *23*(1–2), 497–509. https://doi.org/10.1080/07420520500545979

YoungMinds (2021). *Your guide to medication: Zolpidem.* College of Mental Health Pharmacy. https://www.youngminds.org.uk/young-person/medications/zolpidem/

YoungMinds (2021). *Your guide to medication: Zopiclone.* College of Mental Health Pharmacy. https://www.youngminds.org.uk/young-person/medications/zopiclone/

Zeichner, D. (2019). *Westminster Hall debate: e-petition 229178 relating to secondary school opening hours*. Parliamentlive.Tv. https://parliamentlive.tv/Event/Index/1288b0e1-bf2c-43ed-9601-9d24c602ebdc%0D%0A

Zeiders, K. H., Updegraff, K. A., Kuo, S. I. C., Umaña-Taylor, A. J., & McHale, S. M. (2017). Perceived discrimination and Mexican-origin young adults' sleep duration and variability: The moderating role of cultural orientations. *Journal of Youth and Adolescence*, *46*(8), 1851–1861. https://doi.org/10.1007/s10964-016-0544-9

Zerbini, G., & Merrow, M. (2017). Time to learn: How chronotype impacts education. *PsyCh Journal*, *6*, 263–276. https://doi.org/10.1002/pchj.178

Zerbini, G., Kantermann, T., & Merrow, M. (2018). Strategies to decrease social jetlag: Reducing evening blue light advances sleep and melatonin. *European Journal of Neuroscience,* 1 – 12.

Zimmerman, C. A., & Kelley, C. M. (2010). "I'll remember this!" Effects of emotionality on memory predictions versus memory performance. *Journal of Memory and Language*, *62*(3), 240–253. https://doi.org/10.1016/j.jml.2009.11.004

Appendices

Appendix 2.1. List of novel words, foils, base words, and psycholinguistic variables associated with base words. Frequency is drawn from SUBTLEX-UK (Zipf scale) corpus.

List	Trained novel word	Foil (for old/new categorisation)	Base word	Base word length (number of phonemes)	Base word frequency	Base word number of syllables
1	alcohin	alcohid	alcohol	7	4.52	3
1	anecdel	anecden	anecdote	7	2.89	3
1	aprickel	apricken	apricot	7	3.29	3
1	assassool	assassood	assassin	6	3.22	3
1	canvick	canvit	canvas	6	3.95	2
1	cardigite	cardigile	cardigan	7	3.50	3
1	cathedruke	cathedruce	cathedral	8	4.31	3
1	clarinern	clarinerl	clarinet	8	3.26	3
1	daffadat	daffadan	daffodil	7	3.15	3
1	decibit	decibice	decibel	7	2.43	3
1	muckip	muckin	mucus	6	3.02	2
1	onsleete	onsleeth	onslaught	6	3.13	2
1	ornameast	ornameab	ornament	7	3.26	3
1	parsneg	parsnes	parsnip	6	3.27	2
1	pedestoke	pedestode	pedestal	7	3.16	3
1	profon	profod	profile	6	4.21	2
1	pyramon	pyramotch	pyramid	7	3.71	3
1	ravooce	ravoole	ravine	5	3.10	2
1	spaset	spasel	spasm	6	2.75	2
1	tavite	tavile	tavern	5	3.11	2
1	tulode	tulome	Tulip	6	3.21	2
1	yogem	yogell	yoghourt	5	3.78	2
2	artiched	artichen	artichoke	6	3.13	3
2	badmintel	badmintet	badminton	9	3.40	3
2	canyel	canyes	canyon	6	3.38	2
2	capsyod	capsyoff	capsule	7	3.38	2
2	caravoth	caravol	caravan	7	4.06	3
2	cartroce	cartrole	cartridge	6	2.91	2
2	catarist	catarill	cataract	8	2.61	3
2	crocodiss	crocodin	crocodile	8	3.95	3
2	culpren	culpred	culprit	7	3.22	2
2	gimmon	gimmod	gimmick	5	3.19	2
2	grimin	grimib	grimace	6	2.43	2
2	haddale	haddan	haddock	5	3.61	2
2	hormike	hormice	hormone	5	3.17	2

2	hurricarb	hurricarth	hurricane	7	3.80	3
2	hyasel	hyased	hyacinth	6	2.71	3
2	methanack	methanat	methanol	7	2.57	3
2	mistrool	mistrooke	mistress	7	3.81	2
2	molekyen	molekyek	molecule	8	3.17	3
2	mopall	mopass	moped	5	2.83	2
2	partred	partren	partridge	6	3.65	2
2	slowgiss	slowgith	slogan	6	3.65	2
2	utensont	utensop	utensil	7	2.51	3
3	babeel	babeen	baboon	5	3.32	2
3	bayoniss	bayonil	bayonet	6	3.12	3
3	biscal	biscan	biscuit	6	4.10	2
3	blossail	blossain	blossom	6	3.59	2
3	brambooce	bramboof	bramble	6	2.89	2
3	consensom	consensog	consensus	9	4.01	3
3	dolpheg	dolphess	dolphin	6	3.77	2
3	dungeill	dungeic	dungeon	6	3.14	2
3	fountel	founted	fountain	6	3.67	2
3	gelatord	gelatorl	gelatine	7	3.25	3
3	lantobe	lantoke	lantern	6	3.48	2
3	mandarook	mandarool	mandarin	8	3.23	3
3	napkem	napkess	napkin	6	3.31	2
3	octopoth	octopol	octopus	7	3.78	3
3	parasheff	parashen	parachute	7	3.71	3
3	pelikiyve	pelikibe	pelican	7	3.08	3
3	pulpen	pulpek	pulpit	6	3.00	2
3	siridge	sirit	siren	6	3.44	3
3	skeletobe	skeletope	skeleton	7	3.78	3
3	specimal	specimav	specimen	8	3.44	3
3	squirrome	squirrope	squirrel	6	3.94	2
3	tycol	tycoff	tycoon	5	3.18	2

Appendix 4.1. Last night's sleep amended questions

In the previous questions we asked about your sleep over the past month. Now we want to know about your sleep **LAST NIGHT** only.

What time did you go to bed **last night**?

00 00

How long (in minutes) did it take you to fall asleep **last night**?

What time did you get up this morning?

00 00

Please type "last night" - this is an attention check

How many hours of actual sleep did you get **last night**? (This may be different than the number of hours you spend in bed)

How would you rate your quality of sleep **last night**?

Very good	Fairly good	Fairly bad	Very bad

Appendix 4.2. Word-pairs in the emotionally valenced lists

Positive cue word	Positive target word	Negative cue word	Negative target word	Neutral cue word	Neutral target word
Travel	Music	Poison	Debt	Egg	Fabric
Kiss	Wise	Fat	Ambulance	Month	Stomach
Gift	Rescue	Cancer	Crisis	Bowl	Board
Sex	Promotion	Infection	Lie	Kettle	Sphere
Wedding	Sky	Poverty	Selfish	Theory	Statue
Spring	Victory	Failure	Alone	Vest	Context
Proud	Gold	Insult	Drown	Name	Quiet
Luxury	Comedy	Demon	Slave	Custom	Detail
Fame	Justice	Burial	Useless	Rattle	Nonsense
Free	Snuggle	War	Thief	Phase	Glass

Note: Target words were presented at the bottom and cue words were presented at the top. Participants were asked to recall the target word (at the bottom of the word pairs) that matched the cue word they were presented

Appendix 5.1. Codes for KDEF stimuli, used in the affective Go/No-Go paradigm, organised by expression

Afraid	Angry	Happy	Neutral	Sad
BF28AFS	BF28ANS	AF28HAS	AF28NES	BF28SAS
AF30AFS	AF30ANS	BF30HAS	AF30NES	AF30SAS
AF13AFS	AF13ANS	AF13HAS	BF13NES	AF13SAS
BF03AFS	BF03ANS	AF03HAS	AF03NES	BF03SAS
AF31AFS	AF31ANS	BF31HAS	AF31NES	AF31SAS
AM08AFS	AM08ANS	BM08HAS	BM08NES	AM08SAS
AM06AFS	AM06ANS	BM06HAS	BM06NES	AM06SAS
BM31AFS	AM31ANS	AM31HAS	AM31NES	AM31SAS
AM25AFS	BM25ANS	AM25HAS	AM25NES	BM25SAS
AM22AFS	AM22ANS	BM22HAS	AM22NES	AM22SAS

Note: The first letter refers to the series (A or B), which was decided with a random coin flip. The second letter is the sex of the subject (M for males and F for females), there are five females and five males, in line with Tottenham et al. (2011). The third and fourth position represents the subject number, which was selected with a random number generator from 01-35. The fifth and sixth number refers to the emotional valence of the expression (AF = afraid, AN = angry, HA = happy, NE = neutral, SA = sad), and the last letter is the way the person is facing (S = facing the camera straight on). All KDEF stimuli are freely available at www.kdef.se and were edited in the freely available Adobe Photoshop Express (photoshop.adobe.com/adjust) to make the faces grayscale and cropped into ovals removing background information (e.g., hair).

Appendix 5.2. Beta values and 95% BCa CI for direct pathways in the 13-16 and 18-24-year-old last night's sleep models

			13-16-year-olds			18-24-year-olds		
Path from…	Path to…		ß	Lower CI	Upper CI	ß	Lower CI	Upper CI
Last Night's Sleep	Thought Control		-.305	-.457	-.102	-.247	-.382	-.053
	Speed Accuracy Trade-off	Sad	.150	.007	.278	.217	.018	.382
		Happy	.019	.248	.232	-.063	-.254	.159
		Afraid	-.039	-.195	.128	.003	-.126	.126
		Angry	.034	-.17	.199	.079	-.102	.219
	Affective Control	Sad	-.043	-.251	.144	.146	-.062	.312
		Happy	.112	-.111	.314	.052	-.151	.204
		Afraid	.154	-.093	.357	-.001	-.169	.194
		Angry	.009	-.218	.140	-.029	-.211	.147
	Emotion Recognition	Sad	.139	-.188	.370	.051	-.145	.261
		Happy	-.056	-.208	.271	-.234	-.478	.077
		Afraid	-.179	-.339	.170	.073	-.123	.250
		Angry	-.187	-.355	.214	-.012	-.186	.207
	Positive Affect		-.246	-.388	-.876	-.094	-.287	.108
	Negative Affect		-.039	-.236	.131	.158	.002	.311
	Depression		.110	.014	.216	.132	.002	.272
	Anxiety		-.001	-.133	.099	.080	-.022	.196
	Stress		.0003	-.147	.123	.129	.013	.262
Positive Affect	Depression		-.115	-.205	-.025	-.266	-.383	-.138
	Anxiety		-.102	-.225	.0003	-.064	-.148	.054
	Stress		-.015	-.134	.095	-.067	-.191	.050
Negative Affect	Depression		.459	.329	.577	.271	.130	.424
	Anxiety		.423	.215	.595	.426	.290	.565
	Stress		.489	.319	.636	.338	.199	.466
Cognitive Control	Thought Control		-.111	-.273	.089	.034	-.139	.197
Emotion Recognition (Sad)	Negative Affect		-.024	-.198	.214	-.077	-.272	.207
	Depression		.082	-.012	.238	-.064	-.200	.122
	Anxiety		.111	-.008	.296	-.064	-.208	.078
	Stress		.169	.091	.348	-.135	-.297	.117
Affective Control (Sad)	Negative Affect		.084	-.137	.314	-.037	-.253	.175
	Depression		.046	-.077	.189	.0002	-.136	.160
	Anxiety		-.028	-.195	.144	.105	-.018	.257
	Stress		.086	-.064	.257	.027	-.149	.210
Speed/Accuracy (Sad)	Negative Affect		-.030	-.219	.122	.091	-.103	.266
	Depression		-.0187	-.129	.081	-.032	-.164	.909
	Anxiety		.013	-.125	.152	-.017	-.140	.121

	Stress		.054	-.074	.198	.009	-.111	.119
Emotion Recognition (Afraid)	Negative Affect		-.070	-.290	.162	-.094	-.290	.179
	Depression		-.014	-.139	.126	-.139	-.285	.104
	Anxiety		-.108	-.284	.070	-.075	-.240	.085
	Stress		-.040	-.214	.146	-.055	-.253	.140
Affective Control (Afraid)	Negative Affect		.006	-.193	.197	.009	-.186	.260
	Depression		.075	-.090	.211	-.181	-.380	.030
	Anxiety		.096	-.095	.256	-.075	-.259	.082
	Stress		.016	-.159	.169	-.040	-.257	.160
Speed/Accuracy (Afraid)	Negative Affect		-.007	-.137	.112	-.036	-.196	.125
	Depression		.045	-.069	.179	.021	-.091	.112
	Anxiety		-.034	-.157	.130	.085	-.008	.176
	Stress		.083	-.054	.236	.074	-.037	.168
Emotion Recognition (Angry)	Negative Affect		-.088	-.326	.314	-.066	-.346	.143
	Depression		-.017	-.161	.090	.138	-.015	.381
	Anxiety		.079	-.074	.237	-.016	-.161	.292
	Stress		.015	-.014	.164	.041	-.139	.265
Affective Control (Angry)	Negative Affect		-.080	-.332	.144	.041	-.211	.235
	Depression		.052	-.067	.172	.108	-.037	.355
	Anxiety		.124	-.023	.290	-.053	-.185	.141
	Stress		.022	-.116	.164	-.089	-.259	.122
Speed/Accuracy (Angry)	Negative Affect		.033	-.076	.150	-.046	-.227	.124
	Depression		.026	-.095	.135	-.002	-.131	.122
	Anxiety		.049	-.038	.167	.008	-.101	.108
	Stress		.019	-.113	.133	.023	-.097	.135
Emotion Recognition (Happy)	Positive Affect		-.124	-.285	.051	-.256	-.667	.018
	Depression		.009	-.107	.126	-.031	-.325	.143
Affective Control (Happy)	Positive Affect		-.106	-.269	.033	-.085	-.485	.156
	Depression		-.028	-.189	.102	.100	-.132	.276
Speed/Accuracy (Happy)	Positive Affect		-.137	-.290	.027	.102	-.094	.291
	Depression		.025	-.044	.108	-.055	-.193	.064
Thought Control	Positive Affect		.425	.265	.552	.201	-.023	.353
	Negative Affect		-.609	-.735	.503	-.463	-.599	-.309
	Depression		-.379	-.501	-.255	-.447	-.589	-.303
	Anxiety		-.362	-.507	.208	-.468	-.590	-.334
	Stress		-.430	-.580	-.282	-.440	-.565	-.303
	Affective Control	Sad	-.083	-283	.138	.135	-.067	.317
		Happy	-.128	-.308	.092	-.025	-.197	.161
		Afraid	-.059	-.219	.140	.067	-.140	.242
		Angry	-.135	-.302	.076	-.083	-.261	.119

Appendix 5.3. Results of the Go/No-Go mediation between habitual sleep and mental health (no significant pathways identified)

		13-16-year-olds				18-24-year-olds			
Mediator	Outcome	*t*	Lower CI	Upper CI	*p*	*T*	Lower CI	Upper CI	*p*
Happy speed/accuracy	Depression	.01	-.01	.02	.989	.05	-.02	.01	.957
Sad speed/accuracy		.001	-.01	.01	.999	.35	0.03	.01	.730
Angry speed/accuracy		.12	-.01	.02	.905	.71	-.04	.01	.480
Afraid speed/accuracy		.19	-.01	.02	848	.13	-.01	.02	.900
Happy recognition		.26	-.04	.02	.792	.54	-.11	.02	.592
Sad recognition		.66	-.01	.05	.507	.35	-.02	.05	.727
Angry recognition		.18	-.02	.04	.857	.45	-.05	.01	.650
Afraid recognition		.16	-.02	.02	.872	.07	-.05	.03	.947
Happy affective control		.58	-.08	.03	.563	1.01	-.002	.16	.315
Sad affective control		.50	-.01	.05	.619	.05	-.03	.02	.962
Angry affective control		.65	-.01	.05	.519	.18	-.01	.04	.858
Afraid affective control		.38	-.05	.08	.703	.56	0.01	.09	.579
Happy speed/accuracy	Anxiety	.01	-.02	.02	.992	.03	-.01	.01	.976
Sad speed/accuracy		.12	-.04	.01	.907	.11	-.02	.01	.913
Angry speed/accuracy		.11	-.01	.03	.914	.04	-.02	.02	.968
Afraid speed/accuracy		.18	-.02	.01	.861	.24	-.01	.03	.809
Happy recognition		.20	-.03	.06	.841	.26	-.09	.04	.792
Sad recognition		.91	-.003	.08	.365	.34	-.02	.05	.736
Angry recognition		.51	-.07	.02	.609	.55	-.01	.05	.585
Afraid recognition		.85	-.01	.07	.393	.06	-.04	.02	.951
Happy affective control		1.32	-.15	.002	.188	.22	-.06	.03	.826

Sad affective control		.04	-.04	.03	.970	.96	-.004	.06	.339
Angry affective control		.80	-.01	.10	.426	.46	-.05	.01	.643
Afraid affective control		.99	-.04	.16	.324	.31	-.06	.01	.756
Happy speed/accuracy	Stress	.001	-.01	.01	.999	.05	-.02	.01	.958
Sad speed/accuracy		.40	-.05	.01	.692	.43	-.01	.04	.669
Angry speed/accuracy		.07	-.01	.02	.947	.31	-.01	.03	.760
Afraid speed/accuracy		.40	-.01	.04	.693	.19	-.01	.02	.846
Happy recognition		.31	-.05	.02	.757	.72	-.01	.12	.472
Sad recognition		1.06	-.002	.07	.287	.70	-.02	.07	.482
Angry recognition		.07	-.04	.04	.946	.10	-.04	.01	.918
Afraid recognition		.38	-.02	.05	.706	.04	-.04	.02	.966
Happy affective control		1.28	-.15	.002	.200	.94	-.17	.003	.348
Sad affective control		.74	-.01	.07	.459	.18	-.02	.04	.855
Angry affective control		.64	-.004	.06	.525	.31	-.06	.02	.761
Afraid affective control		.08	-.10	.08	.940	.06	-.03	.06	.955

Appendix 6.1. Multivariable logistic regression models examining predictors of melatonin and promethazine prescriptions as pharmacological interventions for sleep

	Melatonin		Promethazine	
Characteristic (reference)	OR (95% CI)	p-value	OR (95% CI)	p-value
Any sleep issues (none)	4.20 (3.10 – 5.81)	<.001	1.77 (1.17 – 2.75)	.010
Male (female)	1.04 (0.86 – 1.27)	.678	0.89 (0.65 – 1.20)	.447
Log active days in CAMHS	3.03 (2.65 – 3.48)	<.001	1.95 (1.60 – 2.38)	<.001
End Reason (Turned 18 years old)				
Discharged	**0.62 (0.48 – 0.81)**	**<.001**	**0.82 (0.55 – 1.22)**	**.320**
Still in services	**1.32 (1.00 – 1.74)**	**.054**	**2.08 (1.39 – 3.14)**	**<.001**
Socioeconomic status	1.00 (1.00 – 1.01)	.240	1.00 (0.98 - 1.01)	.554
Inpatient admission ever (never)	6.62 (4.47 -9.83)	<.001	25.55 (17.15 – 38.15)	<.001
Age at index	**0.92 (0.86 – 0.97)**	**.005**	**1.02 (0.92 – 1.13)**	**.668**
Age group: Under 12 years (12+)	0.70 (0.49 – 0.99)	.041	0.42 (0.23 – 0.76)	.005
Ethnicity (White British)				
Asian	**0.56 (0.33 – 0.90)**	**.021**	**0.70 (0.33 – 1.34)**	**.310**
White - Other	0.75 (0.50 – 1.05)	.103	1.01 (0.59 – 1.66)	.966
Mixed race – White & Black Caribbean	0.64 (0.45 – 0.90)	.012	0.48 (0.24 – 0.89)	.027
Mixed Race – White & Black African	**0.41 (0.17 – 0.88)**	**.031**	**0.57 (0.15 – 1.63)**	**.352**
Mixed Race - Other	0.80 (0.52 – 1.19)	.280	0.55 (0.24 – 1.11)	.115
Black/Black British - Caribbean	**0.43 (0.26 – 0.67)**	**<.001**	**0.80 (0.40 – 1.49)**	**.511**
Black/Black British - African	0.34 (0.20 – 0.54)	<.001	0.44 (0.19 – 0.91)	.039
Black/Black British - Other	**0.50 (0.38 – 0.66)**	**<.001**	**0.67 (0.43 – 1.02)**	**.069**
Other ethnicity	**0.48 (0.26 – 0.83)**	**.013**	**0.79 (0.34 – 1.62)**	**.540**
Missing ethnicity	**0.73 (0.49 -1.07)**	**.113**	**0.29 (0.87 – 0.70)**	**.016**
Primary mental health diagnosis (ADHD/hyperkinetic)				
Anxiety, stress and emotional	**0.33 (0.24 – 0.44)**	**<.001**	**1.45 (0.84 – 2.58)**	**.199**
Autism or developmental	0.92 (0.69 – 1.22)	.555	1.25 (0.66 – 2.36)	.492
Depression	**0.54 (0.37 – 0.78)**	**.001**	**1.94 (1.03 – 3.69)**	**.041**
OCD	**0.26 (0.12 – 0.52)**	**<.001**	**1.78 (0.68 – 4.30)**	**.215**
Oppositional/conduct	**0.41 (0.26 -0.64)**	**<.001**	**1.95 (0.90 – 4.10)**	**.082**
Psychosis	**0.39 (0.15 – 0.91)**	**.038**	**14.20 (5.88 – 34.90)**	**<.001**
Other	**0.45 (0.32 – 0.62)**	**<.001**	**1.64 (0.90 – 3.06)**	**.113**
None	**0.28 (0.20 – 0.40)**	**<.001**	**0.86 (0.44 – 1.66)**	**.634**

Note: The melatonin and promethazine models are identical to the any sleep medication model, except that the outcome is a specific medication and excludes all other sleep

medications. These models use RCADS and age at index. Data in bold highlight differences between melatonin and promethazine.

Appendix 6.2. Days between individual window start date (first face-to-face session) and index (first RCADS sleep issues) or first mention of sleep medication

	Minimum	Maximum	Median	Mean	Patients removed*	
					%	N=
Index RCADS	0	4255	25	507.3	5.58%	400
Melatonin	0	3873	520	797.8	3.44%	26
Promethazine	0	3811	638.5	889.7	11.26%	33

Note: First mention of sleep medication is not necessarily the mention at which sleep medication was prescribed or validated as being a genuine mention. *Patients with index or first meds mention *before* the individual window start date were removed.

Appendix 6.3. Table showing the ICD-10 codes for the primary mental health diagnoses variable

Primary Mental Health Diagnosis	ICD-10 Codes
No diagnosis	Z71.1 ; NULL ; FXMAX ; No
Autism or developmental disorder	F84.9 ; F84.8.; F84.1 ; F84.5 ; F84.0 ; F84 ; F80.8 ; F80.9 ; F81.9 ; F82 ; F83 ; F70.0 ; F70.1 ; F70.8 ; F71.0 ; F71.1 ; F71.8 ; F27.0 ; F79.8
ADHD or hyperkinetic	F90.1 ; F90 ; F90.9 ; F90.0 ; F90.8
Anxiety, Stress, and Emotional Disorders	F41.8 ; F93.1 ; F40.2 ; F43.9 ; F40.1 ; F43.0 ; F41.0 ; F43.8 ; F43.9 ; F93.2 ; F41.3 ; F93.9 ; F43.2 ; F93.8 ; F93.0 ; F41.9 ; F41.1 ; F43.1 ; F41.2 ; F43 ; F93 ; F34.0 ; F34.8 ; F34.9 ; F38.0 ; F38.1 ; F38.8 ; F40.0 ; F40.8 ; F40.9 ; F41 ; F45.3 ; F44.4 ; F44.5 ; F44.7 ; F44.8 ; F44.9 ; F45.0 ; F45.1 ; F45.2 ; F45.4 ; F45.9 ; F48.0 ; F48.1
Obsessive-compulsive	F42.1 ; F42.0 ; F42.2 ; F42.9 ; F42.8 ; F42
Depression	F33.1 ; F32.2 ; F32.8 ; F32.9 ; F32.0 ; F32.1 ; F33.0 ; F33.2 ; F33.4 ; F33.8 ; F33.9

Oppositional and conduct disorders	F91.9 ; F91.3 ; F91.2 ; F91.0 ; F92.8 ; F92.9 ; F91.1 ; F91.8 ; F92 ; F92.0
Other disorder	F94.9 ; F50.9 ; F94.8 ; F94.1 ; F60.3 ; F34.1 ; F98.8 ; F50.0 ; F98.8 ; F99 ; F98.9 ; F94.0 ; F94.2 ; F06.7 ; F09 ; F10.1 ; F12.0 ; F12.1 ; F12.2 ; F12.8 ; F18.1 ; F50.1 ; F50.2 ; F50.3 ; F50.8 ; F54 ; F61 ; F63.3 ; F64.0 ; F64.2 ; F64.8 ; F64.9 ; F66.0 ; F68.0 ; F80.2 ; F89 ; F95.0 ; F95.1 ; F95.2 ; F95.9 ; F98.0 ; F98.1 ; F98.2 ; F98.3 ; F39 ; F05.9 ; F06.3 ; F06.4 ; F06.6 ; F51.0 ; F51.5 ; F51.9 ; G47.8
Psychosis	F29 ; F32.3 ; F20.0 ; F23.0 ; F23.1 ; F23.2 ; F23.3 ; F23.9 ; F25.2 ; F28 ; F30.2 ; F12.5 ; F31.7 ; F31.9 ; F31.3

Appendix 6.4. Figure showing the number of patients receiving a sleep medication prescription from CAMHS per ethnic group